COMMUNAUTÉ D'UNIVERSITÉS ET ÉTABLISSEMENTS
l'université nantes angers le mans

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Christophe A. N. Biscio

Mémoire présenté en vue de l'obtention du
grade de Docteur de l'Université de Nantes
Label européen
sous le label de l'Université de Nantes Angers Le Mans

École doctorale : Sciences et technologies de l'information, et mathématiques

Discipline : Mathématiques et leurs interactions, section CNU 26
Unité de recherche : Laboratoire de Mathématiques Jean Leray (LMJL)

Soutenue le 24 septembre 2015

Contribution to the modelling and the parametric estimation of determinantal point processes

JURY

Président :	**M. Bernard Delyon**, Professeur, Université de Rennes I
Rapporteur :	**M^{me} Patricia Reynaud-Bouret**, Directeur de recherche CNRS, Université de Nice
Examinateurs :	**M. Jean-François Coeurjolly**, Maître de conférences, Université de Grenoble **M^{me} Anne Philippe**, Professeur, Université de Nantes
Directeur de thèse :	**M. Frédéric Lavancier**, Maître de conférences, Université de Nantes

* * *

Rapporteur :	**M. Rasmus Waagepetersen**, Professeur, Université de Aalborg

Contents

Remerciements

Passage obligé et également mon préféré, voilà le moment des remerciements, moment de liberté où ma grammaire approximative et mon expression chaotique trouveront leur place sans contrainte. Cette thèse fut pour moi un accomplissement mais surtout une fierté. À ce titre je remercie avant toute chose et en premier lieu mes grands parents, Ange Napoléon, Onéglia, Christophe et Marie sans qui rien de cela n'aurait été possible. C'est par leurs encouragements et leur dévouement à leur famille en ces périodes difficiles d'immigration que je leur dois mon parcours. J'espère par cette thèse rattraper les diplômes que Christophe aurait passés si il n'avait pas du travailler le bois pour nourrir sa famille. Cette thèse ainsi que tous mes travaux leurs sont et leurs seront toujours dédiés. Je remercie également mon cher "paire" qui m'a appris l'importance de la grosseur des tiers ainsi que Hubert et Claude pour leur soutien continu durant mes années d'études.

Je remercie Bernard Delyon, Anne Philippe et Jean-François Coeurjolly pour avoir accepté d'être jury de ma thèse. Merci également à Ronan Le Guével pour avoir suivi mes progrès tout au long de la thèse. Je remercie Rasmus Waagepetersen qui je pense partage mes aspirations pour des statistiques propres et rigoureusement prouvées. Suivre ses traces même de loin caché derrière une forêt tropicale, au position d'arbres inhomogène ou non, serait pour moi un honneur et je le remercie d'avoir accepté de rapporter ma thèse. Je remercie également Patricia Reynaud-Bouret en tant que rapporteuse pointue de ma thèse mais surtout comme enseignante et véritable muse des statistiques. Je me souviens du temps où sur les bancs de l'université de Nice, Anis, Gilles et moi-même suivions assidûment son cours et furent tous trois aspirés dans les statistiques, m'éloignant ainsi de la topologie algébrique, ha que De Rham va me manquer.

Je remercie toute l'équipe du laboratoire de mathématiques Jean Leray pour leur disponibilité et leur soutien, que cela soit pour des ordres de mission ou pour des demandes de mise en disponibilité. Merci à Claude pour ces bons moments à la bibliothèque et pour son expo.

Outre les statisticiens, cette thèse doit beaucoup aux commentaires, avis et schémas de preuves donnés par des gens divers et variés. Je remercie Silvia Serfaty pour m'avoir guidé vers son papier et un livre d'analyse qui ont débloqué de nombreux points dans ma thèse. Je remercie Jean François Coulombel pour ne pas s'être jeté par la fenêtre à la vue de $\mathcal{F}(C) \leq 1$ et pour m'avoir ensuite illuminé sur une démonstration importante qui dépassait mes compétences de l'époque. Bien que j'aurais aimé découvrir l'existence de la vérité absolue, il m'a offert l'existence de mon minimum sous des contraintes peu banales et je l'en remercie. Sur ce dernier point je remercie également les doctorants d'analyse, Antoine Benoit, Virgile Robbe et Damien Gobin, qui du rez-de-chaussé auraient pu se jeter par la fenêtre sans trop de risque mais qui ont su m'écouter et m'aider. Je remercie Guillaume Roux pour m'avoir fait découvrir la distance de Haussdorf et m'avoir montré le chemin au travers d'un labyrinthe de compacts aléatoires convexes. Je remercie Gilles Carron, dit Gilou l'avocat dans certains milieux lyonnais, pour m'avoir fait découvrir le livre de Taylor, mine incroyable d'informations sur les edps. Je remercie Samuel Tapie pour ces références sur les valeurs propres qui m'ont aidé à prendre du recul et je remercie finalement Philippe Carmona pour son aide sur les mesures symétriques.

La thèse fut une épreuve mais surtout une aventure riche en événements, amis et ennemis ainsi que beaucoup de fiers et fidèles compagnons qui m'ont aidé à aller jusqu'au bout. Je remercie évidement les doctorants du laboratoire de mathématiques Jean Leray qui ont sut me faire adorer ces trois années, supporter ma folie, et me faire amèrement regretter mon départ. Ces trois années furent les plus riches de ma vie en partie grâce à eux et à leur amitié qui dépassait le cadre du laboratoire, ils ne pourront que me manquer. Je remercie Céline, ma confidente et amie pour toutes ces conversations, ses bruits de vélociraptor et les musiques écoutées ensemble en cette première année de thèse. J'en profite au passage pour remercier Doudou et Lulu pour leur soutien et leurs sourires. Je remercie Virgile et Hélène pour leur accueil et leur amitié ainsi que pour la belle et mignonne Marinette, arrivée plus tard mais pour qui j'ai craqué au premier regard. À elle qui a su me tenir en éveil par ses nombreux pleurs et su m'illuminer par ses rires, merci. Merci aussi à Victor et Caroline pour leur amitié sincère ainsi que leurs rires et délires partagés. Je remercie Vytaute, ma seule camarade statisticienne dans cette enfer d'analystes et d'algébristes, les discussions avec toi lors de tes trop courts passages sur Nantes ont été enrichissantes même si le sujet était, tel tes champs aléatoires, limite et partait dans tous les sens et pas toujours de la même façon. Je remercie Antoine pour avoir supporté toutes mes emportées lyriques mentales (folie) et pour avoir accepté tous mes câlins et questions sur les edps sans jamais avoir porté plainte (pour les câlins), puisse la Belgique te rendre heureux et t'épanouir personnellement et scientifiquement.

Je remercie également tous les autres thésards de ces trois dernières années, Ilaria l'italienne au caractère farouche mais qui a su se montrer délicate dans les moments les plus importants, Gilberto le moins italien mais plus Sarde, Olivier le claqueur de langue, Guillaume le jeune (comme Louis VII ou Philippe), Vincent le marseillais aux drops du second tour, Salim le bricoleur au cœur sur la main, Carlos le multilingue, Thomas G. le cannibal pour ne pas m'avoir mangé, Thomas B. le vegan pour ne pas m'avoir mangé le poireau, Thomas W. le musclé, Noémie la buveuse de thé, Florian l'ingénieur pour ses nombreux dépannages informatiques, Pierre à qui on lui a dit pour lui faire plaisir et pour m'avoir fait découvrir l'existence de lieux nous permettant de mieux se connaître intérieurement, Valentin qui cache bien son jeu, Damien aux beaux yeux bruns, au visage mignon et à la voix douce et suave (rrr), Victor le portos au slip violet qui va bien en fond d'écran, Vivien le champion de Trivial Pursuit et Radek le barman aux multiples bières.

Une thèse ne se ferait pas sans conférences et conférenciers. À ce titre je remercie toutes les personnes que j'ai rencontrées en ces occasions pour avoir ouvert mon esprit à de nouveaux aspects des statistiques mais aussi pour m'avoir fait découvrir de nouvelles langues et cultures. Merci à tous les jeunes conférenciers d'Aalborg pour avoir trompé ma solitude doctorale statisticienne nantaise et à tous les "vieux" de cette école d'été pour leurs encouragements et leurs soutiens parfois à coup de chope de 1L, merci Peter. Je remercie en particulier Ottman pour ses encouragements et ses idées ingénieuses. Merci à Henrike pour tous ces moments de rire en conférence. Merci à ceux que j'ai rencontré à Avignon, fussent ils à Aalborg ou non, pour une conférence qui m'a apporté un regard neuf sur les statistiques. J'aurais pu, avec eux, arpenter les rues de la cité des papes toute la nuit.

Merci à tous ceux du GDR Geosto pour des découvertes telles que le fait que 3^3 est le seul

âge en n^n dont on aura conscience, merci Jeff, mais aussi merci à eux pour leurs exposés divers et variés. Je remercie en particulier Marianne pour ses nombreux encouragements qui ont beaucoup compté pour moi. Ma gratitude va également à Pierre pour ses remarques et conseils, Nicolas pour ses encouragements et son intérêt, Hermine pour sa gentillesse et sa bienveillance et Julien pour une partie de belote des plus malchanceuses qui soit, ainsi que pour plein d'infos et potins dont je suis friand sur son camarade lyonnais.

Je remercie les doctorants et seniors des RJS pour leur accueil et leurs nombreux surnoms: Matthieu, Florian, Guillaume, Antoine, Jérôme et Thomas. Merci aux biostaticiens qui me font regretter de ne pas améliorer la détection d'Alzeihmer en France, merci de m'avoir fait découvrir la géométrie d'Aitchison et merci aussi de m'envoyer autant de Lasso. Merci pour cette conférence, pour cette balade en canoë et pour cette traversée difficile de la dune du Pyla. Merci Johan pour tes conseils et remarques sur le Danemark.

Je remercie (décidément) les enseignants qui m'ont fait apprécier les mathématiques tout au long de mon cursus. Je pense en particulier à Joseph Ventura, professeur de classe préparatoire au lycée Dumont D'Urville mais aussi entre autres à Yannick Baraud, Benoit Cadre, Mihai Gradinaru et Jean Christophe Breton.

Je remercie mes élèves pour avoir supporté sans doute l'un de leurs professeur le plus fou et m'avoir fait découvrir et aimer les joies de l'enseignement en fac. Vous avoir enseigné les statistiques et les probabilités durant ces 3 années fut l'une des expériences les plus enrichissantes de ma vie. Merci en particulier à Marie Claire qui a su montrer qu'une bonne élève se cache là où personne ne l'attend plus, merci à Mathieu pour ses raisonnements brillants, merci à Hélène pour avoir été ma première étudiante agrégée et merci à Sarah qui selon moi aura accès à tout ce qu'elle souhaitera dans sa vie professionnelle, ton potentiel et tes capacités sont pour moi la plus grande des motivations pour devenir enseignant. Enfin merci à mes tous premiers L2 dont coco et sa bande de joyeux lurons, qui cette année je l'espère deviendront agrégés à leurs tour.

L'heure et maintenant venue des remerciements spéciaux pour gens spéciaux ou pour un simple collègue.

Je remercie Julie sans qui cette thèse n'aurait jamais eu lieu. Merci pour tout tes conseils et encouragements. Bien qu'initialement ma camarade tu es devenue pour moi une mentor et un exemple à suivre, merci.

Je remercie ma camarade de galère thèsarienne et amie, Émilie avec qui j'ai combattu envers et contre tout dans les tranchées de la recherche mathématiques, bien qu'accompagnés de Death Cab for Cutie au lieu des Valkyries de Wagner, je n'en reste pas moins grandi par ces moments, merci pour tout ma chère amie.

Je remercie mes amis qui ont toujours su être là pour moi, merci à Angelina et Vincent, merci à Lisa et Romain et merci à Maman Eugénie et Pierre. Pour votre existence et présence, je vous remercie de tout mon cœur et je m'excuse de ne pas avoir pris plus le temps de venir vous voir.

Merci à toi mon seigneur et maître, merci à toi chef, merci à toi patron (alors que me viens la chanson des charlots en tête), merci à toi Frédéric pour m'avoir encadré, accompagné, encouragé et ouvert à des idées, des notions et un monde que je ne voyais que de loin. Tu as supporté tant bien que mal ma grammaire tristement légendaire pour toujours me permettre de m'améliorer. Encore une fois, merci pour ton dévouement à ta fonction de directeur de thèse. Tu as été un encadrant remarquable aussi bien par ta pédagogie, ta sympathie que par tes idées et je ne

pourrais que faire des éloges à ton sujet dans l'avenir. J'ai d'ailleurs pris un peu d'avance dans certaines conférences pour ça.

En fredonnant du Maxime Le Forestier, je tiens à remercier Marcel Boujna pour son soutien, merci à toi e-mood pour avoir supporté mes délires dans une colocation épique. Nous fûmes réellement un duo de choc, je repense à notre première pizza coupée au tournevis et mangée sur un meuble de cuisine alors que nous étions entourés de carton. Je repense à nos galères, missions en voiture, bouffes, soirées télé, conversations nocturnes sur un peu tous les sujets, discussions mathématiques et complaintes des étudiants etc. Encore merci Moody pour ne pas avoir eu peur quand je parlais tout seul, quand je frappais les murs, quand je chantais, quand je râlais pour un rien devant un ami imaginaire bref merci d'être resté l'ami fidèle que tu es. Pour tous ces moments de soutien dans la thèse et dans la vie, je te remercie Moudhaffar.

Finalement, merci à toi mia bella piccola pigna baffuta di cozza con un colto coltello qui est quelques fois appelée Silvia, pour m'avoir guidé dans les rues de Lisbonne et pour tous ces moments où tu m'as rendu heureux dans une vie que je considérais bien trop triste, en me faisant découvrir qu'on pouvait avancer à deux.

En conclusion, à vous tous qui lisez ces lignes et m'avait connu et/ou accompagné durant ces 3 années, merci. Même si je n'ai pas mis votre nom dans ces lignes, par paresse d'écriture ou par oubli, votre rencontre m'a permis d'avancer durant ces années et du plus profond de moi-même, je vous en remercie.

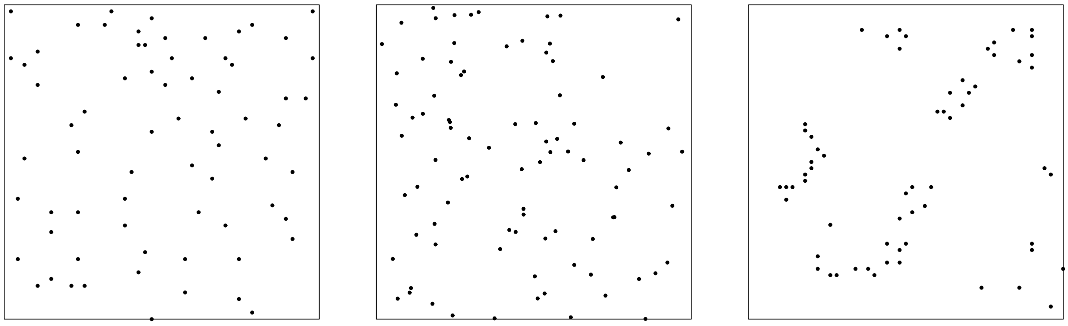
1

Introduction (version française)

1.1 Contexte

De nombreux domaines scientifiques sont confrontés à l'observation de configurations de points, comme par exemple la position d'arbres en Sylviculture, la position de patients malades en Épidémiologie ou la position de particules en Mécanique Quantique. Beaucoup d'autres exemples peuvent être trouvés dans la littérature, voir entre autres [33]. Les processus ponctuels spatiaux sont les outils mathématiques utilisés pour modéliser ces configurations de points. Ils peuvent être vus comme des points aléatoirement distribués dans l'espace, généralement un compact de $\mathbb{R}^2$. La statistique spatiale est la branche des mathématiques qui cherche à décrire, modéliser et inférer le comportement de ces processus. En particulier, la dépendance entre les positions des différents points est de première importance.

Beaucoup de modèles ont déjà été introduits pour représenter une grande variété de configurations de points. Dans ce manuscrit, nous nous limitons à l'étude de processus stationnaires, pour lesquels la répartition des points est supposée invariante par translation. L'exemple le plus simple est le **processus ponctuel de Poisson**. Un processus ponctuel $\mathbf{X}$ est un processus ponctuel de Poisson sur $\mathbb{R}^d$ d'intensité ρ si, sur tout borné B de $\mathbb{R}^d$, le nombre de points $\mathbf{X}(B)$ dans B suit une loi de Poisson de paramètre $\rho|B|$, où $|B|$ est le volume de B, et étant donné $\mathbf{X}(B) = n$, les n points sont répartis indépendamment et uniformément dans B. Si $\rho = 1$, le processus de Poisson est dit standard. Ces processus modélisent des situations où aucune interaction n'est observée et sont bien évidement trop restrictifs pour représenter n'importe quelle configuration. En fait, les processus ponctuels exhibent souvent de la répulsion ou de l'attraction comme par exemple des particules en physique statistique qui se repousseraient entre elles, ou au contraire s'attireraient pour former des amas de points. En pratique, les processus ponctuels sont séparés en trois catégories, voir par exemple [33, Section 1] ou [3] : réguliers (ou répulsifs, ou inhibitifs) si les points ne sont pas trop proches les uns des autres, aléatoires s'il n'y a aucune interaction entre eux et attractifs s'il y a formation de clusters. Cette classification est illustrée sur la Figure 1.1, où de gauche à droite les configurations de points sont respectivement régulière, aléatoire et attractive. Cette classification est arbitraire et ne reflète pas la complexité de tous les jeux de données mais reste cependant suffisante pour un grand nombre d'applications.



(a) (b) (c)

Figure 1.1 – Trois exemples de configurations de points : (a) Positions de 69 villes espagnoles dans une région de 40 miles par 40 miles, (b) réalisation sur $[0,1]^2$ d'un processus de Poisson stationnaire d'intensité $\rho = 100$, (c) Positions de séquoias dans une forêt. Les jeux de données représentés en (a) et (c) sont respectivement disponibles dans les packages `R` `spatial` et `spatstat`.

Nous avons déjà mentionné le processus ponctuel de Poisson, utilisé pour modéliser des situations où il n'y a aucune interaction entre les points. Pour modéliser de l'attraction, les **processus de Cox**, aussi appelés processus de Poisson doublement stochastiques, sont généralement utilisés. Ces processus sont définis à partir d'un processus ponctuel de Poisson inhomogène, c'est à dire un processus ponctuel de Poisson où l'intensité ρ est une fonction de $\mathbb{R}^d$ dans $\mathbb{R}^+$. Un processus ponctuel $\mathbf{X}$ sur $\mathbb{R}^d$ est un processus de Cox stationnaire dirigé par l'intensité aléatoire stationnaire Z si, conditionnellement à $Z(x) = \rho(x)$, $\mathbf{X}$ est un processus de Poisson d'intensité $\rho(x)$. Cela comprend par exemple les processus de Neyman-Scott et les processus ponctuels à clusters, voir [49]. Parallèlement, il existe plusieurs modèles de processus ponctuels pour représenter de la répulsion. Nous présentons les plus usuels en Section 1.2 et notamment la classe importante des **processus ponctuels de Gibbs**. Récemment, une autre famille, les **processus ponctuels déterminantaux**, a été considérée pour modéliser des configurations de points répulsives. Ces modèles s'avèrent être une alternative intéressante aux processus de Gibbs comme cela a été illustré sur différentes applications dont l'exemple de la Figure 2.1 (a), voir [43]. Ce manuscrit est dédié à l'étude des DPPs.

1.2 Processus ponctuels répulsifs et processus ponctuels déterminantaux

1.2.1 Préliminaires sur les processus ponctuels

Pour $d \geq 1$, nous munissons $\mathbb{R}^d$ de la tribu $\mathcal{B}(\mathbb{R}^d)$ générée par les Boréliens et notons $\mathcal{B}_0(\mathbb{R}^d)$ l'ensemble de tous les Boréliens bornés de $\mathbb{R}^d$. Nous notons $\mathbf{x} := \{x_1, , x_2, \dots, x_n, \dots\}$ une configuration arbitraire de points dans $\mathbb{R}^d$ et $\mathbf{x}(B)$ le nombre de points de $\mathbf{x}$ qui tombent dans B pour un B arbitraire dans $\mathcal{B}_0(\mathbb{R}^d)$. Soit $\mathcal{N} := \{\mathbf{x} \subset \mathbb{R}^d,\ \mathbf{x}(B) < \infty,\ \forall B \in \mathcal{B}_0(\mathbb{R}^d)\}$

l'espace des configurations de points dans $\mathbb{R}^d$ localement finies. Nous munissons $\mathcal{N}$ de la tribu engendrée par les ensembles $\{\mathbf{x} \subset \mathbb{R}^d,\ \mathbf{x}(B) = n\}$ pour tout $B \in \mathcal{B}_0(\mathbb{R}^d)$ et tout $n \in \mathbb{N} \cup \{0\}$ où $\mathbb{N}$ est l'ensemble des entiers strictement positifs. Un processus ponctuel sur $\mathbb{R}^d$ est une application mesurable d'un espace probabilisé $(\Omega, \mathcal{F}, P)$ dans l'ensemble $\mathcal{N}$. Nous écrirons un processus ponctuel par une lettre en gras, habituellement $\mathbf{X}$, et nous identifions l'application $\mathbf{X}$ et l'ensemble aléatoire de points associé. Pour indiquer qu'un point x appartient à $\mathbf{X}$, nous écrirons $x \in \mathbf{X}$. De plus, nous supposerons toujours que les processus ponctuels sont simples, i.e. presque sûrement, deux points d'un processus ne coïncident jamais. Pour une introduction détaillée aux processus ponctuels, nous référons à [9, 10].

Les moments factoriels, définis ci-dessous, figurent parmi les quantités importantes pour l'étude des processus ponctuels.

Définition 1.2.1. *La mesure moment factorielle d'ordre k ($k \geq 1$) d'un processus ponctuel simple $\mathbf{X}$ est la mesure sur $\mathbb{R}^{dk}$, notée $\alpha^{(k)}$, telle que pour toute famille de sous-ensembles $D_1, \ldots, D_k$ de $\mathbb{R}^d$,*

$$\alpha^{(k)}\left(D_1 \times \ldots \times D_k\right) = \mathbb{E}\left(\sum_{(x_1,\ldots,x_k)\in\mathbf{X}^k}^{\neq} \mathbf{1}_{\{x_1\in D_1,\ldots,x_k\in D_k\}}\right)$$

où $\mathbb{E}$ désigne l'espérance par rapport à la distribution de $\mathbf{X}$ et le symbole $\neq$ au-dessus du signe somme indique que nous ne considérons que les k-uplets $x_1, \ldots, x_k$ mutuellement disjoints.

Les cas les plus importants sont la mesure moment factorielle d'ordre un, appelée la mesure d'intensité, et la mesure moment factorielle d'ordre deux. Sous notre hypothèse de stationnarité, la mesure d'intensité est proportionnelle à la mesure de Lebesgue sur $\mathbb{R}^d$. Ceci signifie qu'il existe un $\rho > 0$ tel que pour tout $S \subset \mathbb{R}^d$ nous avons $\alpha^{(1)}(S) = \rho|S|$. Dans ce cas, ρ est appelé l'**intensité** et représente le nombre moyen de points par unité de volume. Remarquons aussi que ρ est la densité de la mesure d'intensité par rapport la mesure de Lebesgue. Si $\alpha^{(2)}$ admet une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}^{2d}$, cette densité est notée $\rho^{(2)}$ et est appelée **la fonction d'intensité d'ordre deux**. Intuitivement, pour $(x, y) \in \mathbb{R}^{2d}$ et $x \neq y$, $\rho^{(2)}(x, y)$ peut être vu comme la probabilité qu'un point du processus se trouve dans un petit voisinage autour de x et qu'un autre point se trouve dans un petit voisinage autour de y. Remarquons que sous notre hypothèse de stationnarité, $\rho^{(2)}(x, y) = \rho^{(2)}(0, y - x)$ est une fonction symétrique et dépend seulement de $y - x$.

En statistique spatiale, les propriétés de second ordre d'un processus ponctuel sont souvent étudiées au travers de statistiques telles que la **fonction de corrélation par paires** (pcf) ou la **fonction de Ripley**, définies comme suit. La pcf est définie, pour presque tout $x \in \mathbb{R}^d$, par

$$g(x) = \frac{\rho^{(2)}(0, x)}{\rho^2},$$

et pour tout $t \geq 0$, la fonction de Ripley K est définie par

$$K(t) = \int_{B(0,t)} g(x)dx,$$

où $B(0, t)$ est la boule Euclidienne ouverte de centre 0 et de rayon t.

Intuitivement, $g(x)$ est le rapport entre la probabilité que deux points du processus se trouvent séparé de x (en tenant compte de l'interaction entre les points induite par le processus) et cette même probabilité pour un processus de Poisson (situation sans interaction). Il s'ensuit que pour tout $x \in \mathbb{R}^d$, une interprétation usuelle, voir [59], est que $g(x) > 1$ caractérise de l'attraction tandis que $g(x) < 1$ caractérise de la répulsion. Concernant la fonction de Ripley K, $\rho K(t)$ s'interprète comme le nombre moyen de points dans $B(0, t)$ sachant que 0 est un point du processus, voir [49] pour des détails supplémentaires.

Par exemple, pour le processus ponctuel de Poisson stationnaire sur $\mathbb{R}^d$ avec intensité $\rho > 0$, la fonction d'intensité d'ordre deux existe et pour tout $(x, y) \in \mathbb{R}^{2d}$ tel que $x \neq y$, on a $\rho^{(2)}(x, y) = \rho^2$. Ainsi, pour tout $x \in \mathbb{R}^d \setminus \{0\}$,

$$g(x) = 1$$

et pout tout $t \geq 0$,

$$K(t) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(1 + \frac{d}{2}\right)} t^d.$$

1.2.2 Bref catalogue des processus ponctuels répulsifs

Un premier exemple de processus ponctuel répulsif est la grille aléatoire définie comme suit, voir aussi [61, Section 14.4.2]. Soit x_0 un point uniformément distribué dans le cube de $\mathbb{R}^d$ centré en 0 et de coté de longueur $a > 0$. Les points de la grille aléatoire de maille de longueur $a > 0$ sont les sommets de la grille carrée commençant en x_0 et de maille de longueur a. Ce processus exhibe une répulsion stricte puisque les points ne peuvent être à distance inférieure à a les uns de autres. De plus la configuration est très régulière puisque la distance entre les points est nécessairement multiple de a.

Une généralisation de la grille aléatoire est la grille aléatoire perturbée. Elle est définie à partir d'une grille aléatoire de maille de longueur $a > 0$ mais où chacun des points est soumis indépendamment à une translation aléatoire dont le vecteur est uniformément distribué dans $B(0, r)$ pour un $r > 0$ donné. Une réalisation sur $[-5, 5]^2$ avec $a = 1$ et $r = 0.2$ est représentée sur la Figure 1.2. Notons que dans cet exemple, l'intensité est égale à un. La grille aléatoire perturbée est moins répulsive que la grille aléatoire avec même longueur de maille, mais reste quand même très répulsive si r est petit par rapport à a. La grille aléatoire et la grille aléatoire perturbée (si $r < a/2$) font en fait partie d'une classe plus générale de processus appelée les **processus hardcore**.

Les processus hardcore sont l'exemple typique de processus strictement répulsifs. En effet, ils interdisent deux points d'être à distance inférieure à un certain $R > 0$ appelé le rayon hardcore. Un premier exemple est le **processus ponctuel de Matérn de type I** qui est obtenu à partir d'un processus ponctuel de Poisson où toutes les paires de points distants de moins de R on été retirées. Il existe plusieurs généralisations de ce processus, obtenues par différentes règles de piochage d'un processus de Poisson, ou en imposant un aléa sur le rayon hardcore, voir par exemple [59], [48], [63] et [60]. La généralisation la plus connue est le **processus ponctuel de Matérn de type II**, implémenté dans le package `spatstat` de `R` et que nous avons représenté sur la Figure 1.2 avec le rayon hardcore maximale pour un processus ponctuel de Matérn de type II d'intensité un. Cette simulation correspond à un processus ponctuel de

Poisson sous-jacent d'intensité infinie et notre simulation est seulement une approximation. Beaucoup d'autres processus hardcore proviennent de la Physique, par exemple le modèle **RSA** (random sequential absorption), le **processus des feuilles mortes**, ou le **hard sphere packing model**, voir [33] pour une présentation de ces processus. Cependant, la répulsion stricte est une condition restrictive qui ne permet pas toujours de représenter la réalité. Afin d'avoir plus de souplesse, les **processus de Gibbs** sont généralement utilisés.

Les processus de Gibbs sont la classe de processus la plus utilisée pour modéliser des configurations de points régulières. Un processus ponctuel $\mathbf{X}$ sur un borné S de $\mathbb{R}^d$ est un processus ponctuel de Gibbs s'il admet une densité f par rapport au processus de Poisson standard. Notons $\mathcal{N}_S := \{\mathbf{x} \subset S,\ \mathbf{x}(S) < \infty\}$ l'espace de toutes les configurations finies de points dans S et $\sigma(\mathcal{N}_S)$ la tribu engendrée associée, générée par tous les ensembles $\{\mathbf{x} \subset S,\ \mathbf{x}(S) = n\}$ où $n \in \mathbb{N}$. Nous avons, pour tout $F \in \sigma(\mathcal{N}_S)$,

$$P(\mathbf{X} \in F) = \sum_{n \geq 0} \frac{e^{-|S|}}{n!} \int_{S^n} \mathbf{1}_{\big\{\{x_1,\ldots,x_n\} \in F\big\}} f\big(\{x_1, \ldots, x_n\}\big) dx_1 \ldots dx_n,$$

la seule configuration avec zéro point étant l'ensemble vide. Habituellement, la densité f dans la formule précédente s'écrit

$$f(\mathbf{x}) = \frac{1}{c} \prod_{\mathbf{y} \subset \mathbf{x}} \phi(\mathbf{y}),$$

où $\mathbf{x}$ et $\mathbf{y}$ appartiennent à $\mathcal{N}_S$, ϕ est une fonction mesurable de $\mathcal{N}_S$ dans $\mathbb{R}^+$ et c est une constante normalisatrice, voir [49] pour plus de détails. Selon le choix de f, une très grande variété d'interactions peut être proposée. Un premier exemple est le modèle hardcore avec rayon hardcore $R > 0$, obtenu pour

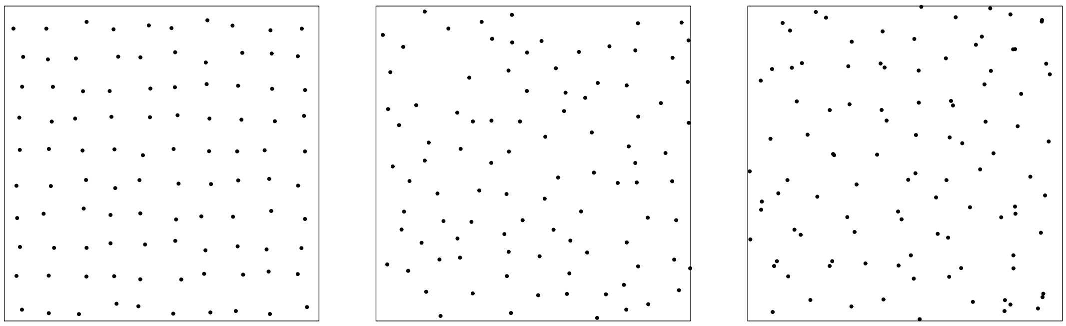
$$f(\mathbf{x}) \propto \beta^{\mathbf{x}(S)} \mathbf{1}_{\{s_R(\mathbf{x})=0\}},$$

où $\beta \geq 0$, et $s_R(\mathbf{x})$ est le nombre de paires de points de $\mathbf{x}$ qui sont distants d'au plus R. Un autre processus ponctuel de Gibbs bien connu est le **processus ponctuel de Strauss** défini sur tout borné $S \subset \mathbb{R}^d$ par la densité

$$f(\mathbf{x}) \propto \beta^{\mathbf{x}(S)} \gamma^{s_R(\mathbf{x})},$$

où $\beta \geq 0$, $\gamma \in [0, 1]$ et $R \geq 0$ sont des paramètres. Le paramètre γ est appelé le paramètre d'interaction : si $\gamma = 0$ on retrouve le processus précédent avec rayon hardcore R, si $0 < \gamma < 1$ le processus exhibe de la répulsion, et si $\gamma = 1$ c'est un processus ponctuel de Poisson. Une réalisation dans le cas où $S = [-5, 5]^2$, $\gamma = 0.1$, $R = 1$ et $\beta = 110$ est représentée sur la Figure 1.2. Il existe beaucoup d'autres exemples de processus de Gibbs, tels que le processus “area-interaction ” ou le processus de Lennard-Jones, voir par exemple [49] ou [65]. Remarquons que nous n'avons défini les processus ponctuels de Gibbs que sur des bornés. Il est toutefois possible de les définir sur $\mathbb{R}^d$ mais sous des hypothèses techniques supplémentaires, voir [49].

En conclusion, il existe déjà beaucoup de modèles de processus ponctuels répulsifs adaptés à de multiples configurations de points régulières. Parmi eux, les processus ponctuels de Gibbs sont les plus flexibles, dans le sens où ils peuvent modéliser une grande variété d'interactions. Cependant, le modèle de Gibbs peut être difficile à manipuler dans la pratique. La principale



(a) (b) (c)

Figure 1.2 – Trois exemples sur $[-5,5]^2$ de configurations de points répulsives avec intensité $\rho = 1$: (a) est une grille aléatoire perturbée avec $a = 1$ et $r = 0.2$, (b) est un processus ponctuel hardcore de Matérn de type II avec rayon $R = 1/\sqrt{\pi}$, (c) est un processus ponctuel de Strauss avec $R = 1$, $\gamma = 0.1$ et $\beta = 110$. Notons que β a été calculé numériquement afin d'avoir $\rho = 1$.

raison provient de la constante de normalisation c, qui apparaît dans la densité, et qui est impossible à calculer dans la plupart des cas. Ainsi, certaines méthodes de Monte-Carlo par chaîne de Markov sont nécessaires pour leur simulation. De même, l'inférence par vraisemblance est impossible sans des méthodes MCMC compliquées, et d'autres méthodes moins efficaces n'impliquant pas c, telles que la pseudo-vraisemblance, ont été développées pour estimer les paramètres. Enfin, les moments d'un processus ponctuel de Gibbs sont inconnus. Il s'ensuit que des quantités telles que la mesure d'intensité, la pcf ou la fonction de Ripley sont inconnues et doivent être approchées par simulation.

1.2.3 Les DPPs, une classe intéressante de processus ponctuels répulsifs

Les DPPs ont été introduits dans leur forme actuelle par Macchi dans [46], afin de modéliser la position de particules qui se repoussent entre elles. Ils ont la particularité d'être définis par leurs moments, et plus précisément par leurs intensités jointes.

Définition 1.2.2. *Si elle existe, l'intensité jointe d'ordre k ($k \geq 1$) d'un processus ponctuel simple, notée $\rho^{(k)}$, est la densité par rapport à la mesure de Lebesgue de la mesure moment factorielle $\alpha^{(k)}$ (voir Définition 1.2.1).*

Sous notre hypothèse de stationnarité, nous avons $\rho^{(k)}(x_1, \dots, x_k) = \rho^{(k)}(0, x_2 - x_1, \dots, x_k - x_1)$. En particulier, l'intensité ρ et la fonction d'intensité du second ordre $\rho^{(2)}$ introduites dans la Section 1.2.1 correspondent respectivement à $k = 1$ et $k = 2$.

Définition 1.2.3. *Soit une fonction $C : \mathbb{R}^d \to \mathbb{R}$. Un processus ponctuel $\mathbf{X}$ sur $\mathbb{R}^d$ est un processus ponctuel déterminantal stationnaire de noyau C, en abrégé $\mathbf{X} \sim DPP(C)$, si pour tout $k \geq 1$, les intensités jointes d'ordre k vérifient la relation*

$$\rho^{(k)}(x_1, \dots x_k) = \det[C](x_1, \dots, x_k)$$

pour presque tout $(x_1, \dots, x_k) \in \mathbb{R}^{dk}$, où $[C](x_1, \dots, x_k)$ est la matrice avec entrées $C(x_i - x_j)$, pour $1 \leq i, j \leq k$.

L'existence d'un DPP requiert plusieurs conditions sur le noyau C. Nous renvoyons à [32] pour une présentation détaillée de ces conditions. Pour simplifier, nous donnons ci-dessous des conditions suffisantes d'existence d'un DPP stationnaire, basées sur la transformée de Fourier et aisément vérifiables en pratique. Nous définissons la transformée de Fourier d'une fonction $h \in L^1(\mathbb{R}^d)$ comme

$$\mathcal{F}(h)(t) = \int_{\mathbb{R}^d} h(x) e^{2i\pi x \cdot t} dx, \quad \forall t \in \mathbb{R}^d,$$

et nous prolongeons sa définition aux fonctions $L^2(\mathbb{R}^d)$ par le théorème de Plancherel, voir [58]. Ainsi, nous avons le résultat d'existence suivant.

Proposition 1.2.4 ([44])**.** *Pour C une fonction symétrique, continue à valeurs réelles et dans $L^2(\mathbb{R}^d)$, $DPP(C)$ existe si et seulement $0 \leq \mathcal{F}(C) \leq 1$.*

Autrement dit, par la Proposition 1.2.4, toutes les fonctions de covariance C continue, à valeurs réelles et qui sont dans $L^2(\mathbb{R}^d)$ avec $\mathcal{F}(C) \leq 1$ définissent un DPP. Dans ce manuscrit, nous ne considérerons que des DPPs stationnaires avec noyau à valeurs réelles C. Il est toutefois possible de considérer des noyaux à valeurs complexes et/ou des DPPs non stationnaires. Nous renvoyons vers [32] pour une présentation sur les DPPs dans le cas général. Pour toute la suite, nous supposerons la condition suivante vérifiée.

Condition $\mathcal{K}(\rho)$. Un noyau C vérifie la condition $\mathcal{K}(\rho)$ si C est une fonction symétrique, continue, à valeurs réelles, dans $L^2(\mathbb{R}^d)$, et telle que $C(0) = \rho$ et $0 \leq \mathcal{F}(C) \leq 1$.

Remarquons que si $\mathbf{X} \sim DPP(C)$ pour C vérifiant $\mathcal{K}(\rho)$ avec un $\rho > 0$ donné, alors l'intensité de $\mathbf{X}$ est ρ. La condition $\mathcal{K}(\rho)$ est vérifiée par de nombreuses fonctions de covariance, rendant ainsi aisée la définition de familles paramétriques de DPPs. Par exemple, nous pouvons considérer la famille à noyaux Gaussiens

$$C(x) = \rho e^{-\left|\frac{x}{\alpha}\right|^2}, \quad x \in \mathbb{R}^d, \tag{1.2.1}$$

où $|.|$ désigne la norme Euclidienne sur $\mathbb{R}^d$, $\rho > 0$ et $\alpha \leq 1/(\sqrt{\pi}\rho^{1/d})$, la dernière contrainte sur les paramètres étant une conséquence de la condition d'existence $\mathcal{F}(C) \leq 1$ dans $K(\rho)$. Nous introduisons de nouvelles familles paramétriques au cours du Chapitre 3, telles que la famille de noyaux type Bessel, définie par

$$C(x) = \kappa \frac{J_{\frac{\sigma+d}{2}}\left(2\left|\frac{x}{\alpha}\right|\sqrt{\frac{\sigma+d}{2}}\right)}{\left(2\left|\frac{x}{\alpha}\right|\sqrt{\frac{\sigma+d}{2}}\right)^{\frac{\sigma+d}{2}}}, \quad x \in \mathbb{R}^d, \tag{1.2.2}$$

où $\rho > 0$, $\sigma \geq 0$, $\alpha < \alpha_{\max}(\rho, \sigma)$ et $J_{\frac{\sigma+d}{2}}$ est la fonction de Bessel de premier genre d'ordre $\frac{\sigma+d}{2}$. Les constantes κ et $\alpha_{\max}(\rho, \sigma)$ sont données en Section 3.5. Cette famille est particulièrement intéressante dans la mesure où elle peut couvrir toute la plage de répulsion des DPPs, du processus ponctuel de Poisson stationnaire (la situation sans interaction) au DPP le plus répulsif (tel que nous le déterminons dans la Section 3.3).

D'un point de vue méthodologique, les DPPs possèdent de nombreuses propriétés intéressantes :

- des familles paramétriques peuvent facilement être définies comme nous l'avons vu précédemment,
- de par leur définition, les moments de tout ordre sont connus. En particulier, la fonction de Ripley K et la pcf sont explicitement connus,
- la densité des DPPs par rapport à un processus ponctuel de Poisson standard est connu sur tout compact $S \in \mathbb{R}^d$ et la constante normalisatrice est explicitement connue, contrairement aux processus de Gibbs, pourvu qu'une représentation spectrale du noyau soit accessible. Si cette dernière condition n'est pas remplie, une approximation efficace de la représentation spectrale via la transformée de Fourier de C a été proposée et étudiée dans [44],
- en conséquence, la méthode du maximum de vraisemblance est faisable pour estimer les paramètres (pourvu qu'une représentation spectrale de C soit connue ou approchée),
- une estimation par minimum de contraste basée K ou g peut aussi être utilisée pour estimer les paramètres,
- la simulation des DPPs sur un compact peut être effectuée parfaitement (si une décomposition spectrale de C est connue) et rapidement. Un algorithme de simulation est présenté dans [44] et implémenté dans le package `spatstat` de `R`.

Pour $\rho = 100$, nous avons simulé et représenté sur la Figure 1.3 différentes réalisations sur $[0, 1]^2$ de : (a)-(c) DPPs avec noyaux (1.2.1) et $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$, (d)-(f) DPPs avec noyaux (1.2.2), $\sigma = 0$ et $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$. La dernière valeur de α correspond à la valeur maximale possible due à la restriction sur la transformée de Fourier du noyau dans $\mathcal{K}(\rho)$, voir Section 3.5. De plus, les paramètres de ces modèles ont été estimés par minimum de contraste basé sur K, sur g et par maximum de vraisemblance (en utilisant l'approximation de la représentation spectrale du noyau de [44]). Ces méthodes d'estimation sont disponibles dans `spatstat`. Par exemple, nous donnons dans les Tableaux 1.1-1.2, l'erreur quadratique moyenne des trois méthodes sur 500 réalisations de DPPs, avec les mêmes noyaux et paramètres que les DPPs de la Figure 1.3, mais simulés sur $[0, 1]^2$, $[0, 2]^2$ et $[0, 3]^2$.

Pour toutes les méthodes d'estimation considérées dans les Tableaux 1.1-1.2, les estimateurs semblent consistants, et la précision, au sens de l'erreur quadratique moyenne, augmente avec la taille de la fenêtre. D'après ces résultats, l'estimation par maximum de vraisemblance semble la meilleure méthode, ce qui confirme les observations déjà faites dans [44]. Cependant, l'estimation par minimum de contraste, particulièrement celle basée sur g, semble donner de bons résultats. De plus, leur temps de calcul est bien plus court que pour l'estimation par maximum de vraisemblance. Par exemple, l'estimation du paramètre α pour les DPPs sur $[0, 3]^2$, avec noyaux (1.2.1) ou (1.2.2), prend quelques minutes avec l'estimation par minimum de contraste, contre plusieurs heures avec l'estimation par maximum de vraisemblance. Il semblerait enfin que les estimateurs ont asymptotiquement un comportement Gaussien, comme on le voit sur la Figure 1.4 où nous avons représenté l'histogramme obtenu à partir des estimations de α pour le DPP avec noyau (1.2.2), sur 500 réalisations sur $[0, 1]^2$ avec $\sigma = 0$ et $\alpha = 0.03$.

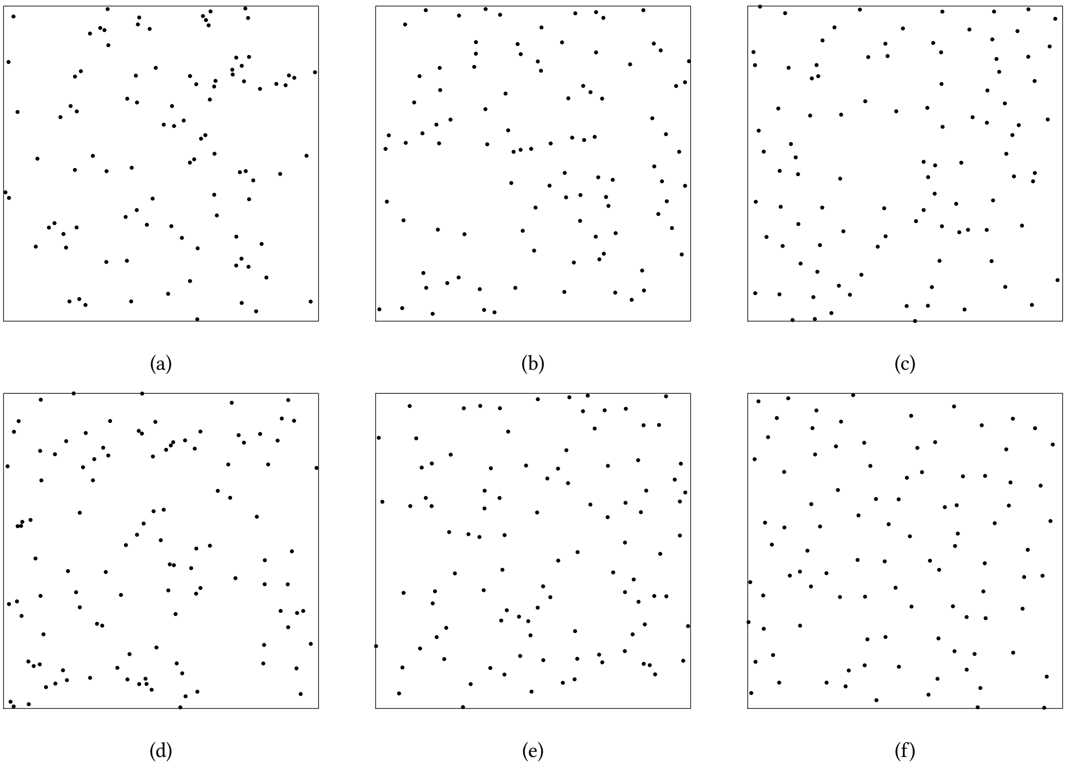


Figure 1.3 – Réalisations sur $[0,1]^2$ de (a)-(c) DPPs avec noyau (1.2.1) où, de gauche à droite, $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$, (d)-(f) DPPs avec noyau (1.2.2) où $\sigma = 0$ et, de gauche à droite, $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$. L'intensité ρ est égale à 100 pour tous les modèles.

	$[0,1]^2$			$[0,2]^2$			$[0,3]^2$		
	K	pcf	ML	K	pcf	ML	K	pcf	ML
$\alpha = 0.01$	2.026	1.039	1.032	0.848	0.309	0.220	0.521	0.175	0.096
$\alpha = 0.03$	1.214	0.706	0.786	0.419	0.248	0.175	0.231	0.180	0.084
$\alpha = 1/(10\sqrt{\pi})$	0.356	0.588	0.225	0.113	0.258	0.061	0.051	0.176	0.022

Table 1.1 – Erreur quadratique moyenne de l'estimateur obtenu par minimum de contraste basé sur K et g, tel que présenté dans Chapitre 5, et par maximum de vraisemblance (ML), comme présenté dans [44]. L'estimation a été faite sur 500 réalisations de DPPs sur $[0,1]^2$, $[0,2]^2$ et $[0,3]^3$, avec noyau (1.2.1) et $\rho = 100$. Toutes les entrées ont été multipliées par 10^4 pour rendre le tableau plus compact.

	$[0,1]^2$			$[0,2]^2$			$[0,3]^2$		
	K	pcf	ML	K	pcf	ML	K	pcf	ML
$\alpha = 0.01$	1.023	0.511	0.428	0.426	0.220	0.110	0.280	0.107	0.048
$\alpha = 0.03$	0.441	0.403	0.322	0.164	0.162	0.068	0.090	0.110	0.029
$\alpha = 1/(10\sqrt{\pi})$	0.068	0.194	0.046	0.021	0.091	0.012	0.008	0.055	0.002

Table 1.2 – Erreur quadratique moyenne de l'estimateur obtenu par minimum de contraste basé sur K et g, tel que présenté dans Chapitre 5, et par maximum de vraisemblance (MLE), comme présenté dans [44]. L'estimation a été faite sur 500 réalisations de DPPs sur $[0,1]^2$, $[0,2]^2$ et $[0,3]^3$, avec noyau (1.2.2), $\sigma = 0$ et $\rho = 100$. Toutes les entrées ont été multipliées par 10^4 pour rendre le tableau plus compact.

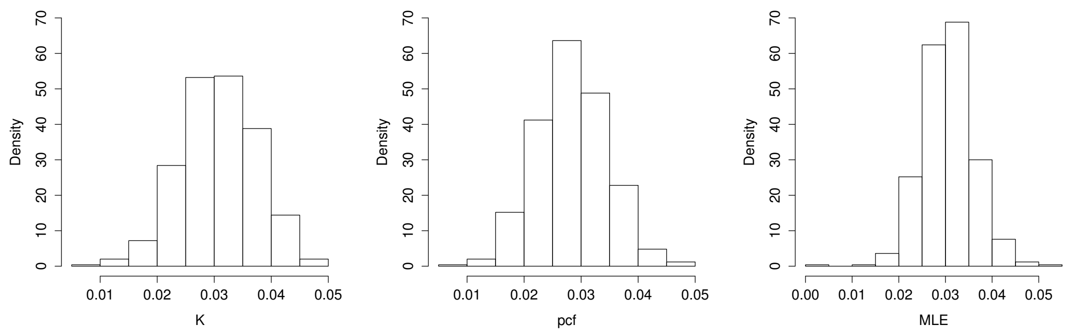


Figure 1.4 – Histogramme obtenu pour l'estimation du paramètre α des DPPs avec noyaux (1.2.2), sur 500 réalisations sur $[0,1]^2$, avec $\sigma = 0$ et $\alpha = 0.03$. De gauche à droite, les méthodes utilisées sont : minimum de contraste basé sur K, minimum de contraste basé sur g, maximum de vraisemblance.

En dépit de ces observations, il n'existe à notre connaissance aucun résultat théorique sur le comportement asymptotique de l'estimation par maximum de vraisemblance ou de l'estimation par minimum de contraste pour les DPPs. L'étude conduite au Chapitre 5 comble partiellement ce vide en démontrant la consistance et la normalité asymptotique des estimateurs obtenus par minimum de contraste basé sur la fonction de Ripley K ou la pcf g. Les propriétés asymptotiques du maximum de vraisemblance demeurent un problème ouvert.

En conclusion, les DPPs sont une classe particulièrement intéressante pour simuler et modéliser des configurations répulsives de points. Ils peuvent être une bonne alternative aux processus ponctuels de Gibbs dans plusieurs cas, par exemple lorsque la connaissance des moments est requise. Cependant, depuis leur introduction par Macchi dans [46], les DPPs ont été largement étudiés en Probabilité mais en comparaison assez peu d'un point de vue statistique jusqu'à récemment avec [44] et sa version longue [43]. D'autre travaux statistiques ont également été effectués avec [11, 47] en télécommunication, [41] dans le cas discret pour l'apprentissage statistique, [52] et [5]. Cette étude a pour ambition de fournir les bases théoriques à certaines méthodes statistiques pour les DPPs.

1.3 Contenu de la thèse

Le Chapitre 3 présente une étude détaillée de la répulsion possible pour un DPP stationnaire. En particulier, nous introduisons dans ce chapitre de nouvelles familles paramétriques de DPPs qui peuvent couvrir toute la plage de répulsion offerte pas les DPPs, du processus de Poisson stationnaire au DPP le plus répulsif. Dans le Chapitre 4, nous prouvons que les DPPs sont Brillinger-mélangeants et donnons plusieurs applications statistiques de cette propriété. Finalement, dans le Chapitre 5, nous prouvons la consistance et la normalité asymptotique d'estimateurs paramétriques obtenus par une méthode de minimum de contraste basée sur le fonction de Ripley K ou sur la fonction de corrélation par paires (pcf) g.

1.3.1 Chapitre 3

Dans ce chapitre, nous étudions la répulsion des processus déterminantaux stationnaires. Afin d'y parvenir, nous introduisons deux définitions de la répulsion pour un processus ponctuel stationnaire, toutes deux basées sur la pcf g que nous nommons répulsion globale et locale.

Définition. *Soient $\mathbf{X}$ et $\mathbf{Y}$ deux processus ponctuels stationnaires avec même intensité ρ et fonction de corrélation par paires respectives $g_{\mathbf{X}}$ et $g_{\mathbf{Y}}$.*

- *Supposons que $(1-g_{\mathbf{X}})$ et $(1-g_{\mathbf{Y}})$ soient intégrables, nous dirons que $\mathbf{X}$ est plus globalement répulsif que $\mathbf{Y}$ si $\int(1-g_{\mathbf{X}}) \geq \int(1-g_{\mathbf{Y}})$.*
- *Supposons que $g_{\mathbf{X}}$ est deux fois différentiable en 0 et vérifie $g_{\mathbf{X}}(0)=0$, nous dirons que $\mathbf{X}$ est plus localement répulsif que $\mathbf{Y}$ si $g_{\mathbf{Y}}(0)>0$, ou $g_{\mathbf{Y}}$ n'est pas deux fois différentiable en 0, ou $g_{\mathbf{Y}}$ est deux fois différentiable en 0 avec $g_{\mathbf{Y}}(0)=0$ et $\Delta g_{\mathbf{Y}}(0) \geq \Delta g_{\mathbf{X}}(0)$.*

Tandis que la répulsion globale considère le comportement de g sur tout l'espace $\mathbb{R}^d$, la répulsion locale ne considère que la courbure de g en l'origine (plus la courbe de g est plate, plus le processus est répulsif). Par exemple, les processus ponctuels hardcore en sont le cas extrême : en effet, pour eux, g est nul sur un voisinage de l'origine.

La répulsion globale a déjà été introduite dans [44] où les auteurs trouvent une infinité de DPPs étant les plus globalement répulsifs. En considérant la répulsion locale, nous avons prouvé la proposition suivante (voir Théorème 3.3.2 et Corollaire 3.3.3) qui exprime l'existence de l'unique DPP stationnaire sur $\mathbb{R}^d$ qui est le plus localement répulsif. De plus, c'est également l'un des plus globalement répulsifs. Pour ν un réel, nous notons J_ν la fonction de Bessel de premier genre à l'ordre ν.

Proposition. *Le DPP avec noyau*

$$C_B(x) = \frac{\sqrt{\rho\Gamma(\frac{d}{2}+1)}}{\pi^{d/4}} \frac{J_{\frac{d}{2}}\left(2\sqrt{\pi}\Gamma(\frac{d}{2}+1)^{\frac{1}{d}}\rho^{\frac{1}{d}}|x|\right)}{|x|^{\frac{d}{2}}}, \ \forall x \in \mathbb{R}^d,$$

est l'unique DPP avec noyau vérifiant $\mathcal{K}(\rho)$, qui est à la fois globalement et localement le DPP le plus répulsif parmi tous les DPPs stationnaires d'intensité ρ.

En particulier, les DPPs stationnaires ne peuvent pas inclure de distance hardcore. Ainsi, ils ne sont pas aussi flexibles que les processus ponctuels de Gibbs. Nous illustrons néanmoins dans ce Chapitre 3 que les DPPs demeurent une classe de modèle relativement flexible.

Nous étudions aussi dans ce Chapitre 3 la répulsion pour la sous classe des DPPs R-dépendant. Nous disons qu'un processus $\mathbf{X}$ est R-dépendant, pour un $R > 0$ fixé, si $\mathbf{X} \cap A$ et $\mathbf{X} \cap B$ sont indépendant pour tous A et B boréliens de $\mathbb{R}^d$ séparés par une distance plus grande que R. Pour un DPP, cela est équivalent à supposer que son noyau C est de portée finie, i.e. $C(x) = 0$ pour $|x| \geq R$. Cette étude s'est révélée délicate et nous n'avons obtenu que des résultats partiels sur le plus répulsif des DPPs R-dépendant à $R > 0$ fixé.

Nous concluons ce chapitre en proposant de nouvelles familles paramétriques de DPPs. Pour ces familles, nous donnons une forme explicite du noyau et de sa transformée de Fourier, ce qui est particulièrement intéressant pour l'algorithme de simulation implémenté dans `spatstat`. De plus, ces familles paramétriques sont suffisamment flexibles par rapport à leurs paramètres pour couvrir toute la plage de répulsion offerte par les DPPs stationnaires, du processus ponctuels de Poisson (la situation sans interaction) au DPP le plus répulsif tel que déterminé dans la proposition précédente.

1.3.2 Chapitre 4

Un processus ponctuel est dit Brillinger-mélangeant si, pour tout $k \geq 2$, la variation totale de la mesure cumulante factorielle réduite d'ordre k est finie. Dans ce chapitre, nous rappelons la définition de ces mesures, et, pour les DPPs avec noyau C, nous l'exprimons en fonction de C. Ainsi, nous avons été en mesure de prouver le théorème suivant (voir Théorème 4.2.2).

Théorème. *Un DPP avec noyau vérifiant la condition $\mathcal{K}(\rho)$, pour un $\rho > 0$ donné, est Brillinger-mélangeant.*

En utilisant le mélange de Brillinger, nous avons prouvé un théorème limite central pour une large classe de fonctionnelles d'ordre $p \geq 1$ d'un processus ponctuel stationnaire, généralisant ainsi un théorème donné dans [40]. En particulier, ce théorème s'applique aux DPPs avec noyaux vérifiant $\mathcal{K}(\rho)$, et la variance asymptotique obtenue est explicitement connue. En application de ce théorème quand $p = 1$, nous retrouvons un résultat bien connu sur l'estimateur naturel de l'intensité, voir [57] et Corollaire 4.3.3.

Théorème. *Soit $\mathbf{X}$ un DPP avec noyau vérifiant pour un $\rho > 0$ la condition $\mathcal{K}(\rho)$ et $\{D_n\}_{n\in\mathbb{N}}$ une famille d'ensembles réguliers qui tendent vers $\mathbb{R}^d$ quand n tend vers l'infini. Soit, pour tout $n \in \mathbb{N}$,*

$$\widehat{\rho}_n = \frac{1}{|D_n|} \sum_{x \in \mathbf{X}} \mathbf{1}_{\{x \in D_n\}}.$$

Nous avons la convergence

$$\sqrt{|D_n|}\,(\widehat{\rho}_n - \rho) \xrightarrow[n\to+\infty]{distr.} N(0, \sigma^2)$$

où $\sigma^2 = \lim_{n\to+\infty} Var\left(\sqrt{|D_n|}\widehat{\rho}_n\right) = \rho - \int_{\mathbb{R}^d} C(x)^2 dx$.

Plusieurs autres applications du mélange de Brillinger aux processus ponctuels stationnaires peuvent être trouvées dans [26, 27, 29, 36, 37] et peuvent ainsi s'appliquer aux DPPs avec noyau vérifiant $\mathcal{K}(\rho)$ pour un $\rho > 0$ donné. Par exemple, nous déduisons plusieurs résultats concernant l'estimateur à noyau suivant de la pcf. Soit $\mathbf{X}$ un DPP stationnaire et isotrope, nous définissons l'estimateur avec noyau k et fenêtre $\{b_n\}_{n\in\mathbb{N}}$ de la pcf,

$$\widehat{g}_n(t) = \frac{1}{\sigma_d t^{d-1}\widehat{\rho}_n^2} \sum_{\substack{(x,y)\in\mathbf{X}^2 \\ x\neq y}} \mathbf{1}_{\{x\in D_n,\ y\in D_n\}} \frac{1}{b_n|D_n\cap D_n^{x-y}|} k\left(\frac{t-|x-y|}{b_n}\right)$$

où $\sigma_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ est l'aire de la sphère unité en dimension d, $\{D_n\}_{n\in\mathbb{N}}$ est une suite d'ensembles dans $\mathbb{R}^d$, et pour tous $n\in\mathbb{N}$ et $z\in\mathbb{R}^d$, on a $D_n^z := \{u, u+z\in D_n\}$. La normalité asymptotique de cet estimateur est une conséquence de [27] et du fait que les DPPs sont Brillinger-mélangeants (voir Proposition 4.3.4).

Proposition. *Supposons que $\{D_n\}_{n\in\mathbb{N}}$ soit une séquence de sous-ensembles réguliers de $\mathbb{R}^d$ tels que $b_n^3|D_n|\to+\infty$ et $b_n^5|D_n|\to 0$. Soit k une fonction symétrique, bornée, à support compact inclus dans $[-T,T]$ pour un $T>0$ donné et vérifiant $\int_{\mathbb{R}} k(x)dx = 1$. Soit également C un noyau deux fois différentiable sur $\mathbb{R}^d\setminus\{0\}$, vérifiant la condition $\mathcal{K}(\rho)$ pour un $\rho>0$ donné et tel que $DPP(C)$ soit isotrope. En notant g la pcf de $DPP(C)$, nous avons pour tout $t>0$, la convergence*

$$\sqrt{b_n|D_n|}\left(\widehat{g}_n(t)-g(t)\right) \xrightarrow[n\to+\infty]{distr.} N(0,\tau^2),$$

où $\tau^2 = 2\frac{g(t)}{\rho^2\sigma_d t^{d-1}}\sqrt{\int_{\mathbb{R}} k^2(x)dx}$.

Une conclusion similaire sur la normalité asymptotique de l'erreur quadratique intégrée de $\widehat{g}_n$ est aussi donnée au Chapitre 4 en utilisant certains résultats de [26].

1.3.3 Chapitre 5

Nous étudions dans ce chapitre les propriétés asymptotiques d'une méthode d'estimation par minimum de contraste d'un DPP paramétrique. Nous considérons une famille paramétrique de DPPs avec noyau $C_{\rho,\theta}$, où $\rho = C_{\rho,\theta}(0) > 0$ avec θ appartenant à un sous-ensemble Θ_ρ de $\mathbb{R}^p$ pour un $p\geq 1$ donné, et tel que pour tout $\rho>0$ et $\theta\in\Theta_\rho$, $C_{\rho,\theta}$ vérifie $\mathcal{K}(\rho)$. Nous supposons que pour tout $\rho>0$ et pour tout θ_0 point intérieur de Θ_{ρ_0} (cet intérieur étant supposé non vide), nous observons $\mathbf{X}\sim DPP(C_{\rho_0,\theta_0})$ sur une séquence $\{D_n\}_{n\in\mathbb{N}}$ de sous-ensembles de $\mathbb{R}^d$ qui grossissent vers $\mathbb{R}^d$ dans toutes les directions.

L'estimation de ρ_0 faite par $\widehat{\rho}_n$ est déjà discutée au Chapitre 4 et nous nous concentrons ici sur l'estimation de θ_0 par la méthode de minimum de contraste suivante. Pour $\rho>0$ et $\theta\in\Theta_\rho$, définissons $J(.,\theta)$ une fonction de $\mathbb{R}^d$ dans $\mathbb{R}^+$, qui est une statistique de $DPP(C_{\rho,\theta})$ ne dépendant pas de ρ. Pour les DPPs, les cas les plus importants et naturels sont la fonction de Ripley K et la pcf g.

Pour tout $n\in\mathbb{N}$, considérons $\widehat{J}_n$ un estimateur de J basé sur les observations de $\mathbf{X}$ sur D_n. Définissons la fonction de contraste

$$U_n(\theta) = \int_{r_{min}}^{r_{max}} w(t)\left\{\widehat{J}_n(t)^c - J(t,\theta)^c\right\}^2 dt,$$

où $r_{max} > r_{min} \geq 0$ et $c\neq 0$ sont des paramètres, et w est une fonction poids généralement choisie pour être lisse sur (r_{min}, r_{max}), ou au moins intégrable. Finalement, nous définissons l'estimateur de θ_0 par

$$\widehat{\theta}_n = \arg\min_{\theta \in \Theta_{\hat{\rho}_n}} U_n(\theta).$$

Nous prouvons au Chapitre 5 un théorème général sur les propriétés asymptotiques de $\widehat{\theta}_n$ dans le cas d'un processus ponctuel stationnaire quelconque. Appliqué aux DPPs pour les cas particuliers où J est la fonction de Ripley K ou la pcf g, nous obtenons le théorème suivant (voir Théorèmes 5.2.2 et 5.2.3).

Théorème. *Soit* $\mathbf{X}$ *un DPP avec noyau vérifiant* C_{ρ_0,θ_0} *pour un* $\rho_0 > 0$ *donné et* θ_0 *un point intérieur de* $\Theta_{\rho_0} \subset \mathbb{R}^p$*, où* $p \geq 1$*. Supposons que pour tout* $\theta \in \Theta_{\rho_0}$*,* $C_{\rho_0,\theta}$ *vérifie la condition* $\mathcal{K}(\rho)$ *et est deux fois différentiable par rapport à* θ*. Si* $J = K$ *ou* $J = g$*, nous avons, sous de faibles hypothèses techniques supplémentaires, la convergence*

$$\sqrt{|D_n|}(\widehat{\theta}_n - \theta_0) \xrightarrow[n \to +\infty]{distr.} \mathcal{N}(0, \Sigma),$$

où Σ *s'exprime en fonction de* C_{ρ_0,θ_0}*.*

2

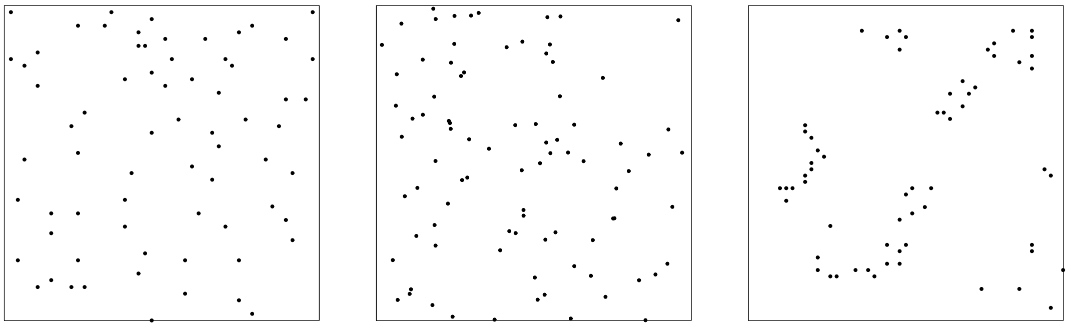
Introduction (english version)

2.1 Background

Spatial point patterns arise from many areas of science, for instance in Forestry as the location of trees, in Epidemiology as the location of diseased persons or in Quantum Mechanics as the position of particles. Many other examples may be found in the literature, see for instance [33]. Spatial point processes are the mathematical tool used to model these patterns. They may be seen as points randomly distributed in space, generally a compact set of $\mathbb{R}^2$. Spatial statistics are concerned with the description, the modelling and the inference of these distributions, in particular the dependence between the location of points is of main importance.

Many point process models have been introduced to represent a wide varieties of point patterns. In this manuscript, we restrict our study to stationary point processes, i.e. the distribution of points is assumed to be invariant by translation. The most simple example is the **Poisson point process**. A point process $\mathbf{X}$ is a Poisson point process on $\mathbb{R}^d$ with intensity $\rho > 0$ if on any bounded set B of $\mathbb{R}^d$, the number of points $\mathbf{X}(B)$ in B follows a Poisson distribution with parameter $\rho|B|$, where $|B|$ denotes the volume of B, and given $\mathbf{X}(B) = n$, the n points in B are independently and uniformly distributed in B. If $\rho = 1$, the Poisson point process is said to be standard. These processes model situations where no interaction is observed and they are obviously too simple to model all datasets. In fact, point patterns often exhibit repulsion or clustering as for example particles in statistical physics may repel each others or conversely attract each others to form clusters of points. Point processes are usually summarized in three categories, see for instance [33, Section 1] or [3]: regular (or repulsive, or inhibitive) if points are not too close to each others, random if there is no interaction and clustered if there is appearance of clusters. This classification is illustrated in Figure 2.1 where, from left to right, the point patterns are regular, random and clustered, respectively. Notice that this categorization is arbitrary and does not reflect the complexity of all datasets but is sufficient for a wide variety of applications.

The already mentioned Poisson processes model the case of no interaction between the points. To model attraction or clustering, **Cox processes**, also called doubly stochastic Poisson



(a) (b) (c)

Figure 2.1 – Three examples of point patterns: (a) locations of 69 Spanish towns in a 40 mile by 40 mile region, (b) realization on $[0,1]^2$ of a stationary Poisson point process with intensity $\rho = 100$, (c) locations of California redwood trees. The datasets in (a) and (c) are available in the R library `spatial` and `spatstat`, respectively.

point processes, are generally used. They are defined from inhomogeneous Poisson point processes, i.e. Poisson point processes where the intensity ρ is a function from $\mathbb{R}^d$ to $\mathbb{R}^+$. A point process $\mathbf{X}$ on $\mathbb{R}^d$ is a stationary Cox process driven by the stationary random intensity Z if conditionally to $Z(x) = \rho(x)$, $\mathbf{X}$ is a Poisson point process with intensity $\rho(x)$. This includes, for example, the Neyman-Scott and clusters point processes, see [49]. On the other hand, there exist several point process models for pure inhibition or repulsiveness. We review the usual ones in Section 2.2 below, including the wide class of **Gibbs point processes**. Recently, another class of point processes known as **determinantal point processes** (DPPs) has been studied to model repulsiveness. They have shown to be, in some cases, a good alternative to Gibbs point processes for the study of several regular point patterns, as the one in Figure 2.1 (a), see [43]. DPPs are briefly presented in Section 2.2.3. This manuscript is devoted to their study.

2.2 Repulsive point processes and determinantal point processes

2.2.1 Preliminaries on point processes

For $d \geq 1$, let $\mathbb{R}^d$ be equipped with the σ-algebra $\mathcal{B}(\mathbb{R}^d)$ of Borel sets and $\mathcal{B}_0(\mathbb{R}^d)$ be the class of bounded Borel sets on $\mathbb{R}^d$. Denote by $\mathbf{x} := \{x_1, , x_2, \ldots, x_n, \ldots\}$ an arbitrary configuration of points in $\mathbb{R}^d$ and by $\mathbf{x}(B)$ the number of points of $\mathbf{x}$ that belongs to B, for an arbitrary $B \in \mathcal{B}_0(\mathbb{R}^d)$. Define $\mathcal{N} := \{\mathbf{x} \subset \mathbb{R}^d,\ \mathbf{x}(B) < \infty,\ \forall B \in \mathcal{B}_0(\mathbb{R}^d)\}$ as the space of locally finite configurations of points in $\mathbb{R}^d$. We equipped $\mathcal{N}$ with the σ-algebra generated by the sets $\{\mathbf{x} \subset \mathbb{R}^d,\ \mathbf{x}(B) = n\}$ for all $B \in \mathcal{B}_0(\mathbb{R}^d)$ and all $n \in \mathbb{N} \cup \{0\}$, where $\mathbb{N}$ is the set of positive integers. A point process on $\mathbb{R}^d$ is a measurable application from a probability space $(\Omega, \mathcal{F}, P)$ into the set $\mathcal{N}$. We denote a point process by a bold capital letter, usually $\mathbf{X}$, and identify the mapping $\mathbf{X}$ and the associated random set of points. Then, to indicate that a point x belongs to $\mathbf{X}$, we write $x \in \mathbf{X}$. Further, we always assume that the point processes are simple, i.e. two points of the process never coincide, almost surely. For further details, we refer to [9, 10].

Important quantities for the study of point processes are the so-called factorial moment measures defined as follow.

Definition 2.2.1. *The factorial moment measure of order k ($k \geq 1$) of a simple point process $\mathbf{X}$ is the measure on $\mathbb{R}^{dk}$, denoted by $\alpha^{(k)}$, such that for any family of subsets $D_1, \ldots, D_k$ in $\mathbb{R}^d$,*

$$\alpha^{(k)} \left(D_1 \times \ldots \times D_k \right) = \mathbb{E} \left(\sum_{(x_1,\ldots,x_k) \in \mathbf{X}^k}^{\neq} \mathbf{1}_{\{x_1 \in D_1, \ldots, x_k \in D_k\}} \right)$$

where $\mathbb{E}$ is the expectation over the distribution of $\mathbf{X}$ and the symbol $\neq$ over the sum means that we consider only mutually disjoints k-tuples of points $x_1, \ldots, x_k$.

Cases of main interest are the factorial moment measure of order one, called the intensity measure, and the factorial moment measure of order two. Under our stationary assumption, the intensity measure is proportional to the Lebesgue measure on $\mathbb{R}^d$ which means that for all $S \subset \mathbb{R}^d$, $\alpha^{(1)}(S) = \rho|S|$, for a given $\rho > 0$. In this case, ρ is called the **intensity of the process** and represents the expected number of points per unit volume. Note that ρ is the density of the intensity measure with respect to the Lebesgue measure. If $\alpha^{(2)}$ admits a density with respect to the Lebesgue measure on $\mathbb{R}^{2d}$, this density is denoted by $\rho^{(2)}$ and called **second order intensity function**. Heuristically, for $(x, y) \in \mathbb{R}^{2d}$ and $x \neq y$, $\rho^{(2)}(x, y)$ may be viewed as the probability that there is a point of the process in a small neighbourhood around x and another point in a small neighbourhood around y. Note that under our stationary assumption, $\rho^{(2)}(x, y) = \rho^{(2)}(0, y - x)$ is a symmetric function and depends only on $y - x$.

In spatial statistics, the second order properties of a point process are often studied through summary statistics such as the **pair correlation function (pcf)** or the **Ripley's K-function** defined as follows. The pcf is defined for almost every $x \in \mathbb{R}^d$ by

$$g(x) = \frac{\rho^{(2)}(0, x)}{\rho^2}$$

and for all $t \geq 0$, the K-function is defined by

$$K(t) = \int_{B(0,t)} g(x) dx$$

where $B(0, t)$ denotes the Euclidean ball centred at 0 with radius t.

Intuitively, $g(x)$ is the quotient of the probability that two points separate by x occur (taking into account the interaction induced by the process) over the same probability for a Poisson process (the case of no interaction). Consequently, for $x \in \mathbb{R}^d$, a common interpretation, see for instance [59], is that $g(x) > 1$ characterizes clustering while $g(x) < 1$ characterizes repulsiveness. Regarding the K-function, $\rho K(t)$ is interpreted as the mean number of points in $B(0, t)$ assuming that 0 is a point of the process, see [49] for further details.

For instance, for the stationary Poisson point process on $\mathbb{R}^d$ with intensity $\rho > 0$, the second order intensity function exists and for all $(x, y) \in \mathbb{R}^{2d}$ such that $x \neq y$, $\rho^{(2)}(x, y) = \rho^2$. Thus, for all $x \in \mathbb{R}^d \setminus \{0\}$,

$$g(x) = 1$$

and for all $t \geq 0$,

$$K(t) = \frac{\pi^{\frac{d}{2}}}{\Gamma\left(1+\frac{d}{2}\right)}\, t^d.$$

2.2.2 Short catalogue of repulsive point processes

A first simple example of repulsive point process is the **lattice point process** defined as follows, see also [61, Section 14.4.2]. Let a point x_0 be uniformly distributed in the square of $\mathbb{R}^d$ centred at 0 with side length equal to a given $a > 0$. The points of the lattice point process with mesh width a are the vertices of the quadratic lattice with mesh width a starting at x_0. This process exhibits pure inhibition since the points cannot be closer than a distance a apart and the associated point patterns are very regular since the interpoint distances are necessarily multiple of a.

A generalization of the lattice point process is the **jittered lattice point process**. It is defined from a lattice point process with given mesh width $a > 0$ where each point is subjected to an independent random displacement with random vector uniformly distributed in $B(0, r)$ for a given $r > 0$. A realization in $[-5, 5]^2$, with $a = 1$ and $r = 0.2$ is represented in Figure 2.2. Note that for this example, the intensity equals one. The jittered lattice point process is less regular than the lattice point process with same mesh width but remains strongly regular if r is small with respect to a. Both the jittered lattice (if $r < a/2$) and lattice point processes are in fact examples of a more general class of processes named the **hardcore point processes**.

The hardcore point processes are the typical examples of purely inhibitive point processes. They are processes with interpoint distance at least greater than a given $R > 0$, named the hardcore radius. A first example is the **type I Matérn hardcore point process** that is obtained from a Poisson point processes where all R-close pairs of points are deleted. There exist several generalizations obtained by a different thinning of the Poisson process or randomization of the hardcore radius, see for instance [59], [48], [63] and [60]. The most known generalization that is implemented in the `R` library `spatstat` is the **type II Matérn hardcore process** represented in Figure 2.2 with maximal hardcore radius that a type II Matérn hardcore process with unit intensity can reach, see [33, Section 6.5]. It corresponds to an infinite intensity of the underlying Poisson process and our simulation is only an approximation. Other hardcore point processes arise from Physics, for instance the **RSA** (random sequential absorption), **dead leaves** or **hard sphere packing models**, see [33] for a review on these processes. However, the pure inhibition condition is restrictive and does not always reflect the reality. To introduce more flexibility, **Gibbs point processes** are usually used.

Gibbs point processes are the most used class of models for regular point patterns. A point process $\mathbf{X}$ on a bounded subset S of $\mathbb{R}^d$ is a Gibbs point process if it admits a density f with respect to the standard Poisson point process. Denoting $\mathcal{N}_S := \{\mathbf{x} \subset S,\ \mathbf{x}(S) < \infty\}$ the space of all finite configurations of points in S and $\sigma(\mathcal{N}_S)$ its associated σ-algebra generated by the sets $\{\mathbf{x} \subset S,\ \mathbf{x}(S) = n\}$ for all $n \in \mathbb{N}$, we have for all $F \in \sigma(\mathcal{N}_S)$,

$$P(\mathbf{X} \in F) = \sum_{n \geq 0} \frac{e^{-|S|}}{n!} \int_{S^n} \mathbf{1}_{\left\{\{x_1,\dots,x_n\} \in F\right\}} f\big(\{x_1, \dots, x_n\}\big) dx_1 \dots dx_n,$$

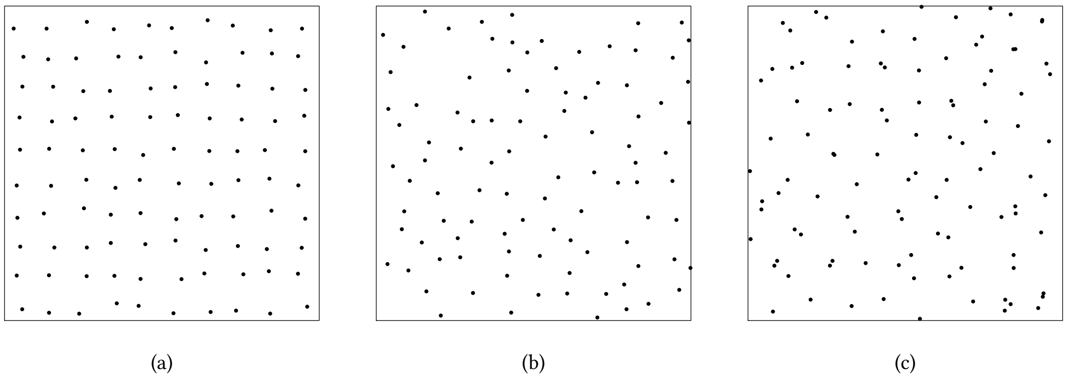


Figure 2.2 – Three examples of repulsive point patterns on $[-5,5]^2$ and with intensity $\rho = 1$: (a) a jittered lattice point process with $a = 1$ and $r = 0.2$, (b) a type II Matérn hardcore point process with radius $R = 1/\sqrt{\pi}$, (c) a Strauss point process with $R = 1$, $\gamma = 0.1$ and $\beta = 110$. Note that β has been numerically calibrated in order to have $\rho = 1$.

the only configuration with zero points being the null set. Usually, the density f in the last formula writes

$$f(\mathbf{x}) = \frac{1}{c} \prod_{\mathbf{y} \subset \mathbf{x}} \phi(\mathbf{y})$$

where $\mathbf{x}$ and $\mathbf{y}$ belongs to $\mathcal{N}_S$, ϕ is a measurable function from $\mathcal{N}_S$ into $\mathbb{R}^+$ and c is a normalizing constant, see [49] for further details. A large variety of interactions can be considered, depending on the choice of f. A first example is the hardcore model with hardcore radius $R \geq 0$ obtained for

$$f(\mathbf{x}) \propto \beta^{\mathbf{x}(S)} \mathbf{1}_{\{s_R(\mathbf{x})=0\}}$$

where $\beta \geq 0$ and $s_R(\mathbf{x})$ denotes the number of R-close pairs of points in $\mathbf{x}$. Another well known Gibbs point process is the **Strauss point process** defined on a bounded set $S \subset \mathbb{R}^d$ by the density

$$f(\mathbf{x}) \propto \beta^{\mathbf{x}(S)} \gamma^{s_R(\mathbf{x})}$$

where $\beta \geq 0$, $\gamma \in [0,1]$ and $R \geq 0$ are parameters. The parameter γ is called the interaction parameter: if $\gamma = 0$ we retrieve the previous purely inhibitive model with hardcore radius R, if $0 < \gamma < 1$ the process shows inhibition and if $\gamma = 1$ it is a Poisson point process. A realization in the case $S = [-5,5]^2$, $\gamma = 0.1$, $R = 1$ and $\beta = 110$ is represented in Figure 2.2. There are many other examples of Gibbs models as the area-interaction or Lennard-Jones processes, see for instance [49] or [65]. Note that we have only defined Gibbs point processes on a bounded set but it is in fact possible to define them on $\mathbb{R}^d$ under some technical assumptions on their densities, see [49].

In conclusion, there already exist many repulsive point process models adapted to a large variety of regular point patterns. Among them, the Gibbs point processes are the most flexible in the sense that they can model a wide variety of interactions. However Gibbs models can be difficult to handle in practice. The main reason is that the normalizing constant c involved in the density of a Gibbs point process is unknown in most cases. Consequently, some Markov chain Monte-Carlo methods (MCMC) are necessary for their simulations. Similarly likelihood inference is infeasible without elaborate MCMC methods and some less efficient procedures that do not involve c, such as the pseudo-likelihood method, have been developed to estimate the parameters. Further, the moments of a Gibbs point process are unknown. Hence, quantities such as the intensity measure, the pcf or the Ripley's K-function are unknown and can be only approximated by simulations.

2.2.3 DPPs as an appealing class of repulsive point processes

DPPs have been introduced in their current form by Macchi in [46] to model the position of particles that repel each others. They have the particularity of being defined by their moments, namely the joint intensities defined as follow.

Definition 2.2.2. *If it exists, the joint intensity of order k ($k \geq 1$) of a simple point process, denoted by $\rho^{(k)}$, is the density with respect to the Lebesgue measure of the factorial moment measure $\alpha^{(k)}$ defined in Definition 2.2.1.*

Under our stationary assumption, we have $\rho^{(k)}(x_1, \dots, x_k) = \rho^{(k)}(0, x_2 - x_1, \dots, x_k - x_1)$. In particular, the intensity ρ and the second order intensity function $\rho^{(2)}$ introduced in Section 2.2.1 are the cases associated to $k = 1$ and $k = 2$ respectively.

Definition 2.2.3. *Let $C : \mathbb{R}^d \to \mathbb{R}$ be a function. A point process $\mathbf{X}$ on $\mathbb{R}^d$ is a stationary DPP with kernel C, in short $\mathbf{X} \sim DPP(C)$, if for all $k \geq 1$ its joint intensity of order k satisfies the relation*

$$\rho^{(k)}(x_1, \dots x_k) = \det[C](x_1, \dots, x_k)$$

for almost every $(x_1, \dots, x_k) \in \mathbb{R}^{dk}$, where $[C](x_1, \dots, x_k)$ denotes the matrix with entries $C(x_i - x_j)$, $1 \leq i, j \leq k$.

The existence of a DPP requires several conditions on the kernel C. We refer to [32] for a detailed presentation of these conditions. For convenience, we give sufficient conditions in the stationary case, based on the Fourier transform, that are easy to verify in practice. We define the Fourier transform of a function $h \in L^1(\mathbb{R}^d)$ as

$$\mathcal{F}(h)(t) = \int_{\mathbb{R}^d} h(x) e^{2i\pi x \cdot t} dx, \quad \forall t \in \mathbb{R}^d$$

and extend this definition to $L^2(\mathbb{R}^d)$ by Plancherel's theorem, see [58]. We have the following existence result.

Proposition 2.2.4 ([44]). *Assume C is a symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$. Then $DPP(C)$ exists if and only if $0 \leq \mathcal{F}(C) \leq 1$.*

In other words, by Proposition 2.2.4 any continuous real-valued covariance function C in $L^2(\mathbb{R}^d)$ with $\mathcal{F}(C) \leq 1$ defines a DPP. In the following, we consider only stationary DPPs with real-valued kernel C but it is actually possible to consider complex-valued kernels and/or non-stationary DPPs. We refer to [32] for a review on DPPs in the general case. Henceforth, we assume the following condition.

Condition $\mathcal{K}(\rho)$. A kernel C is said to verify condition $\mathcal{K}(\rho)$ if C is a symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$ with $C(0) = \rho$ and $0 \leq \mathcal{F}(C) \leq 1$.

Note that if $\mathbf{X} \sim DPP(C)$ where C satisfies $\mathcal{K}(\rho)$ for a given $\rho > 0$, then the intensity of $\mathbf{X}$ is ρ. The condition $K(\rho)$ is verified by numerous covariance functions, making easy the definition of parametric families of DPPs. For instance we can consider the family with Gaussian kernels

$$C(x) = \rho e^{-\left|\frac{x}{\alpha}\right|^2}, \quad x \in \mathbb{R}^d \tag{2.2.1}$$

where $|.|$ denote the Euclidean norm on $\mathbb{R}^d$, $\rho > 0$ and $\alpha \leq 1/(\sqrt{\pi}\rho^{1/d})$, the latter constraint on the parameter space being a consequence of the existence condition $\mathcal{F}(C) \leq 1$ in $K(\rho)$. We introduce new parametric families in Chapter 3, as the Bessel-type family with kernels

$$C(x) = \kappa \frac{J_{\frac{\sigma+d}{2}}\left(2\left|\frac{x}{\alpha}\right|\sqrt{\frac{\sigma+d}{2}}\right)}{\left(2\left|\frac{x}{\alpha}\right|\sqrt{\frac{\sigma+d}{2}}\right)^{\frac{\sigma+d}{2}}}, \quad x \in \mathbb{R}^d \tag{2.2.2}$$

where $\rho > 0$, $\sigma \geq 0$, $\alpha < \alpha_{\max}(\rho, \sigma)$ and $J_{\frac{\sigma+d}{2}}$ is the Bessel function of the first kind of order $\frac{\sigma+d}{2}$. The constants κ and $\alpha_{\max}(\rho, \sigma)$ are given in Section 3.5. This family is particularly interesting in that it can cover a large range of DPPs, from the stationary Poisson process (the case of no interaction) to what can be considered as the most repulsive DPP (as determined in Section 3.3).

From a methodological point of view, DPPs have numerous appealing properties:

- parametric families can easily be considered, as illustrated above;
- from the very definition, their moments of any order are known. In particular, important summary statistics such as the K-function or the pcf g are completely known;
- the density of DPPs with respect to the standard Poisson point process is known on any compact $S \subset \mathbb{R}^d$ without involving an incalculable constant as Gibbs point processes, providing we know a spectral representation for the kernel. If the latter is unknown, an efficient approximation in the Fourier domain is proposed and studied in [44];
- consequently, the maximum likelihood method is available to estimate the parameters (provided a spectral representation for C is known or approximated);
- minimum contrast methods based on summary statistics like K or g can also be used to estimate the parameters;
- the simulation of a DPP on a compact set can be done perfectly (if a spectral representation for C is known) and quickly. A simulation algorithm is presented in [44] and implemented in the `R` library `spatstat`.

As an illustration, letting $\rho = 100$, we have simulated and represented in Figure 2.3 realizations on $[0, 1]^2$ of: (a)-(c) DPPs with kernels (2.2.1) and $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$, (d)-(f) DPPs with kernels (2.2.2), $\sigma = 0$ and $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$. The last value of α corresponds to the maximal possible value due to the restriction on the Fourier transform of the kernel in $\mathcal{K}(\rho)$, see Section 3.5. Moreover, the parameters of these models have been estimated with the minimum contrast method based on K, on g and with the maximum likelihood method (using the approximation of the spectral representation of the kernel of [44]). These procedures are also implemented in `spatstat`. For instance, we give in Tables 2.1-2.2 the mean squared errors of the three mentioned methods over 500 realisations of DPPs with the same kernels and parameters than DPPs in Figure 2.3 but simulated on $[0,1]^2$, $[0,2]^2$ and $[0,3]^2$.

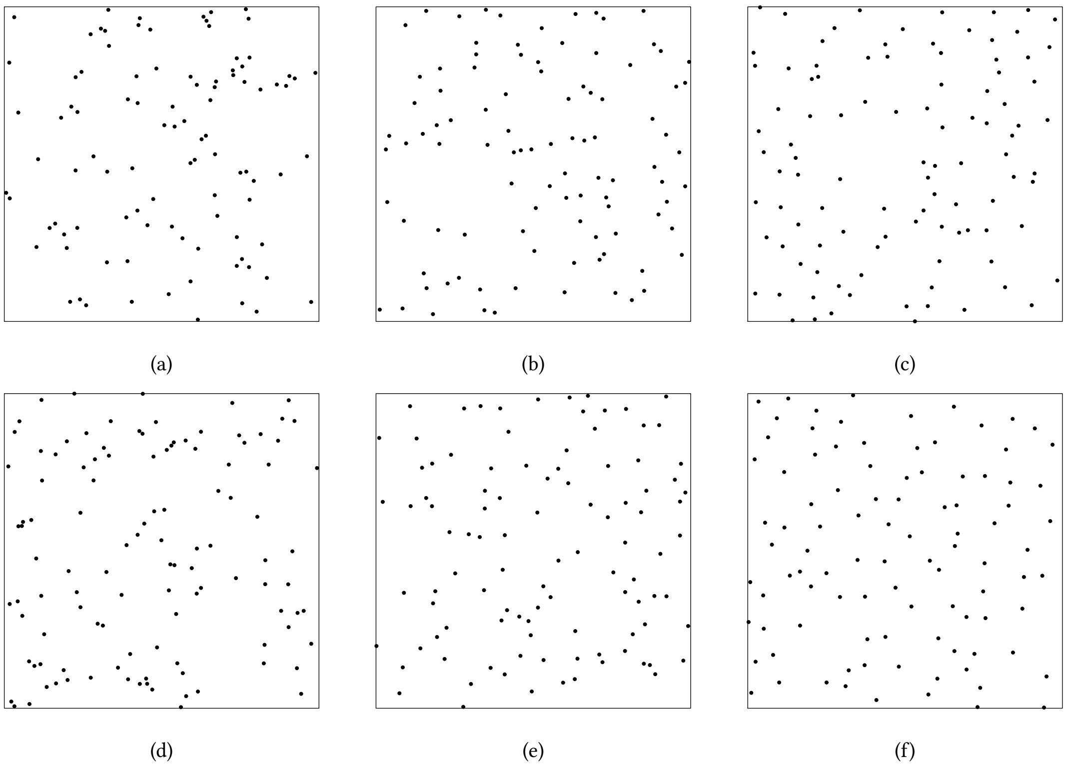


Figure 2.3 – Realizations on $[0,1]^2$ of (a)-(c) DPPs with kernel (2.2.1) where from left to right $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$, (d)-(f) DPPs with kernel (2.2.2) where $\sigma = 0$ and from left to right $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$. The intensity is $\rho = 100$ for all models.

	$[0,1]^2$			$[0,2]^2$			$[0,3]^2$		
	K	pcf	ML	K	pcf	ML	K	pcf	ML
$\alpha = 0.01$	2.026	1.039	1.032	0.848	0.309	0.220	0.521	0.175	0.096
$\alpha = 0.03$	1.214	0.706	0.786	0.419	0.248	0.175	0.231	0.180	0.084
$\alpha = 1/(10\sqrt{\pi})$	0.356	0.588	0.225	0.113	0.258	0.061	0.051	0.176	0.022

Table 2.1 – Mean squared error of the estimators obtained by minimum contrast method based on K and g as presented in Chapter 5 and by maximum likelihood method (ML) as presented in [44]. The estimation is done over 500 realizations of DPPs on $[0,1]^2$, $[0,2]^2$ and $[0,3]^3$ with kernel (2.2.1) and $\rho = 100$. All entries are multiplied by 10^4 to make the table more compact.

	$[0,1]^2$			$[0,2]^2$			$[0,3]^2$		
	K	pcf	ML	K	pcf	ML	K	pcf	ML
$\alpha = 0.01$	1.023	0.511	0.428	0.426	0.220	0.110	0.280	0.107	0.048
$\alpha = 0.03$	0.441	0.403	0.322	0.164	0.162	0.068	0.090	0.110	0.029
$\alpha = 1/(10\sqrt{\pi})$	0.068	0.194	0.046	0.021	0.091	0.012	0.008	0.055	0.002

Table 2.2 – Mean squared error of the estimators obtained by minimum contrast method based on K and g as presented in Chapter 5 and by maximum likelihood method as presented in [44]. The estimation is done over 500 realizations of DPPs on $[0,1]^2$ and $[0,2]^2$ and $[0,3]^3$ with kernel (2.2.2), $\sigma = 0$ and $\rho = 100$. All entries are multiplied by 10^4 to make the table more compact.

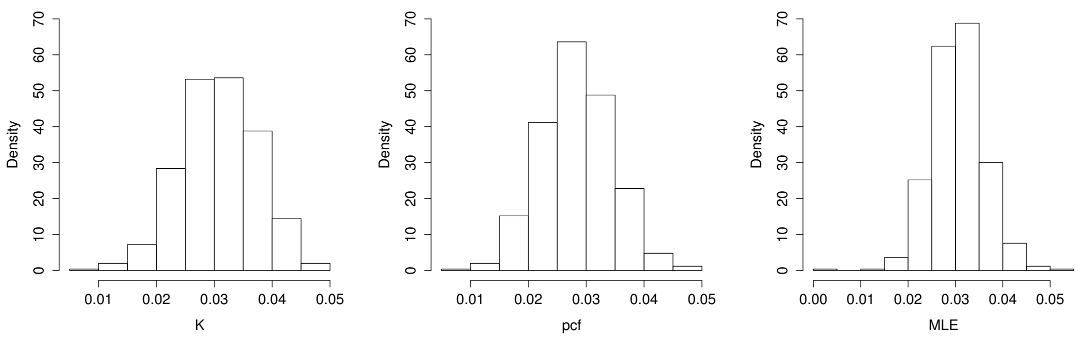


Figure 2.4 – Histogram obtained for the estimations of the parameter α of DPPs with kernel (2.2.2) over 500 realizations on $[0,1]^2$ with $\sigma = 0$ and $\alpha = 0.03$. From left to right the methods used are: minimum contrast based on K, minimum contrast based on g, maximum likelihood.

For all methods considered in Tables 2.1-2.2, the estimators seem consistent and the precision, in the sense of the mean squared error, increases with the windows size. From these results, maximum likelihood method seems to be the best method, which agrees with the observations made in [44]. However, minimum contrast methods, especially the one based on g, seem to perform reasonably well. Moreover, their computation is faster than the maximum likelihood method. For instance, the estimation of the parameter α for DPPs on $[0,3]^2$, with kernels (2.2.1) or (2.2.2), takes some minutes by minimum contrast methods while it takes several hours by the maximum likelihood method. Finally, it seems that each estimator has asymptotically a Gaussian behaviour, as illustrated in Figure 2.4 where we have represented the histogram obtained from the estimation of α for a DPP with kernel (2.2.2) over 500 realizations on $[0,1]^2$ with $\sigma = 0$ and $\alpha = 0.03$.

Despite these observations, to the best of our knowledge, no asymptotic results were available on maximum likelihood or minimum contrast methods for DPPs. The study conducted in Chapter 5 partially fills this gap by proving the consistency and asymptotic normality of the estimators obtained by minimum contrast method with the K-function or the pcf g. The asymptotic properties of the maximum likelihood estimator still remain an open problem.

In conclusion, DPPs are very attractive to simulate and to fit repulsive point patterns. They may present a good alternative to Gibbs point processes in several cases where, for instance, the knowledge of moments is preferred. However, since their introduction by Macchi in [46], DPPs have been largely studied in Probability but mostly unexplored from a statistical point of view point until recently with [44] and its associated supplementary materials [43]. Other statistical works include [11, 47] in telecommunication, [41] in the discrete case for statistical learning, [52] and [5]. This study has for ambition to provide theoretical backgrounds of some statistical methods for DPPs.

2.3 Contents of the PhD

Chapter 3 provides a detailed study of the possible repulsiveness of stationary DPPs. In particular, we present in this chapter new parametric families of DPPs that cover a large range of DPPs, from the stationary Poisson process to the most repulsive DPP. In Chapter 4, we prove a mixing properties for DPPs, namely Brillinger mixing, and give some of its statistical consequences. We finally prove in Chapter 5 consistency and asymptotic normality of the parametric estimators obtained by the minimum contrast method based on the Ripley's K-function and the pcf.

2.3.1 Chapter 3

In this chapter we study how repulsive stationary DPPs can be. To this end, we introduce two definitions of the repulsiveness of a stationary point process, named global and local repulsiveness, that are both based on the pcf.

Definition. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two stationary point processes with the same intensity ρ and respective pair correlation function $g_{\mathbf{X}}$ and $g_{\mathbf{Y}}$.*

- *Assuming that both $(1-g_{\mathbf{X}})$ and $(1-g_{\mathbf{Y}})$ are integrable, we say that $\mathbf{X}$ is globally more repulsive than $\mathbf{Y}$ if $\int(1-g_{\mathbf{X}}) \geq \int(1-g_{\mathbf{Y}})$.*
- *Assuming that $g_{\mathbf{X}}$ is twice differentiable at 0 with $g_{\mathbf{X}}(0) = 0$, we say that $\mathbf{X}$ is more locally repulsive than $\mathbf{Y}$ if either $g_{\mathbf{Y}}(0) > 0$, or $g_{\mathbf{Y}}$ is not twice differentiable at 0, or $g_{\mathbf{Y}}$ is twice differentiable at 0 with $g_{\mathbf{Y}}(0) = 0$ and $\Delta g_{\mathbf{Y}}(0) \geq \Delta g_{\mathbf{X}}(0)$.*

While the global repulsiveness deals with the behaviour of g over the whole space $\mathbb{R}^d$, the local repulsiveness quantifies the flatness of the pcf at the origin (the flatter, the more repulsive). For instance, the hardcore point processes are the extreme cases where g equals to 0 on a neighbourhood of the origin.

The global repulsiveness has been introduced in [44] where the authors find an infinity of DPPs that may be considered as the most repulsive DPP (in the global sense). By looking at the local repulsiveness, we have proved the following proposition (see Theorem 3.3.2 and Corollary 3.3.3) that gives the existence of a unique stationary DPP on $\mathbb{R}^d$ which is the most locally repulsive. Moreover, it is one of the most globally repulsive DPPs. For ν real, we denote by J_ν the Bessel function of the first kind of order ν.

Proposition. *The DPP with kernel*

$$C_B(x) = \frac{\sqrt{\rho\Gamma(\frac{d}{2}+1)}}{\pi^{d/4}} \frac{J_{\frac{d}{2}}\left(2\sqrt{\pi}\Gamma(\frac{d}{2}+1)^{\frac{1}{d}}\rho^{\frac{1}{d}}|x|\right)}{|x|^{\frac{d}{2}}}, \quad \forall x \in \mathbb{R}^d,$$

is the unique DPP with kernel verifying $\mathcal{K}(\rho)$ which is, both globally and locally, the most repulsive DPP among all stationary DPPs with intensity ρ.

In particular, stationary DPPs can not include a hardcore distance and so are not as flexible as Gibbs point processes. We nonetheless illustrate in Chapter 3 that DPPs models still remain quite flexible.

Additionally, we study in Chapter 3 repulsiveness for the particular class of R-dependent DPPs. We say that a point process $\mathbf{X}$ is R-dependent, for a given $R > 0$, if $\mathbf{X} \cap A$ and $\mathbf{X} \cap B$ are independent, for A and B two Borel sets in $\mathbb{R}^d$ separated by a distance larger than R. In the DPP's case, this is equivalent to DPP's with finite range kernels C, i.e. $C(x) = 0$ for $|x| \geq R$. This study has revealed to be challenging and we give only partial results on the most repulsive R-dependent DPP, for a given $R > 0$.

We conclude this chapter by giving new parametric families of DPPs as the already mentioned Bessel-type family defined in (2.2.2). For these families, we provide a closed form expression of the DPP's kernel and of its Fourier transform. This is particularly convenient for the simulation algorithm implemented in `spatstat`. Moreover these parametric families are sufficiently flexible, with respect to their parameters, to cover a wide range of DPPs from the Poisson point process (the case of no interaction) to the most repulsive DPP (as determined in the last proposition).

2.3.2 Chapter 4

A stationary point process is Brillinger mixing if for any $k \geq 2$ the total variation of the reduced factorial cumulant measure of order k is finite. In this chapter, we recall the definition of these measures and in case of a DPP with kernel C, we relate them to C. This allows us to prove the following theorem (see Theorem 4.2.2).

Theorem. *A DPP with kernel verifying the condition $\mathcal{K}(\rho)$, for a given $\rho > 0$, is Brillinger mixing.*

As a consequence of the Brillinger mixing, we prove a central limit theorem on a wide class of functionals of order p of a stationary point process, for a given $p \geq 1$, generalizing a theorem given in [40]. In particular, this theorem applies to stationary DPPs with kernel verifying the condition $\mathcal{K}(\rho)$ and the asymptotic variance is known in closed form. As an application of this theorem when $p = 1$, we retrieve a well known result for the natural estimator of the intensity, see [57] and Corollary 4.3.3.

Theorem. *Let $\mathbf{X}$ be a DPP with kernel verifying the condition $\mathcal{K}(\rho)$ for a given $\rho > 0$ and $\{D_n\}_{n\in\mathbb{N}}$ be a family of regular sets tending to $\mathbb{R}^d$ as $n \to \infty$. Define for all $n \in \mathbb{N}$,*

$$\widehat{\rho}_n = \frac{1}{|D_n|} \sum_{x\in\mathbf{X}} \mathbf{1}_{\{x\in D_n\}}.$$

We have the convergence

$$\sqrt{|D_n|}\,(\widehat{\rho}_n - \rho) \xrightarrow[n\to+\infty]{distr.} N(0,\sigma^2)$$

where $\sigma^2 = \lim_{n\to+\infty} Var\left(\sqrt{|D_n|}\widehat{\rho}_n\right) = \rho - \int_{\mathbb{R}^d} C(x)^2 dx$.

Several other applications of Brillinger mixing for a stationary point process may be found in [26, 27, 29, 36, 37] and may therefore be applied to DPPs satisfying $\mathcal{K}(\rho)$ for a given $\rho > 0$. For instance, we deduce several results concerning the following kernel estimator of the pcf. Let $\mathbf{X}$ be a stationary and isotropic DPP, we consider the estimator with kernel k and bandwidth $\{b_n\}_{n\in\mathbb{N}}$ of the pcf,

$$\widehat{g}_n(t) = \frac{1}{\sigma_d t^{d-1}\widehat{\rho}_n^2} \sum_{\substack{(x,y)\in\mathbf{X}^2\\ x\neq y}} \mathbf{1}_{\{x\in D_n,\ y\in D_n\}} \frac{1}{b_n|D_n\cap D_n^{x-y}|} k\left(\frac{t-|x-y|}{b_n}\right)$$

where $\sigma_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ denotes the surface-area of the d-dimensional unit sphere, $\{D_n\}_{n\in\mathbb{N}}$ is a sequence of subset of $\mathbb{R}^d$ and for all $n \in \mathbb{N}$ and $z \in \mathbb{R}^d$, $D_n^z := \{u, u+z \in D_n\}$. The asymptotic normality of this estimator follows directly from [27] and the Brillinger mixing of DPPs (see Proposition 4.3.4).

Proposition. *Assume that $\{D_n\}_{n\in\mathbb{N}}$ is a regular sequence of subsets of $\mathbb{R}^d$ such that $b_n^3|D_n| \to +\infty$ and $b_n^5|D_n| \to 0$. Let k be a symmetric and bounded function with compact support included in $[-T,T]$, for a given $T > 0$, and $\int_{\mathbb{R}} k(x)dx = 1$. Let C be a twice differentiable kernel on $\mathbb{R}^d \setminus \{0\}$ such that it verifies the condition $\mathcal{K}(\rho)$ for a given $\rho > 0$ and $DPP(C)$ is isotropic. Denoted g the pcf of $DPP(C)$, we have for all $t > 0$ the convergence*

$$\sqrt{b_n|D_n|}\,(\widehat{g}_n(t) - g(t)) \xrightarrow[n\to+\infty]{distr.} N(0,\tau^2)$$

where $\tau^2 = 2\frac{g(t)}{\rho^2\sigma_d t^{d-1}}\sqrt{\int_{\mathbb{R}} k^2(x)dx}$.

A similar conclusion concerning the asymptotic normality of the integrated squared error of $\widehat{g}_n$ is also given in Chapter 4, using some results of [26].

2.3.3 Chapter 5

We study in this chapter the asymptotic properties of a minimum contrast procedure to fit DPPs to a point pattern. We consider a parametric family of DPPs with kernel $C_{\rho,\theta}$ where $\rho = C_{\rho,\theta}(0) > 0$ and θ belongs to a subset Θ_ρ of $\mathbb{R}^p$ for a given $p \geq 1$ such that for any $\rho > 0$ and any $\theta \in \Theta_\rho$, $C_{\rho,\theta}$ satisfies $\mathcal{K}(\rho)$. We assume that for a given $\rho_0 > 0$ and θ_0 in the interior of Θ_{ρ_0} (provided this interior is non-empty) we observe $\mathbf{X} \sim DPP(C_{\rho_0,\theta_0})$ on a sequence $\{D_n\}_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}^d$ that grows to $\mathbb{R}^d$ in all directions.

The estimation of ρ_0 is carried out by $\widehat{\rho}_n$ discussed in Chapter 4 and we focus on the estimation of θ_0 by the following minimum contrast procedure. For $\rho > 0$ and $\theta \in \Theta_\rho$, let $J(.,\theta)$ be a function from $\mathbb{R}^d$ into $\mathbb{R}^+$ which is a summary statistic of $DPP(C_{\rho,\theta})$ that does not depend on ρ. In the DPP's case, the most important and natural examples are the K-function and the pcf g. Then, for all $n \in \mathbb{N}$, consider $\widehat{J}_n$ an estimator of J from the observation of $\mathbf{X}$ on D_n and define the discrepancy measure

$$U_n(\theta) = \int_{r_{min}}^{r_{max}} w(t) \left\{ \widehat{J}_n(t)^c - J(t,\theta)^c \right\}^2 dt$$

where $r_{max} > r_{min} \geq 0$ and $c \neq 0$ are parameters and w is a weight function generally chosen to be smooth on (r_{min}, r_{max}) or at least integrable. Finally, we consider the following estimator of θ_0,

$$\widehat{\theta}_n = \underset{\theta\in\Theta_{\hat{\rho}_n}}{\arg\min}\, U_n(\theta).$$

We prove in Chapter 5 a general theorem dealing with the asymptotic properties of $\widehat{\theta}_n$ for an arbitrary stationary point process. Applied to the DPPs for the important particular cases when J is the Ripley's K-function or the pcf g, we obtain the following theorem (see Theorems 5.2.2 and 5.2.3).

Theorem. *Let $\mathbf{X}$ be a DPP with kernel C_{ρ_0,θ_0} for a given $\rho_0 > 0$ and θ_0 an interior point of $\Theta_{\rho_0} \subset \mathbb{R}^p$, for $p \geq 1$. Assume that for all $\theta \in \Theta_{\rho_0}$, $C_{\rho_0,\theta}$ verifies the condition $\mathcal{K}(\rho_0)$ and is twice differentiable with respect to θ. If $J = K$ or $J = g$, we have under supplementary mild assumptions, the convergence*

$$\sqrt{|D_n|}(\widehat{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(0,\Sigma)$$

where Σ is completely known with respect to C_{ρ_0,θ_0}.

3

Repulsiveness of stationary determinantal point processes

This chapter is an article accepted for publication in Bernoulli journal so some considerations and definitions are redundant with the introduction.

We consider two ways to quantify the repulsiveness of a point process, both based on its second order properties, and we address the question of how repulsive a stationary DPP can be. We determine the most repulsive stationary DPP, when the intensity is fixed, and for a given $R > 0$ we investigate repulsiveness in the subclass of R-dependent stationary DPPs, i.e. stationary DPPs with R-compactly supported kernels. Finally, in both the general case and the R-dependent case, we present some new parametric families of stationary DPPs that can cover a large range of DPPs, from the stationary Poisson process (the case of no interaction) to the most repulsive DPP.

3.1 Introduction

Determinantal point processes (DPPs) were introduced in their general form by O. Macchi in 1975 [46] to model fermions in quantum mechanics, though some specific DPPs appeared much earlier in random matrix theory. DPPs actually arise in many fields of probability and have deserved a lot of attention from a theoretical point of view, see for instance [32] and [56].

DPPs are repulsive (or regular, or inhibitive) point processes, meaning that nearby points of the process tend to repel each other (this concept will be clearly described in the following). This property is adapted to many statistical problems where DPPs have been recently used, for instance in telecommunication to model the locations of network nodes [11, 47] and in statistical learning to construct a dictionary of diverse sets [41]. Other examples arising from biology, ecology and forestry are studied in [44] and its associated on-line supplementary file [43].

The growing interest for DPPs in the statistical community is due to that their moments are explicitly known, parametric families can easily been considered, their density on any compact set admits a closed form expression making likelihood inference feasible and they can be simulated easily and quickly. Section 3.2 summarizes some of these properties and we refer to [44] for a detailed presentation. These features make the class of DPPs a competitive alternative to the usual class of models for repulsiveness, namely the Gibbs point processes. In contrast, for Gibbs point processes, no closed form expression is available for the moments, the likelihood involves an intractable normalizing constant and their simulation requires Markov chain Monte Carlo methods.

However, DPPs can not model all kinds of repulsive point patterns. For instance, as deduced from Section 3.3, stationary DPPs can not involve a hardcore distance between points, contrary to the Matérn hardcore point processes, the RSA (random sequential absorption) model and hardcore Gibbs models, see [33, Section 6.5]. In this paper, we address the question of how repulsive a stationary DPP can be. We also investigate for a given $R > 0$ the repulsiveness in the subclass of R-dependent stationary DPPs, i.e. stationary DPPs with R-compactly supported kernels, which are of special interest for statistical inference in high dimension, see Section 3.4. In both cases, we present in Section 3.5 some parametric families of stationary DPPs that cover a large range of DPPs, from the stationary Poisson process to the most repulsive DPP.

To quantify the repulsiveness of a stationary point process, we consider its second-order properties. Let $\mathbf{X}$ be a stationary point process in $\mathbb{R}^d$ with intensity (i.e. expected number of points per unit volume) $\rho > 0$ and second order intensity function $\rho^{(2)}(x, y)$. Denoting dx an infinitesimal region around x and $|dx|$ its Lebesgue measure, $\rho|dx|$ may be interpreted as the probability that $\mathbf{X}$ has a point in dx. For $x \neq y$, $\rho^{(2)}(x, y)|dx||dy|$ may be viewed as the probability that $\mathbf{X}$ has a point in dx and another point in dy. A formal definition is given in Section 3.2. Note that $\rho^{(2)}(x, y) = \rho^{(2)}(0, y - x)$ is a symmetric function and depends only on $y - x$ because of our stationarity assumption.

In spatial statistics, the second order properties of $\mathbf{X}$ are generally studied through the pair correlation function (in short pcf), defined for any $x \in \mathbb{R}^d$ by

$$g(x) = \frac{\rho^{(2)}(0, x)}{\rho^2}.$$

Since $\rho^{(2)}$ is unique up to a set of Lebesgue measure zero (see [9]), so is g. As it is implicitly done in the literature, see [33, 59], we choose the version of g with as few discontinuity points as possible. It is commonly accepted, see for example [59], that if $g(x) = 1$ then there is no interaction between two points separated by x, whereas there is attraction if $g(x) > 1$ and repulsiveness if $g(x) < 1$. Therefore, when we below compare the global repulsiveness of two stationary point processes we assume they share the same intensity.

Definition 3.1.1. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two stationary point processes with the same intensity ρ and respective pair correlation function $g_{\mathbf{X}}$ and $g_{\mathbf{Y}}$. Assuming that both $(1 - g_{\mathbf{X}})$ and $(1 - g_{\mathbf{Y}})$ are integrable, we say that $\mathbf{X}$ is globally more repulsive than $\mathbf{Y}$ if $\int(1 - g_{\mathbf{X}}) \geq \int(1 - g_{\mathbf{Y}})$.*

The quantity $\int(1 - g)$ is already considered in the on-line supplementary material [43] of [44] as a measure for repulsiveness. It can be justified in several ways. First, it is a natural geometrical method to quantify the distance from g to 1 (corresponding to no interaction), where the area between g and 1 contributes positively to the measure of repulsiveness when $g < 1$ and negatively if $g > 1$. Second, denoting K and K_0 the Ripley's K-functions of $\mathbf{X}$ and of the stationary Poisson process with intensity ρ respectively, see [49, Definition 4.6], we have $\int(1-g) = \lim_{r\to\infty}(K_0(r) - K(r))$. We also refer to [43] for an equivalent interpretation in terms of the reduced Palm distribution. Finally, it is worth mentioning that for any stationary point processes, we have $\int(1-g) \leq 1/\rho$, see [42, Equation (2.5)].

Additional criteria could be introduced to quantify the global repulsiveness of a point process, relying for instance on $\int(1-g)^p$ for a given $p > 0$, or involving higher moments of the point process through the joint intensities of order $k > 2$ (see Definition 3.2.1). However the theoretical study becomes more challenging in these cases and we do not consider these extensions.

Repulsiveness is often interpreted in a local sense: This is the case for hardcore point processes, where a minimal distance δ is imposed between points and so $g(x) = 0$ whenever $|x| < \delta$ where for a vector x, $|x|$ denotes its Euclidean norm. As already mentioned, a DPP can not involve any hardcore distance, but we may want its pcf to satisfy $g(0) = 0$ and stay as close as possible to 0 near the origin. This leads to the following criteria to compare the *local repulsiveness* of two point processes. We denote by ∇g and Δg the gradient and the Laplacian of g, respectively.

Definition 3.1.2. *Let $\mathbf{X}$ and $\mathbf{Y}$ be two stationary point processes with the same intensity ρ and respective pair correlation function $g_{\mathbf{X}}$ and $g_{\mathbf{Y}}$. Assuming that $g_{\mathbf{X}}$ is twice differentiable at 0 with $g_{\mathbf{X}}(0) = 0$, we say that $\mathbf{X}$ is more locally repulsive than $\mathbf{Y}$ if either $g_{\mathbf{Y}}(0) > 0$, or $g_{\mathbf{Y}}$ is not twice differentiable at 0, or $g_{\mathbf{Y}}$ is twice differentiable at 0 with $g_{\mathbf{Y}}(0) = 0$ and $\Delta g_{\mathbf{Y}}(0) \geq \Delta g_{\mathbf{X}}(0)$.*

As suggested by this definition, a stationary point process is said to be locally repulsive if its pcf is twice differentiable at 0 with $g(0) = 0$. In this case $\nabla g(0) = 0$ because $g(x) = g(-x)$. Therefore to compare the behavior of two such pcfs near the origin, specifically the curvatures of their graphs near the origin, the Laplacian operator is involved in Definition 3.1.2. As an example, a stationary hardcore process is locally more repulsive than any other stationary point process because in this case $g(0) = 0$ and $\Delta g(0) = 0$.

We show in Section 3.3 that Definitions 3.1.1 and 3.1.2 agree for the natural choice of what can be considered as the most repulsive DPP. A realization of the latter on $[-5, 5]^2$ is represented in Figure 3.1 (d) when $\rho = 1$. For comparison, letting $\rho = 1$ for all plots, Figure 3.1 shows realizations of: (a) the stationary Poisson process, which is a situation with no interaction; (b)-(c) two DPPs with intermediate repulsiveness, namely DPPs with kernels (3.5.1) where $\sigma = 0$ and $\alpha = 0.2, 0.4$ respectively, as presented in Section 3.5.1; (e) the type II Matérn hardcore process with hardcore radius $\frac{1}{\sqrt{\pi}}$. Notice that $\frac{1}{\sqrt{\pi}}$ is the maximal hardcore radius that a type II Matérn hardcore process with unit intensity can reach, see [33, Section 6.5]. It corresponds to an infinite intensity of the underlying Poisson process and our simulation is only an approximation. These models are sorted from (a) to (e) by their ascending repulsiveness in the sense of Definition 3.1.2. Specifically $g(0) = 1$ for (a) while $g(0) = 0$ and $\Delta g(0)$ is $50,\ 12.5,\ 2\pi$ and 0 from (b) to (e) respectively. This order is clearly apparent in Figure 3.1 (f), where the theoretical pcfs are represented as radial functions, all aforementioned models being isotropic. Concerning global repulsiveness, we have that $\int(1-g)$ is $0,\ 0.12,\ 0.50,\ 1$ and 0.76 from (a) to (e) respectively. The fact that the Matérn hardcore model is globally less repulsive than the DPP in (d) is due to that its pcf can be larger than one. This shows the limitation of Definition 3.1.1 in the study of repulsiveness and the importance of introducing Definition 3.1.2. Overall, Figure 3.1 illustrates that even if stationary DPPs cannot be as (locally) repulsive as hardcore point processes, which may be an important limitation in practice, they nonetheless cover a rather large variety of repulsiveness from (a) to (d).

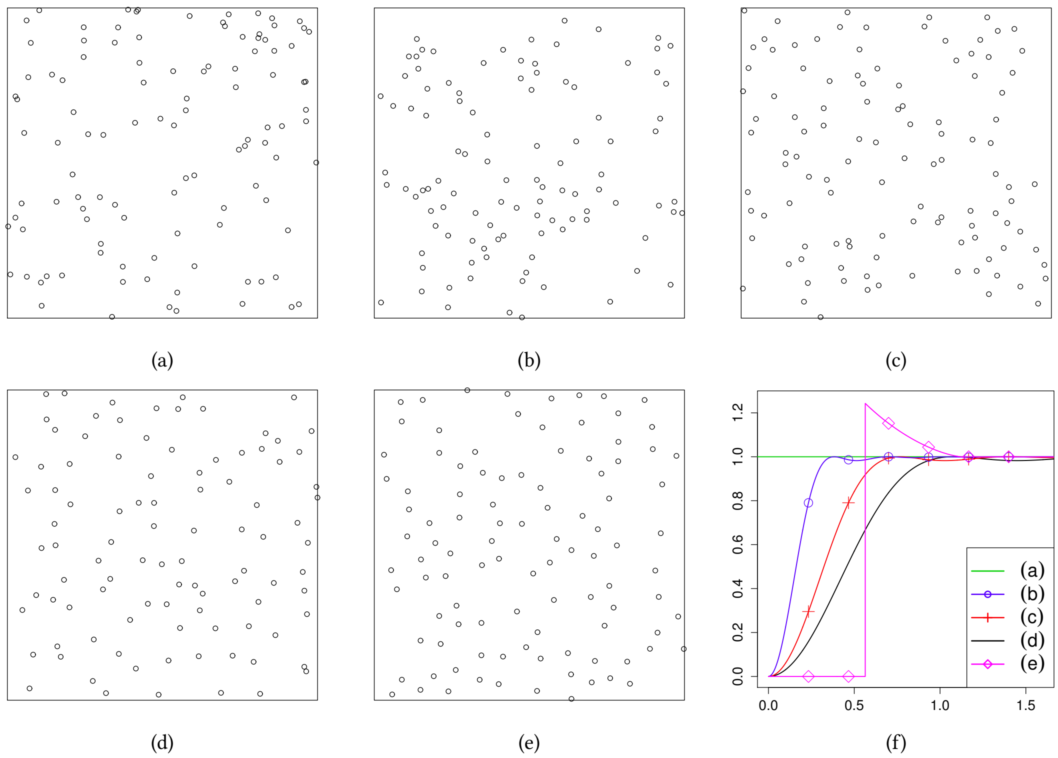


Figure 3.1 – Realizations on $[-5,5]^2$ of (a) the stationary Poisson process, (b)-(d) DPPs with kernels (3.5.1) where $\sigma = 0$ and $\alpha = 0.2, 0.4, \frac{1}{\sqrt{\pi}}$, (e) the type II Matérn hardcore process with hardcore radius $\frac{1}{\sqrt{\pi}}$. (f) Their associated theoretical pcfs. The intensity is $\rho = 1$ for all models and (d) represents the most repulsive stationary DPP in this case.

We recall the definition of a stationary DPP and some related basic results in Section 3.2. Section 3.3 is devoted to the study of repulsiveness in stationary DPPs, both in the sense of Definition 3.1.1 and Definition 3.1.2. In Section 3.4, we focus on repulsiveness for the subclass of stationary DPPs with compactly supported kernels. Then, in Section 3.5, we present three parametric families of DPPs which cover a large range of repulsiveness and have further interesting properties. Section 3.6 gathers the proofs of our theoretical results. Further comments and illustrations are provided in a supplementary material.

3.2 Stationary DPPs

In this section, we review the definition and some properties of stationary DPPs. For a detailed presentation, including the non stationary case, we refer to the survey by Hough et al. [32].

Basics of point processes may be found in [9, 10]. Let us recall that a point process $\mathbf{X}$ is simple if two points of $\mathbf{X}$ never coincide, almost surely. The joint intensities of $\mathbf{X}$ are defined as follows.

Definition 3.2.1. *If it exists, the joint intensity of order k ($k \geq 1$) of a simple point process $\mathbf{X}$ is the function $\rho^{(k)} : (\mathbb{R}^d)^k \to \mathbb{R}^+$ such that for any family of mutually disjoint subsets $D_1, \dots, D_k$ in $\mathbb{R}^d$,*

$$E \prod_{i=1}^{k} \mathbf{X}(D_i) = \int_{D_1} \dots \int_{D_k} \rho^{(k)}(x_1, \dots, x_k) dx_1 \dots dx_k,$$

where $\mathbf{X}(D)$ denotes the number of points of $\mathbf{X}$ in D and E is the expectation over the distribution of $\mathbf{X}$.

In the stationary case, $\rho^{(k)}(x_1, \dots, x_k) = \rho^{(k)}(0, x_2 - x_1, \dots, x_k - x_1)$, so that the intensity ρ and the second order intensity function $\rho^{(2)}$ introduced previously become the particular cases associated to $k = 1$ and $k = 2$ respectively.

Definition 3.2.2. *Let $C : \mathbb{R}^d \to \mathbb{R}$ be a function. A point process $\mathbf{X}$ on $\mathbb{R}^d$ is a stationary DPP with kernel C, in short $\mathbf{X} \sim DPP(C)$, if for all $k \geq 1$ its joint intensity of order k satisfies the relation*

$$\rho^{(k)}(x_1, \dots x_k) = \det[C](x_1, \dots, x_k)$$

for almost every $(x_1, \dots, x_k) \in (\mathbb{R}^d)^k$, where $[C](x_1, \dots, x_k)$ denotes the matrix with entries $C(x_i - x_j)$, $1 \leq i, j \leq k$.

It is actually possible to consider a complex-valued kernel C, but for simplicity we restrict ourselves to the real case. A first example of stationary DPP is the stationary Poisson process with intensity ρ. It corresponds to the kernel

$$C(x) = \rho \mathbf{1}_{\{x=0\}}, \ \forall x \in \mathbb{R}^d. \tag{3.2.1}$$

However, this example is very particular and represents in some sense the extreme case of a DPP without any interaction, while DPPs are in general repulsive as discussed at the end of this section.

Definition 3.2.2 does not ensure existence or unicity of $DPP(C)$, but if it exists, then it is unique, see [32]. Concerning existence, a general result, including the non stationary case, was proved by O. Macchi in [46]. It relies on the Mercer representation of C on any compact set. Unfortunately this representation is known only in a few cases, making the conditions impossible to verify in practice for most functions C. Nevertheless, the situation becomes simpler in our stationary framework, where the conditions only involve the Fourier transform of C. We define the Fourier transform of a function $h \in L^1(\mathbb{R}^d)$ as

$$\mathcal{F}(h)(t) = \int_{\mathbb{R}^d} h(x) e^{2i\pi x \cdot t} dx, \quad \forall t \in \mathbb{R}^d. \tag{3.2.2}$$

By Plancherel's theorem, this definition is extended to $L^2(\mathbb{R}^d)$, see [58]. If C is a covariance function, as assumed in the following, we have $\mathcal{F}\mathcal{F}(C) = C$ so $\mathcal{F}^{-1} = \mathcal{F}$ and from [51, Theorem 1.8.13], $\mathcal{F}(C)$ belongs to $L^1(\mathbb{R}^d)$.

Proposition 3.2.3 ([44]). *Assume C is a symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$. Then $DPP(C)$ exists if and only if $0 \leq \mathcal{F}(C) \leq 1$.*

In other words, Proposition 3.2.3 ensures existence of $DPP(C)$ if C is a continuous real-valued covariance function in $L^2(\mathbb{R}^d)$ with $\mathcal{F}(C) \leq 1$. Henceforth, we assume the following condition.

Condition $\mathcal{K}(\rho)$. A kernel C is said to verify condition $\mathcal{K}(\rho)$ if C is a symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$ with $C(0) = \rho$ and $0 \leq \mathcal{F}(C) \leq 1$.

The assumption $0 \leq \mathcal{F}(C) \leq 1$ is in accordance with Proposition 3.2.3, while the others assumptions in condition $\mathcal{K}(\rho)$ are satisfied by most statistical models of covariance functions, the main counterexample being (3.2.1). Standard parametric families of kernels include the Gaussian, the Whittle-Matérn and the generalized Cauchy covariance functions, where the condition $\mathcal{F}(C) \leq 1$ implies some restriction on the parameter space, see [44].

By Definition 3.2.2, all moments of a DPP are explicitly known. In particular, assuming condition $\mathcal{K}(\rho)$, the intensity of $DPP(C)$ is ρ and denoting g its pcf we have

$$1 - g(x) = \frac{C(x)^2}{\rho^2} \tag{3.2.3}$$

for almost every $x \in \mathbb{R}^d$. Consequently $g \leq 1$, and so we have repulsiveness. Moreover, the study of repulsiveness of stationary DPPs, as defined in Definitions 3.1.1 and 3.1.2, reduces to considerations on the kernel C when condition $\mathcal{K}(\rho)$ is assumed.

3.3 Most repulsive DPPs

We first present the most globally repulsive DPPs in the sense of Definition 3.1.1. They are introduced in the on-line supplementary file associated to [44], see [43], from which the following proposition is easily deduced.

Proposition 3.3.1 ([44]). *In the sense of Definition 3.1.1, $DPP(C)$ is the most globally repulsive DPP among all DPPs with kernel satisfying condition $\mathcal{K}(\rho)$ if and only if $\mathcal{F}(C)$ is even and equals almost everywhere an indicator function of a Borel set with volume ρ.*

According to Proposition 3.3.1, the set of the most globally repulsive DPPs in the sense of Definition 3.1.1 is infinite. This is illustrated in the supplementary material. A natural choice is $DPP(C_B)$ where $\mathcal{F}(C_B)$ is the indicator function of the Euclidean ball centered at 0 with volume ρ. In dimension d, this gives $C_B = \mathcal{F}\left(\mathbf{1}_{\{|\cdot|^d \leq \rho\tau^d\}}\right)$ with $\tau = \left\{\Gamma(d/2+1)/\pi^{\frac{d}{2}}\right\}^{\frac{1}{d}}$ and by [22, Appendix B.5],

$$C_B(x) = \frac{\sqrt{\rho\Gamma(\frac{d}{2}+1)}}{\pi^{d/4}} \frac{J_{\frac{d}{2}}\left(2\sqrt{\pi}\Gamma(\frac{d}{2}+1)^{\frac{1}{d}}\rho^{\frac{1}{d}}|x|\right)}{|x|^{\frac{d}{2}}}, \ \forall x \in \mathbb{R}^d, \tag{3.3.1}$$

where $J_{\frac{d}{2}}$ is the Bessel function of the first kind. For example, we have

- for $d = 1$, $C_B(x) = \operatorname{sinc}(x) = \frac{\sin(\pi\rho|x|)}{\pi|x|}$,
- for $d = 2$, $C_B(x) = \operatorname{jinc}(x) = \sqrt{\rho}\frac{J_1(2\sqrt{\pi\rho}|x|)}{\sqrt{\pi}|x|}$.

This choice was already favored in [44]. However, there is no indication from Proposition 3.3.1 to suggest C_B instead of another kernel given by the proposition. This choice becomes clear if we look at the local repulsiveness as defined in Definition 3.1.2.

Proposition 3.3.2. *In the sense of Definition 3.1.2, the most locally repulsive DPP among all DPPs with kernel satisfying condition $\mathcal{K}(\rho)$ is $DPP(C_B)$.*

Thus, from Propositions 3.3.1 and 3.3.2, we deduce the following corollary.

Corollary 3.3.3. *The kernel C_B is the unique kernel C verifying condition $\mathcal{K}(\rho)$ such that $DPP(C)$ is both the most globally and the most locally repulsive DPP among all stationary DPPs with intensity $\rho > 0$.*

Borodin and Serfaty in [8] characterize in dimension $d \leq 2$ the disorder of a point process by its "renormalized energy". In fact, the smaller the renormalized energy, the more repulsive the point process. Theorem 3 in [8] establishes that $DPP(C_B)$ minimizes the renormalized energy among all stationary DPPs. This result confirms Corollary 3.3.3, that the most repulsive stationary DPP, if any has to be chosen, is $DPP(C_B)$. However, a stationary DPP has a finite renormalized energy if and only if it is given by Proposition 3.3.1 [8, Theorem 1], which indicates that most stationary DPPs have an infinite renormalized energy. Hence this criteria is not of practical use to compare the repulsiveness between two arbitrary DPPs.

3.4 Most repulsive DPPs with compactly supported kernels

In this section, we assume that the kernel C is compactly supported, i.e. there exists $R > 0$ such that $C(x) = 0$ if $|x| > R$. In this case $\mathbf{X} \sim DPP(C)$ is an R-dependent point process in the sense that if A and B are two Borel sets in $\mathbb{R}^d$ separated by a distance larger than R, then $\mathbf{X} \cap A$ and $\mathbf{X} \cap B$ are independent, which is easily verified using Definition 3.2.2. This situation can be particularly interesting for likelihood inference in presence of a large number of points. Assume we observe $\{x_1, \dots, x_n\}$ on a compact window $W \subset \mathbb{R}^d$, then the likelihood is proportional to $\det[\tilde{C}](x_1, \dots, x_n)$ where $\tilde{C}$ expresses in terms of C and inherits the compactly supported property of C, see [44, 46]. While this determinant is expensive to compute if $\tilde{C}$ is not compactly supported and n is large, the situation becomes more convenient in the compactly supported case, since $[\tilde{C}](x_1, \dots, x_n)$ is sparse when R is small with respect to the size of W. We are thus interested in DPPs with kernels satisfying the following condition.

Condition $\mathcal{K}_c(\rho, R)$. A kernel C or $DPP(C)$ is said to verify condition $\mathcal{K}_c(\rho, R)$ if C verifies condition $\mathcal{K}(\rho)$ and C is compactly supported with range R, i.e. $C(x) = 0$ for $|x| \geq R$.

The following proposition shows that any kernel satisfying condition $\mathcal{K}(\rho)$ can be arbitrarily approximated by kernels verifying $\mathcal{K}_c(\rho, r)$ for r large enough. We define the function h by

$$h(x) = \exp\left(\frac{1}{|x|^2 - 1}\right) \mathbf{1}_{\{|x|<1\}}, \quad \forall x \in \mathbb{R}^d. \tag{3.4.1}$$

For a function $f \in L^2(\mathbb{R}^d)$, put $\|f\| = \sqrt{\int |f(t)|^2 dt}$ and denote $[f * f]$ the self-convolution product of f.

Proposition 3.4.1. *Let C be a kernel verifying condition $\mathcal{K}(\rho)$ and h be defined by (3.4.1). Then, for all $r > 0$, the function C_r defined by*

$$C_r(x) = \frac{1}{\|h\|^2} [h * h] \left(\frac{2x}{r}\right) C\left(x\right), \quad \forall x \in \mathbb{R}^d, \tag{3.4.2}$$

verifies $\mathcal{K}_c(\rho, r)$. Moreover, we have the convergence

$$\lim_{r \to +\infty} C_r = C, \tag{3.4.3}$$

uniformly on all compact sets.

In particular, by taking $C = C_B$ in Proposition 3.4.1, it is always possible to find a kernel C_r verifying $\mathcal{K}_c(\rho, r)$ that yields a repulsiveness (local or global) as close as we wish to the repulsiveness of C_B, provided that r is large enough. However, given a maximal range of interaction R, it is clear that the maximal repulsiveness implied by kernels verifying $\mathcal{K}_c(\rho, R)$ can not reach the one of C_B, since the support of C_B is unbounded and $DPP(C_B)$ is the unique most repulsive DPP according to Corollary 3.3.3. In the following, we study the DPP's repulsiveness for a given range $R > 0$.

In comparison with condition $\mathcal{K}(\rho)$, the assumption that C is compactly supported in condition $\mathcal{K}_c(\rho, R)$ makes the optimization problems related to Definitions 3.1.1-3.1.2 much more difficult to investigate. As a negative result, we know very little about the most globally repulsive DPP, in the sense of Definition 3.1.1, under condition $\mathcal{K}_c(\rho, R)$. From relation (3.2.3), this is equivalent to find a kernel C with maximal L^2-norm under the constraint that C verifies $\mathcal{K}_c(\rho, R)$. Without the constraint $\mathcal{F}(C) \leq 1$, this problem is known as the square-integral Turán problem with range R, see for example [39]. For this less constrained problem, it is known that a solution exists, but no explicit formula is available, cf. [12]. For $d = 1$, it has been proved that the solution is unique and there exists an algorithm to approximate it, see [19]. In this case, numerical approximations show that the solution with range R verifies condition $\mathcal{K}_c(\rho, R)$ only if $R \leq 1.02/\rho$. This gives the most globally repulsive DPP verifying $\mathcal{K}_c(\rho, R)$ in dimension $d = 1$, when $R \leq 1.02/\rho$, albeit without explicit formula. Its pcf is represented in Figure 3.2. For other values of R, or in dimension $d \geq 2$, no results are available, to the best of our knowledge.

Let us now turn to the investigation of the most locally repulsive DPP, in the sense of Definition 3.1.2, under condition $\mathcal{K}_c(\rho, R)$. Recall that without the compactly supported constraint of the kernel, we showed in Section 3.3 that the most locally repulsive DPP, namely $DPP(C_B)$, is also (one of) the most globally repulsive DPP.

For $\nu > 0$, we denote by j_ν the first positive zero of the Bessel function J_ν and by J'_ν the derivative of J_ν. We refer to [1] for a survey about Bessel functions and their zeros. Further, define the constant $M > 0$ by

$$M^d = \frac{2^{d-2} j^2_{\frac{d-2}{2}} \Gamma\left(\frac{d}{2}\right)}{\rho \pi^{\frac{d}{2}}}.$$

We have $M\rho = \pi^2/8 \approx 1.234$ when $d = 1$, $M\rho^{1/2} = j_0/\pi^{1/2} \approx 1.357$ when $d = 2$ and $M\rho^{1/3} = \pi^{1/3} \approx 1.465$ when $d = 3$.

Proposition 3.4.2. *If $R \leq M$, then, in the sense of Definition 3.1.2, there exists an unique isotropic kernel C_R such that $DPP(C_R)$ is the most locally repulsive DPP among all $DPPs$ with kernel verifying $\mathcal{K}_c(\rho, R)$. It is given by $C_R = u * u$ where*

$$u(x) = \kappa \, \frac{J_{\frac{d-2}{2}}\left(2 j_{\frac{d-2}{2}} \frac{|x|}{R}\right)}{|x|^{\frac{d-2}{2}}} \, \mathbf{1}_{\left\{|x| < \frac{R}{2}\right\}}, \tag{3.4.4}$$

with $\kappa^2 = \frac{4\Gamma(d/2)}{\rho\pi^{d/2}R^2}\left(J'_{\frac{d-2}{2}}(j_{\frac{d-2}{2}})\right)^{-2}$.

In this proposition C_R is only given as a convolution product. Nonetheless, an explicit expression is known in dimension $d = 1$ and $d = 3$, see [14]. On the other hand, the Fourier transform is known in any dimension since $\mathcal{F}(C_R) = \mathcal{F}(u)^2$. We get from the proof in Section 3.6.3, for all $x \in \mathbb{R}^d$,

$$\mathcal{F}(C_R)(x) = \rho\pi^{d/2}R^d j^2_{\frac{d-2}{2}}\Gamma\left(\frac{d}{2}\right)\left(\frac{J_{\frac{d-2}{2}}(\pi R|x|)}{(\pi R|x|)^{\frac{d-2}{2}}\left(j^2_{\frac{d-2}{2}} - (\pi R|x|)^2\right)}\right)^2. \tag{3.4.5}$$

If $R \geq M$, we have not been able to obtain a closed form expression of the most locally repulsive stationary DPP. However, under some extra regularity assumptions, we can state the following general result about its existence and the form of the solution.

Condition $\mathcal{M}(\rho, R)$. A function u is said to verify condition $\mathcal{M}(\rho, R)$ if $u(x) = 0$ for $|x| > \frac{R}{2}$, u is a radial function and $u \in L^2(\mathbb{R}^d)$ with $\|u\|^2 = \rho$.

Proposition 3.4.3. *For any $R > 0$, there exists an isotropic kernel C_R such that $DPP(C_R)$ is the most locally repulsive DPP among all $DPPs$ with kernel C verifying $\mathcal{K}_c(\rho, R)$. It can be expressed as $C_R = u * u$ where u satisfies $\mathcal{M}(\rho, R)$. Furthermore, if we assume that $\sup_{x\in\mathbb{R}^d}\mathcal{F}(C_R)(x) = \mathcal{F}(C_R)(0)$ and u is twice differentiable on its support, then u is of the form*

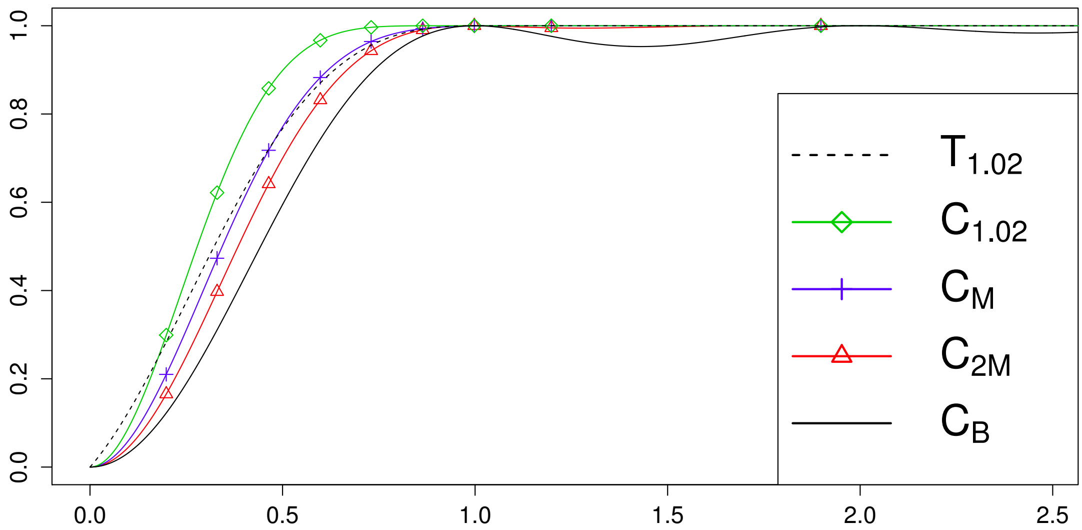
$$u(x) = \left(\beta + \gamma\frac{J_{\frac{d-2}{2}}(|x|/\alpha)}{|x|^{\frac{d-2}{2}}}\right)\mathbf{1}_{\{|x|<\frac{R}{2}\}}, \tag{3.4.6}$$

where $\alpha > 0$, $\beta \geq 0$ and γ are three constants linked by the conditions $\mathcal{M}(\rho, R)$ and $\int_{\mathbb{R}^d} u(x)dx \leq 1$.

In the case $R \leq M$, this proposition is a consequence of Proposition 3.4.2 where $\beta = 0$, $\alpha = R/(2j_{\frac{d-2}{2}})$ and $\gamma = \kappa$. When $R > M$, it is an open problem to find an explicit expression of the kernel C_R without any extra regularity assumptions. Even in this case, (3.4.6) only gives the form of the solution and the constants α, β and γ are not explicitly known. In particular the choice $\beta = 0$ does not lead to the most locally repulsive DPP when $R > M$, contrary to the case $R \leq M$. In fact, the condition $\mathcal{M}(\rho, R)$ allows us to express β and γ as functions of α, R and ρ, but then some numerical approximation are needed to find the value of α in (3.4.6), given R and ρ, such that $DPP(C_R)$ is the most locally repulsive DPP. We detail these relations in Section 3.5.3, where we start from (3.4.6) to suggest a new parametric family of compactly supported kernels.

Contrary to what happens in the non compactly supported case of Section 3.3, the most locally repulsive DPP is not the most globally repulsive DPP under $\mathcal{K}_c(\rho, R)$. This is easily checked in dimension $d = 1$ when $R \leq 1.02/\rho$ implying $R \leq M$: In this case the most globally repulsive DPP under $\mathcal{K}_c(\rho, R)$ is $DPP(T_R)$, where T_R is the solution of the square-integral Turán problem with range R, and the most locally repulsive DPP is $DPP(C_R)$ where C_R is given by (3.4.4). However, according to the results of Section 3.3 corresponding to $R = \infty$, we expect that $DPP(C_R)$ has a strong global repulsiveness even for moderate values of R. This is confirmed in Figure 3.2, that shows the pcf of $DPP(C_R)$ when $d = 1$, $\rho = 1$ and $R = 1.02$, $R = M \approx 1.234$ and $R = 2M$, where in this case we take $C_R = u * u$ with u given by (3.4.6) and the constants are obtained by numerical approximations. The pcfs of $DPP(T_{1.02})$ and $DPP(C_B)$ are added for sake of comparison. Considering the behavior of the pcf near the origin, we note that even if $DPP(T_{1.02})$ is the most globally repulsive DPP under $\mathcal{K}_c(\rho, R)$ when $R \leq 1.02/\rho$, its local repulsiveness is not very strong. On the other hand, $DPP(C_R)$ seems to present strong global repulsiveness for the values of R considered in the figure.

Figure 3.2 – In dimension $d = 1$, comparison between the pcf of $DPP(T_{1.02})$, $DPP(C_B)$ and $DPP(C_R)$ for $R = 1.02, M, 2M$.



3.5 Parametric families of DPP kernels

A convenient parametric family of kernels $\{C_\theta\}_{\theta \in \Theta}$, where $\Theta \subset \mathbb{R}^q$ for some $q \geq 1$, should ideally:

(a) provide a closed form expression for C_θ, for any θ,

(b) provide a closed form expression for $\mathcal{F}(C_\theta)$, for any θ,

(c) be flexible enough to include a large range of DPPs, going from the Poisson point process to $DPP(C_B)$.

The second property above is needed to check the condition of existence $\mathcal{F}(C_\theta) \leq 1$, but it is also useful for some approximations in practice. Indeed, the algorithm for simulating $DPP(C)$ on a compact set S, as presented in [32], relies on the Mercer representation of C on S, which is rarely known in practice. In [44], this decomposition is simply approximated by the Fourier series of C where, up to some rescaling, the k-th Fourier coefficients is replaced by $\mathcal{F}(C)(k)$. The same approximation is used to compute the likelihood. This method has proved to be accurate in most cases, both from a practical and a theoretical point of view, provided ρ is not too small, and to be computationally efficient, see [44].

In addition to (a)-(c), we may also require that C_θ is compactly supported with maximal range R, following the motivation explained in Section 3.4, in which case the maximal possible repulsiveness is given by $DPP(C_R)$. Or we may require that $\mathcal{F}(C_\theta)$ is compactly supported, in which case the Fourier series mentioned in the previous paragraph becomes a finite sum and no truncation is needed in practice. Note however that C_θ and $\mathcal{F}(C_\theta)$ can not both be compactly supported.

Several standard parametric families of kernels are available, including the well-known Whittle-Matérn and the generalized Cauchy covariance functions, where the condition $\mathcal{F}(C_\theta) \leq 1$ implies some restriction on the parameter space, see [44]. Although they encompass a closed form expression for both C_θ and $\mathcal{F}(C_\theta)$, they are not flexible enough to reach the repulsiveness of $DPP(C_B)$. Another family of parametric kernels is considered in [44], namely the power exponential spectral model, that contains as limiting cases C_B and the Poisson kernel (3.2.1). For this reason this family is more flexible than the previous ones, but then only $\mathcal{F}(C_\theta)$ is given and no closed expression is available for C_θ. For all these families, none of C_θ and $\mathcal{F}(C_\theta)$ is compactly supported.

Below, we present alternative families of parametric kernels. The first two ones, so-called Bessel-type and Laguerre-Gaussian families, fulfil the three requirements (a)-(c) above and the Bessel-type family has the additional property that the Fourier transform of the kernels is compactly supported. Moreover we introduce new families of compactly supported kernels, inspired by Proposition 3.4.1 and Proposition 3.4.3.

3.5.1 Bessel-type family

For all $\sigma \geq 0,\ \alpha > 0,\ \rho > 0$, we consider the Bessel-type kernel

$$C(x) = \rho\, 2^{\frac{\sigma+d}{2}} \Gamma\left(\frac{\sigma+d+2}{2}\right) \frac{J_{\frac{\sigma+d}{2}}\left(2|\frac{x}{\alpha}|\sqrt{\frac{\sigma+d}{2}}\right)}{\left(2|\frac{x}{\alpha}|\sqrt{\frac{\sigma+d}{2}}\right)^{\frac{\sigma+d}{2}}}, \quad x \in \mathbb{R}^d. \tag{3.5.1}$$

This positive definite function first appears in [53], where it is called the Poisson function. It has been further studied in [17] and [18], where it is called the Bessel-type function. For obvious reasons, we prefer the second terminology when applied to point processes. For any $x \in \mathbb{R}$, we denote by $x_+ = \max(x, 0)$ its positive part.

Proposition 3.5.1. *Let C be given by* (3.5.1)*, then its Fourier transform is, for all $x \in \mathbb{R}^d$,*

$$\mathcal{F}(C)(x) = \rho \frac{(2\pi)^{\frac{d}{2}} \alpha^d \Gamma(\frac{\sigma+d+2}{2})}{(\sigma+d)^{\frac{d}{2}} \Gamma(\frac{\sigma+2}{2})} \left(1 - \frac{2\pi^2\alpha^2|x|^2}{\sigma+d}\right)_+^{\frac{\sigma}{2}} \tag{3.5.2}$$

and $DPP(C)$ exists if and only if $\alpha \leq \alpha_{\max}$ where

$$\alpha_{\max}^d = \frac{(\sigma+d)^{\frac{d}{2}}\Gamma(\frac{\sigma+2}{2})}{\rho(2\pi)^{\frac{d}{2}}\Gamma(\frac{\sigma+d+2}{2})}.$$

In this case, $DPP(C)$ defines a stationary and isotropic DPP with intensity ρ. Moreover, if $\sigma = 0$ and $\alpha = \alpha_{\max}$, then $C = C_B$ where C_B is defined in (3.3.1)*. In addition, for any $\rho > 0$ and $\alpha > 0$, we have the convergence*

$$\lim_{\sigma \to +\infty} C(x) = \rho\, e^{-\left(\frac{|x|}{\alpha}\right)^2}, \tag{3.5.3}$$

uniformly on all compact sets.

The Bessel-type family contains C_B as a particular case and the Poisson kernel as a limiting case, when $\alpha \to 0$. Moreover, $\mathcal{F}(C)$ is compactly supported, see (3.5.2). The plots in Figure 3.1 (b)-(d) show some realizations of this model when $\sigma = 0$ and $\alpha = 0.2,\ 0.4,\ \alpha_{\max}$, respectively. The supplementary material includes more simulations and shows the behavior of the pcf for different values of the parameters.

3.5.2 Laguerre-Gaussian family

Let us first recall the definition of the Laguerre polynomials. We denote by $\mathbb{N}$ the set $\{1, 2, \ldots\}$. For integers $0 \le k \le m$ and numbers α, define $\binom{m+\alpha}{k} = \frac{(m+\alpha)\ldots(m+\alpha+1-k)}{k}$ if $k > 0$ and $\binom{m+\alpha}{k} = 1$ if $k = 0$.

Definition 3.5.2. *The Laguerre polynomials are defined for all $m \in \mathbb{N} \cap \{0\}$ and $\alpha \in \mathbb{R}$ by*

$$L_m^\alpha(x) = \sum_{k=0}^{m} \binom{m+\alpha}{m-k} \frac{(-x)^k}{k!}, \quad \forall x \in \mathbb{R}.$$

For all $m \in \mathbb{N},\ \alpha > 0,\ \rho > 0$ and $x \in \mathbb{R}^d$, we consider the Laguerre-Gaussian function

$$C(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} L_{m-1}^{\frac{d}{2}} \left(\frac{1}{m} \left| \frac{x}{\alpha} \right|^2 \right) e^{-\frac{1}{m}|\frac{x}{\alpha}|^2}. \tag{3.5.4}$$

This kernel already appears in the literature, see e.g. [16] for an application in approximation theory. The following proposition summarizes the properties that are relevant for its use as a DPP kernel.

Proposition 3.5.3. *Let C be given by* (3.5.4), *then its Fourier transform is, for all $x \in \mathbb{R}^d$,*

$$\mathcal{F}(C)(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} \alpha^d (m\pi)^{\frac{d}{2}} e^{-m(\pi\alpha|x|)^2} \sum_{k=0}^{m-1} \frac{(\pi\sqrt{m}|\alpha x|)^{2k}}{k!} \tag{3.5.5}$$

and $DPP(C)$ exists if and only if $\alpha \le \alpha_{\max}$ where

$$\alpha_{\max}^d = \frac{\binom{m-1+\frac{d}{2}}{m-1}}{\rho (m\pi)^{\frac{d}{2}}}.$$

In this case, $DPP(C)$ is stationary and isotropic with intensity ρ. Moreover, for any $\rho > 0$ and $\alpha > 0$, we have the convergence

$$\lim_{m \to +\infty} C(x) = \rho \Gamma\left(\frac{d}{2} + 1\right) \frac{J_{\frac{d}{2}}\left(2|\frac{x}{\alpha}|\right)}{|\frac{x}{\alpha}|^{\frac{d}{2}}} \tag{3.5.6}$$

uniformly on all compact sets. In particular, for $\alpha = \alpha_{\max}$,

$$\lim_{m \to +\infty} C(x) = C_B(x) \tag{3.5.7}$$

uniformly on all compact sets and where C_B is defined in (3.3.1).

This family of kernels contains the Gaussian kernel, being the particular case $m = 1$, and includes as limiting cases the Poisson kernel (3.2.1) (when $\alpha \to 0$) and C_B, in view of (3.5.7). Some illustrations of this model are provided in the supplementary material, including graphical representations of the pcf and some realizations.

3.5.3 Families of compactly supported kernels

As suggested by Proposition 3.4.1, we can consider the following family of compactly supported kernels, parameterized by the range $R > 0$,

$$C_1(x) = \frac{1}{\|h\|^2}[h * h]\left(\frac{2x}{R}\right) C_B\,(x)\,, \quad \forall x \in \mathbb{R}^d, \tag{3.5.8}$$

where h is given by (3.4.1). The Poisson kernel (3.2.1) and C_B are two limiting cases, when respectively $R \to 0$ and $R \to +\infty$. However this family of kernels has several drawbacks: No closed form expression is available for C_1, nor for $\mathcal{F}(C_1)$. Moreover, when the range R is fixed, $DPP(C_1)$ is not the most repulsive DPP, see Proposition 3.4.3 and the graphical representations in the supplementary material. This is the reason why we turn to another family of compactly supported kernels.

Following Proposition 3.4.3, we introduce a new family of compactly supported kernels with range R, given as a convolution product of functions as in (3.4.6). Specifically, let $R > 0$, $\rho > 0$ and $\alpha > 0$ such that $R/(2\alpha)$ is not a zero of the Bessel function $J_{\frac{d-2}{2}}$ and consider the kernel $C_2 = u * u$ with

$$u(x) = \sqrt{\rho}\,\beta(R,\alpha)\left(1 - \frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}-1}J_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\frac{J_{\frac{d}{2}-1}(|\frac{x}{\alpha}|)}{|x|^{\frac{d}{2}-1}}\right)\mathbf{1}_{\{|x|\le\frac{R}{2}\}}, \tag{3.5.9}$$

where

$$\beta(R,\alpha) = \left[\frac{R^{d-1}\pi^{d/2}}{2^{d-1}\Gamma\left(\frac{d}{2}\right)}\left(\frac{R}{d} - 4\alpha\frac{J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J_{\frac{d}{2}-1}\left(\frac{R}{2\alpha}\right)} + \frac{R}{2}\left(1 - \frac{J_{\frac{d}{2}-2}(\frac{R}{2\alpha})J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J^2_{\frac{d}{2}-1}\left(\frac{R}{2\alpha}\right)}\right)\right)\right]^{-\frac{1}{2}}.$$

Proposition 3.5.4. *Let $C_2 = u * u$ where u is given by* (3.5.9)*, then its Fourier transform is $\mathcal{F}(u)^2$ where for all $x \in \mathbb{R}^d$*

$$\begin{aligned}\mathcal{F}(u)(x) = \sqrt{\rho}\,\beta(R,\alpha)\left(\frac{R}{2|x|}\right)^{\frac{d}{2}-1}&\left(\frac{R}{2|x|}J_{\frac{d}{2}}(\pi R|x|)\right.\\ &\left.+\frac{\pi}{J_{\frac{d-2}{2}}(\frac{R}{2\alpha})}\frac{R\alpha J'_{\frac{d-2}{2}}(\frac{R}{2\alpha})J_{\frac{d-2}{2}}(\pi R|x|) - 2\pi R\alpha^2 J_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|J'_{\frac{d-2}{2}}(\pi R|x|)}{1 - 4\pi^2|\alpha x|^2}\right).\end{aligned}$$

Moreover, $DPP(C_2)$ exists if and only if α is such that $|\mathcal{F}(u)| \le 1$. In this case, $DPP(C_2)$ defines a stationary and isotropic R-dependent DPP with intensity ρ.

The choice of u in (3.5.9) comes from (3.4.6) where γ has been chosen such that u is continuous at $|x| = R/2$ and where β is deduced from the relation $C_2(0) = \|u\|^2 = \rho$. Given ρ and R, the remaining free parameter in this parametric family becomes α. The restriction that $R/(2\alpha)$ must not be a zero of $J_{\frac{d-2}{2}}$ can be alleviated by setting in these cases $\beta = 0$ in (3.4.6) and choose γ so that $C_2(0) = \rho$. Then the most locally repulsive DPP (3.4.4) when $R \le M$ would be part of the parametric family. However, these kernels can be arbitrarily approximated by some kernel given by (3.5.9) for some value of α, so we do not include these particular values of α in the family above.

The condition $|\mathcal{F}(u)| \leq 1$ on α, given R and ρ, must be checked numerically. In most cases, the maximal value of $\mathcal{F}(u)$ holds at the origin and we simply have to check whether $|\mathcal{F}(u)(0)| \leq 1$. No theoretical results are available to claim the existence of an admissible α, but from our experience, there seems to exist an infinity of admissible α for any R and ρ. Moreover, while the most locally repulsive DPP when $R \leq M$ is known and corresponds to (3.4.4), the most repulsive DPP when $R > M$ in the above parametric family seems to correspond to the maximal value of α such that $|\mathcal{F}(u)| \leq 1$, denoted $\alpha_{\max}$.

The parametric family given by C_2 is mainly of interest since it covers a large range of repulsive DPPs while the kernels are compactly supported. Moreover, the closed form expression of $\mathcal{F}(C_2)$ is available and this family contains the most locally repulsive DPP with range R, in view of Proposition 3.4.3, at least when $R \leq M$. Some illustrations are provided in the supplementary material.

3.6 Proofs

3.6.1 Proof of Proposition 3.3.2

As the kernel C_B verifies condition $\mathcal{K}(\rho)$, it defines a DPP with intensity ρ and its associated pcf g_B given by (3.2.3) vanishes at 0. By the analytic definition of Bessel functions, see [1, Relation (9.1.10)],

$$C_B(x) = \frac{\sqrt{\rho\Gamma(\frac{d}{2}+1)}}{\pi^{d/4}} \sum_{n=0}^{+\infty} \frac{(-1)^n \left(\sqrt{\pi}\Gamma(\frac{d}{2}+1)^{\frac{1}{d}}\rho^{\frac{1}{d}}\right)^{2n}}{2^{2n} n! \Gamma\left(n+1+d/2\right)} |x|^{2n}.$$

Thus C_B is twice differentiable at 0 and by (3.2.3), so is g_B. By Definition 3.1.2, any DPP having a pcf g that does not vanish at 0 or is not twice differentiable at 0 is less locally repulsive than $DPP(C_B)$. Consequently, we assume in the following of the proof that $g(0) = 0$ and g is twice differentiable at 0. The problem therefore reduces to minimize $\Delta g(0)$ under the constraint that g is the pcf of a DPP with kernel C verifying condition $\mathcal{K}(\rho)$.

According to condition $\mathcal{K}(\rho)$, the Fourier transform of the kernel C is well defined and belongs to $L^1(\mathbb{R}^d)$, as noticed below (3.2.2). Therefore, we can define the function $f = \frac{\mathcal{F}(C)}{\|\mathcal{F}(C)\|_1}$ where $\|\mathcal{F}(C)\|_1 = \int_{\mathbb{R}^d} |\mathcal{F}(C)(x)|dx$ and consider it as a density function of a random variable $X = (X_1, \cdots, X_d) \in \mathbb{R}^d$. Denote by $\widehat{f}(t) = \mathbb{E}\left(e^{it\cdot X}\right)$ the characteristic function of X. We have

$$\widehat{f}(t) = \frac{C\left(\frac{t}{2\pi}\right)}{\|\mathcal{F}(C)\|_1}, \quad \forall t \in \mathbb{R}^d. \tag{3.6.1}$$

Thus, $\widehat{f}$ is twice differentiable at 0, so by the usual properties of the characteristic function (see [55]), X has finite second order moments and

$$E(X_i^2) = -\frac{\partial^2 \widehat{f}}{\partial x_i^2}(0) + \left(\frac{\partial \widehat{f}}{\partial x_i}(0)\right)^2, \quad i = 1 \ldots d. \tag{3.6.2}$$

On the other hand, as already noticed in Section 3.1, $\nabla g(0) = 0$ and so $\frac{\partial C}{\partial x_i}(0) = 0$ for $i = 1, \ldots, d$. By differentiating both sides of (3.6.1),

$$\frac{\partial \widehat{f}}{\partial x_i}(0) = \frac{1}{2\pi \|\mathcal{F}(C)\|_1} \frac{\partial C}{\partial x_i}(0) = 0, \quad i = 1 \ldots d \tag{3.6.3}$$

and

$$\frac{\partial^2 \widehat{f}}{\partial x_i^2}(0) = \frac{1}{4\pi^2 \|\mathcal{F}(C)\|_1} \frac{\partial^2 C}{\partial x_i^2}(0), \quad i = 1 \ldots d. \tag{3.6.4}$$

Then, by (3.6.2)-(3.6.4),

$$E(|X|^2) = E\left(\sum_{i=1}^{d} X_i^2\right) = -\Delta \widehat{f}(0) = -\frac{1}{4\pi^2 \|\mathcal{F}(C)\|_1} \Delta C(0).$$

Moreover,

$$E(|X|^2) = \int_{\mathbb{R}^d} |x|^2 f(x) dx = \int_{\mathbb{R}^d} |x|^2 \frac{\mathcal{F}(C)}{\|\mathcal{F}(C)\|_1}(x) dx.$$

Hence,

$$\Delta C(0) = -4\pi^2 \int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x) dx. \tag{3.6.5}$$

By (3.2.3) and since $\nabla C(0) = 0$,

$$\begin{aligned}\Delta g(0) = \Delta\left(1 - \frac{C^2}{\rho^2}\right)(0) &= -\frac{1}{\rho^2}\left(\sum_{i=1}^{d} 2C(0)\frac{\partial^2 C}{\partial x_i^2}(0) + 2\left(\frac{\partial C}{\partial x_i}(0)\right)^2\right) \\ &= -\frac{2}{\rho}\sum_{i=1}^{d} \frac{\partial^2 C}{\partial x_i^2}(0) = -\frac{2}{\rho}\Delta C(0). \end{aligned} \tag{3.6.6}$$

Finally, we deduce from (3.6.5) and (3.6.6) that

$$\Delta g(0) = \frac{8\pi^2}{\rho} \int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x) dx.$$

Thus the two following optimization problems are equivalent.

Problem 1: Minimizing $\Delta g(0)$ under the constraint that g is the pcf of a DPP with kernel C satisfying condition $\mathcal{K}(\rho)$.

Problem 2: Minimizing $\int_{\mathbb{R}} |x|^2 \mathcal{F}(C)(x) dx$ under the constraint that C is a kernel which is twice differentiable at 0 and verifies the condition $\mathcal{K}(\rho)$.

The latter optimization problem is a special case of [45, Theorem 1.14], named bathtub principle, which gives the unique solution $\mathcal{F}(C) = \mathbf{1}_{\{|\cdot|^d \leq \rho\tau^d\}}$ in agreement with (3.3.1). This completes the proof.

3.6.2 Proof of Proposition 3.4.1

Notice that h is symmetric, real-valued, infinitely differentiable and verifies $h(x) = 0$ for $x \geq 1$, see [51, Section 3.2]. Thus, $\|h\|$ is finite and $\|h\| \neq 0$, so C_r is well-defined.

Since $h * h(0) = \|h\|^2$, we have $C_r(0) = \rho$. By product convolution properties, $h * h$ is symmetric, real-valued, infinitely differentiable and compactly supported with range 2. Thus, by (3.4.2), C_r is symmetric, real-valued, infinitely differentiable and compactly supported with range r. Then, C_r belongs to $L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. In particular, $\mathcal{F}(C_r)$ is well-defined pointwise. By well-known properties of the Fourier transform, for all $x \in \mathbb{R}^d$,

$$\mathcal{F}(C_r)(x) = \frac{r^d}{2^d\|h\|^2} \left[\mathcal{F}(h)^2 \left(\frac{r}{2}\cdot\right) * \mathcal{F}(C)(\cdot)\right](x). \tag{3.6.7}$$

Since h is symmetric, $\mathcal{F}(h)$ is real valued, so $\mathcal{F}(h)^2 \geq 0$. Thus, as $\mathcal{F}(C) \geq 0$ by condition $\mathcal{K}(\rho)$, we have $\mathcal{F}(C_r) \geq 0$. Further, since $0 \leq \mathcal{F}(C) \leq 1$,

$$\frac{r^d}{2^d\|h\|^2} \int_{\mathbb{R}^d} \mathcal{F}(h)^2 \left(\frac{rt}{2}\right) \mathcal{F}(C)(x-t)dt \leq \frac{r^d}{2^d\|h\|^2} \int_{\mathbb{R}^d} \mathcal{F}(h)^2 \left(\frac{rt}{2}\right) dt. \tag{3.6.8}$$

By the substitution $u = rt/2$ and Parseval's equality, the right-hand side of (3.6.8) equals 1. Finally, (3.6.7) and (3.6.8) give $\mathcal{F}(C_r) \leq 1$, i.e. $0 \leq \mathcal{F}(C_r) \leq 1$.

It remains to show the convergence result (3.4.3), which reduces to prove that $\frac{1}{\|h\|^2}[h*h]\left(\frac{2}{r}\cdot\right)$ tends to 1 uniformly on all compact set when $r \to \infty$. This follows from $h * h(0) = \|h\|^2$ and the uniform continuity of $h * h$ on every compact set.

3.6.3 Proof of Proposition 3.4.2

The proof is based on a theorem from Ehm et al. [14] recalled below with only slight changes in the presentation.

Definition 3.6.1. *Let H denote the normalized Haar measure on the group $SO(d)$ of rotations in $\mathbb{R}^d$ and let C be a kernel verifying condition $\mathcal{K}_c(\rho, R)$. The radialization of the kernel C is the kernel* $\mathrm{rad}(C)$ *defined by*

$$\mathrm{rad}(C)(x) = \int_{SO(d)} C(j(x))H(dj).$$

Note that for any isotropic kernel C, $C = \mathrm{rad}(C)$. We say that $C_1 = C_2$ up to a radialization if C_1 and C_2 are kernels verifying condition $\mathcal{K}_c(\rho, R)$ and $\mathrm{rad}(C_1) = \mathrm{rad}(C_2)$.

Define $\gamma_d > 0$ by $\gamma_d^2 = \frac{4j_{(d-2)/2}^{d-2}}{\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)J_{\frac{d}{2}}^2(j_{(d-2)/2})}$ and set $c_d = \frac{4j_{\frac{d-2}{2}}^2}{4^d\pi^{\frac{d}{2}}\Gamma\left(\frac{d}{2}\right)}$ where $j_{(d-2)/2}$ is introduced before Proposition 3.4.2.

Theorem 3.6.2 ([14]). *Let Ψ be a twice differentiable characteristic function of a probability density f on $\mathbb{R}^d$ and suppose that $\Psi(x) = 0$ for $|x| \geq 1$. Then*

$$-\Delta\Psi(0) = \int |x|^2 f(x)dx \geq 4j_{(d-2)/2}^2$$

with equality if and only if, up to a radialization, $\Psi = \omega_d * \omega_d$*, where* $\omega_d(x) = \gamma_d \frac{\Gamma\left(\frac{d}{2}\right)}{j_{(d-2)/2}^{(d-2)/2}} \frac{J_{\frac{d-2}{2}}(2 j_{\frac{d-2}{2}} |x|)}{|x|^{\frac{d-2}{2}}}$ *for* $|x| \leq \frac{1}{2}$ *and* $\omega_d(x) = 0$ *for* $|x| \geq \frac{1}{2}$*. The corresponding minimum variance density is*

$$f(x) = c_d \Gamma\left(\frac{d}{2}\right)^2 \left(\frac{2^{\frac{d-2}{2}} J_{\frac{d-2}{2}}(\frac{|x|}{2})}{|\frac{x}{2}|^{\frac{d-2}{2}} \left(j_{(d-2)/2}^2 - \left(\frac{|x|}{2}\right)^2 \right)} \right)^2 .$$

According to Definition 3.1.2 and by the same arguments as in the proof of Proposition 3.3.2 and (3.6.6), we seek a kernel C which is twice differentiable at 0 such that $\Delta C(0)$ is maximal among all kernels verifying condition $\mathcal{K}_c(\rho, R)$.

In a first step, we exhibit a candidate for the solution to this optimization problem and in a second step we check that it verifies all required conditions.

Step 1. We say that a function C verifies $\widetilde{\mathcal{K}_c}(\rho, R)$ if it verifies $\mathcal{K}_c(\rho, R)$ without necessarily verifying $\mathcal{F}(C) \leq 1$. Notice that a function C verifies $\widetilde{\mathcal{K}_c}(\rho, R)$ if and only if the function

$$\Psi(x) = \frac{C(Rx)}{\rho}, \ x \in \mathbb{R}^d, \tag{3.6.9}$$

verifies $\widetilde{\mathcal{K}_c}(1, 1)$. Therefore, we have a one-to-one correspondence between $\widetilde{\mathcal{K}_c}(\rho, R)$ and $\widetilde{\mathcal{K}_c}(1, 1)$.

On the other hand, if a function Ψ verifies condition $\widetilde{\mathcal{K}_c}(1, 1)$, it is by Bochner's Theorem the characteristic function of a random variable X. Moreover, the function Ψ is continuous and compactly supported, so it is in $L^1(\mathbb{R}^d)$ and the random variable X has a density f, see [55]. Thus, by Theorem 3.6.2, any function Ψ twice differentiable at 0 and verifying condition $\widetilde{\mathcal{K}_c}(1, 1)$ satisfies

$$\Delta\Psi(0) \leq -4 j_{(d-2)/2}^2. \tag{3.6.10}$$

By differentiating both sides of (3.6.9), we have

$$\Delta\Psi(0) = \frac{R^2}{\rho} \Delta C(0). \tag{3.6.11}$$

Thus, by (3.6.10)-(3.6.11), for any kernel C which is twice differentiable at 0 and verifies $\widetilde{\mathcal{K}_c}(\rho, R)$,

$$\Delta C(0) = \frac{\rho \Delta\Psi(0)}{R^2} \leq -\frac{4\rho j_{(d-2)/2}^2}{R^2}. \tag{3.6.12}$$

By Theorem 3.6.2, the equality in (3.6.12) holds if and only if $\Psi = \omega_d * \omega_d$ and we name C_R the corresponding kernel C given by (3.6.9).

Step 2. The kernel C_R is the candidate to our optimization problem, however it remains to prove that it verifies condition $\mathcal{K}_c(\rho, R)$. We have seen in *Step 1* that C_R verifies $\widetilde{\mathcal{K}_c}(\rho, R)$ and is twice differentiable at 0. We must show that $\mathcal{F}(C_R) \leq 1$. By Theorem 3.6.2, the function $\Psi = \omega_d * \omega_d$ is the characteristic function of a probability density f. Thus, for all $x \in \mathbb{R}^d$,

$$\mathcal{F}(\Psi)(x) = (2\pi)^d f(2\pi x) = (2\pi)^d c_d \Gamma\left(\frac{d}{2}\right)^2 \left(\frac{2^{\frac{d-2}{2}} J_{\frac{d-2}{2}}(|\pi x|)}{|\pi x|^{\frac{d-2}{2}} \left(j_{(d-2)/2}^2 - (|\pi x|)^2 \right)} \right)^2 . \tag{3.6.13}$$

By (3.6.9) and the Fourier transform dilatation we thereby obtain (3.4.5).

Moreover, the Bessel functions are non-negative up to their first non-negative zero so $\omega_d \geq 0$, which implies that $\Psi \geq 0$. Hence by (3.6.13),

$$\mathcal{F}(\Psi)(x) = \left| \int_{\mathbb{R}^d} \Psi(t) e^{2i\pi x \cdot t} dt \right| \leq \int_{\mathbb{R}^d} \Psi(t) dt = \mathcal{F}(\Psi)(0) = \frac{2^d \pi^d c_d}{j^4_{\frac{d-2}{2}}}. \tag{3.6.14}$$

Thus, by (3.6.9) and the Fourier transform dilatation,

$$\mathcal{F}(C_R)(x) \leq \mathcal{F}(C_R)(0) = \frac{2^d R^d \rho \pi^d c_d}{j^4_{\frac{d-2}{2}}} = \frac{R^d}{M^d}. \tag{3.6.15}$$

Since by hypothesis $R \leq M$, we have $\mathcal{F}(C_R) \leq 1$.

3.6.4 Proof of Proposition 3.4.3

According to Definition 3.1.2 and by the same arguments as in the proof of Proposition 3.3.2 and (3.6.6), we seek a kernel C which is twice differentiable at 0 such that $\Delta C(0)$ is maximal among all kernels verifying condition $\mathcal{K}_c(\rho, R)$. By (3.6.5), this is equivalent to solve the following problem A.

Problem A: Minimize $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x) dx$ under the constraints that C is twice differentiable at 0 and verifies $\mathcal{K}_c(\rho, R)$.

The proof of Proposition 3.4.3 is based on the following three lemmas. In the first lemma, the gradient ∇u has to be considered in the sense of distribution when $u \in L^2(\mathbb{R}^d)$ is not differentiable.

Lemma 3.6.3. *A kernel C_R is solution to Problem A if and only if there exists a function u such that, up to a radialization, $C_R = u * u$ where u minimizes $\int_{\mathbb{R}^d} |\nabla u(x)|^2 dx$ among all functions u verifying $\mathcal{M}(\rho, R)$ and $\mathcal{F}(u)^2 \leq 1$.*

The existence statement in Proposition 3.4.3 is given by the following lemma.

Lemma 3.6.4. *There exists a solution to Problem A.*

By Lemma 3.6.3, $C_R = u * u$ where u is the solution of the given optimization problem. Then, under the additional constraint $\sup_{x \in \mathbb{R}^d} \mathcal{F}(C)(x) = \mathcal{F}(C)(0)$, we have $\sup_{x \in \mathbb{R}^d} (\mathcal{F}(u)(x))^2 = (\mathcal{F}(u)(0))^2$. Since $\mathcal{F}(u)^2(0) = \left(\int_{\mathbb{R}^d} u(t) dt\right)^2$, the constraint $\mathcal{F}(u)^2 \leq 1$ in Lemma 3.6.3 becomes $\left(\int_{\mathbb{R}^d} u(t) dt\right)^2 \leq 1$. Notice that $-u$ is also a solution of the optimization problem. Thus, we can assume without loss of generality that $\int_{\mathbb{R}^d} u(t) dt \geq 0$, so that the constraint $\left(\int_{\mathbb{R}^d} u(t) dt\right)^2 \leq 1$ becomes $\int_{\mathbb{R}^d} u(t) dt \leq 1$. In this situation, the optimization problem addressed in Lemma 3.6.3 can be solved by variational calculus. However, an explicit form of the solution is available only if we assume that $u \in \mathcal{C}^2(B\left(0, \frac{R}{2}\right))$, meaning that u is twice continuously differentiable on its support. It is given by the following lemma, which completes the proof of Proposition 3.4.3.

Lemma 3.6.5. *If a function u minimizes $\int_{\mathbb{R}^d} |\nabla u(x)|^2 dx$ among all functions u verifying $\mathcal{M}(\rho, R)$, $u \in \mathcal{C}^2(B\left(0, \frac{R}{2}\right))$ and $\int_{\mathbb{R}^d} u(x)dx \leq 1$, then u is of the form*

$$u(x) = \left(\beta + \gamma \frac{J_{\frac{d-2}{2}}\left(|x|/\alpha\right)}{|x|^{\frac{d-2}{2}}}\right) \mathbf{1}_{\{|x|<\frac{R}{2}\}},$$

where $\alpha > 0, \beta \geq 0$ and γ are three constants linked by the conditions $\mathcal{M}(\rho, R)$ and $\int_{\mathbb{R}^d} u(x)dx \leq 1$.

Proof of Lemma 3.6.3

Let C be a kernel which is twice differentiable at 0 and verifies the condition $\mathcal{K}_c(\rho, R)$. This implies that C is twice differentiable everywhere. Moreover, the quantity $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ is invariant under radialization of the kernel C, see [14, Relation (44)]. Thus, we can consider C as a radial function. Then, by [14, Theorem 3.8], there exists a countable set A and a sequence of real valued functions $\{u_k\}_{k\in A}$ in $L^2(\mathbb{R}^d)$ such that

$$C(x) = \sum_{k\in A} u_k * u_k(x). \tag{3.6.16}$$

Further, the convergence of the series is uniform and for each $k \in A$, the support of u_k lies in $B\left(0, \frac{R}{2}\right)$. Thus,

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx = \int_{\mathbb{R}^d} |x|^2 \sum_{k\in A} |\mathcal{F}(u_k)(x)|^2 \, dx = \sum_{k\in A} \sum_{j=1}^{d} \int_{\mathbb{R}^d} |x_j \mathcal{F}(u_k)(x)|^2 \, dx \tag{3.6.17}$$

where x_j denotes the j-th coordinate of the vector x. In addition, we note that $u_k \in L^2(\mathbb{R}^d)$ so $|\cdot|\mathcal{F}(u_k)(\cdot) \in L^2(\mathbb{R}^d)$ by (3.6.17). Then, by [45, Theorem 7.9], $\nabla u_k \in L^2(\mathbb{R}^d)$ where ∇u_k has to be viewed in the distributional sense and

$$\mathcal{F}\left(\partial_j u_k\right)(x) = 2i\pi x_j \mathcal{F}(u_k)(x). \tag{3.6.18}$$

Thus, from (3.6.17)-(3.6.18) and the Parseval equality,

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx = \sum_{k\in A} \int_{\mathbb{R}^d} \frac{|\nabla u_k(x)|^2}{4\pi^2} dx.$$

As every term in the sum above is positive and since this equality holds for every kernel C, the minimum of $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ is reached if and only if this sum reduces to one term where $u_k = u$. Then we have $C = u * u$ and

$$\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx = \int_{\mathbb{R}^d} \frac{|\nabla u(x)|^2}{4\pi^2} dx. \tag{3.6.19}$$

Therefore, minimizing $\int_{\mathbb{R}^d} |x|^2 \mathcal{F}(C)(x)dx$ is equivalent to minimize $\int_{\mathbb{R}^d} |\nabla u(x)|^2 \, dx$. Hence it remains to see what the constraints on the kernel C means for the function u. Since $C = u * u$, where u is one of the function in the decomposition (3.6.16), u is a so-called real valued Boas-Kac root of C, see [14]. Thus, since C is radial, we have by [14, Theorem 3.1] that u is radial and verifies $u(x) = 0$ for $|x| \geq \frac{R}{2}$. Since C verifies $\mathcal{K}_c(\rho, R)$, we have $C(0) = \rho$ and $0 \leq \mathcal{F}(C) \leq 1$. These constraints are equivalent on u to $\int_{\mathbb{R}^d} u(x)^2 dx = \rho$ and $\mathcal{F}(u)^2 \leq 1$, respectively. Therefore, u verifies condition $\mathcal{M}(\rho, R)$ and $\mathcal{F}(u)^2 \leq 1$.

Proof of Lemma 3.6.4

According to Lemma 3.6.3, C_R is a is solution to Problem A if and only if $C_R = u * u$ where u minimizes $\int_{\mathbb{R}^d} |\nabla u(x)|^2 dx$ among all functions u verifying $\mathcal{M}(\rho, R)$ and $\mathcal{F}(u)^2 \leq 1$. We prove the existence of such a minimum u.

Let Ω denote the open Euclidean ball $B\left(0, \frac{R}{2}\right)$. Consider the Sobolev space

$$H^1(\Omega) = \left\{f : \Omega \to \mathbb{R},\ f \in L^2(\Omega),\ \nabla f \in L^2(\Omega)\right\},$$

with the norm $\|f\|_{H^1(\Omega)} = \left(\|f\|^2 + \|\nabla f\|^2\right)^{\frac{1}{2}}$. For a review on Sobolev spaces, see for example [15] or [45]. For any $f \in H^1(\Omega)$, we consider its extension to $\mathbb{R}^d$ by setting $f(x) = 0$ if $x \notin \Omega$, so that $f \in L^2(\mathbb{R}^d)$. Let us further denote $\mathcal{E}$ the set of functions $f \in H^1(\Omega)$ verifying $\mathcal{M}(\rho, R)$ and $\mathcal{F}(f)^2 \leq 1$.

If the minimum u above exists but $u \notin H^1(\Omega)$, then $\int_\Omega |\nabla u(x)|^2 dx = \infty$, which means that $\mathcal{E}$ is empty, otherwise u would not be the solution of our optimization problem. But $\mathcal{E}$ is not empty, see for instance the functions in Section 3.5.3, so if u exists, $u \in H^1(\Omega)$. Let $(w_k)_{k\in\mathbb{N}}$ be a minimizing sequence in $\mathcal{E}$, i.e.

$$\int_\Omega |\nabla w_k(x)|^2\, dx \underset{k\to+\infty}{\longrightarrow} \inf_{v\in\mathcal{E}} \int_\Omega |\nabla v(x)|^2\, dx, \tag{3.6.20}$$

where for all k, $w_k \in \mathcal{E}$. By (3.6.20) and since for all k, $\int_\Omega |w_k(x)|^2 dx = \rho$, the sequence $\{w_k\}$ is bounded in $H^1(\Omega)$. Then, by the Rellich-Kondrachov compactness theorem (see [15]), it follows that, up to a subsequence, $\{w_k\}$ converges in $L^2(\mathbb{R}^d)$ to a certain function $w \in L^2(\mathbb{R}^d)$ verifying

$$\int_\Omega |\nabla w(x)|^2\, dx = \inf_{v\in\mathcal{E}} \int_\Omega |\nabla v(x)|^2\, dx. \tag{3.6.21}$$

We now prove that $w \in \mathcal{E}$, so that $u = w$ is the solution of our optimization problem. First $w \in H^1(\Omega)$ as justified earlier and so $w \in L^2(\mathbb{R}^d)$. Second, as rotations are isometric functions and since any w_k is radial by hypothesis, we have for any $j \in SO(d)$

$$\begin{aligned}\left\{\int_{\mathbb{R}^d} |w(x) - w_k(x)|^2\, dx \to 0\right\} &\Leftrightarrow \left\{\int_{\mathbb{R}^d} |w(j(x)) - w_k(j(x))|^2\, dx \to 0\right\} \\ &\Leftrightarrow \left\{\int_{\mathbb{R}^d} |w(j(x)) - w_k(x)|^2\, dx \to 0\right\}.\end{aligned}$$

Hence, by uniqueness of the limit, the function w is radial and in particular, its Fourier transform is real. Further, since w is the limit in $L^2(\mathbb{R}^d)$ of w_k, w verifies the following properties:

- w is compactly supported in $B\left(0, \frac{R}{2}\right)$, because $w_k \in \mathcal{E}$ for all k.
- $w \in L^2(\mathbb{R}^d)$ by Rellich-Kondrachov theorem.
- $\int_{\mathbb{R}^d} |w(x)|^2 dx = \int_{\mathbb{R}^d} |w_k(x)|^2 dx = \rho$ since a sphere in $L^2(\mathbb{R}^d)$ is closed.

Therefore, w verifies $\mathcal{M}(\rho, R)$. Third, for every k, w_k being compactly supported and in $L^2(\mathbb{R}^d)$, $w_k \in L^1(\mathbb{R}^d)$ so we can consider $\mathcal{F}(w_k)(x)$ for every $x \in \mathbb{R}^d$ and by the Cauchy-Schwartz inequality

$$|\mathcal{F}(w)(x) - \mathcal{F}(w_k)(x)| \leq a\sqrt{\int_{\mathbb{R}^d} |w(t) - w_k(t)|^2\, dt}, \quad \forall x \in \mathbb{R}^d,$$

where a is a positive constant. Thereby the convergence of w_k to w in $L^2(\mathbb{R}^d)$ implies the pointwise convergence of $\mathcal{F}(w_k)$ to $\mathcal{F}(w)$. Finally, from the relation

$$\mathcal{F}(w_k)(x) \leq 1, \quad \forall x \in \mathbb{R}^d, \quad \forall k \in \mathbb{N},$$

we deduce $\mathcal{F}(w) \leq 1$.

Proof of Lemma 3.6.5

We denote as before $\Omega = B\left(0, \frac{R}{2}\right)$. The optimization problem in Lemma 3.6.5 is a variational problem with isoperimetric constraints. By [20, Chapter 2, Theorem 2], every solution must solve

$$\begin{aligned} \Delta u + \lambda_1 u - \frac{\lambda_2}{2} &= 0 \text{ on } \Omega, \\ u &= 0 \text{ on } \partial\Omega. \end{aligned} \tag{3.6.22}$$

In equation (3.6.22), λ_1 and λ_2 are the Lagrange multipliers associated to the constraints $\int u^2 = \rho$ and $\int u \leq 1$, respectively. By the Karush-Kuhn-Tucker theorem, see [31, Section VII], $\lambda_2 \geq 0$. Moreover, a solution to the partial differential equation with boundary condition (3.6.22) is obtained by linear combination of a homogeneous solution and a particular solution. By [15, Section 6.5, Theorem 2], the Laplacian operator $-\Delta$ has only positive eigenvalues. Hence the associated homogeneous equation $\Delta u + \lambda_1 u = 0$ can have a solution only if $\lambda_1 > 0$.

In addition, the function u is radial by hypothesis, so there exists a function $\tilde{u}$ on $\mathbb{R}$ such that $u(x) = \tilde{u}(|x|)$ for all $x \in \mathbb{R}^d$. The partial differential equation (3.6.22) then becomes

$$\begin{aligned} \tilde{u}''(t) + \frac{d-1}{t}\tilde{u}'(t) + \lambda_1 \tilde{u}(t) - \frac{\lambda_2}{2} &= 0, \quad \forall t \in \left]0, \frac{R}{2}\right[, \\ \tilde{u}\left(\frac{R}{2}\right) &= 0. \end{aligned}$$

As λ_1 is positive, we obtain from [67, Section 4.31, Relations (3) and (4)] that a solution to this equation is of the form

$$\tilde{u}(t) = \left(\frac{\lambda_2}{2\lambda_1} + c_1 \frac{J_{(d-2)/2}(\sqrt{\lambda_1}t)}{t^{(d-2)/2}} + c_2 \frac{Y_{(d-2)/2}(\sqrt{\lambda_1}t)}{t^{(d-2)/2}}\right) \mathbf{1}_{\{0<t<\frac{R}{2}\}}, \tag{3.6.23}$$

where $Y_{(d-2)/2}$ denotes the Bessel function of the second kind. By hypothesis, the function u is continuous on Ω and so at 0. Since $Y_{(d-2)/2}$ has a discontinuity at 0, see for example [1], and the remaining terms in (3.6.23) are continuous, we must have $c_2 = 0$. Then, by renaming the constant c_1 by γ and letting $\alpha = 1/\sqrt{\lambda_1}$, $\beta = \lambda_2/(2\lambda_1)$, we obtain that if u is solution to the optimization problem of Lemma 3.6.5, then u writes

$$u(x) = \left(\beta + \gamma \frac{J_{(d-2)/2}(|x|/\alpha)}{|x|^{(d-2)/2}}\right) \mathbf{1}_{\{x \in \Omega\}}, \tag{3.6.24}$$

where $\alpha > 0$ and $\beta \geq 0$.

3.6.5 Proof of Proposition 3.5.1

Let C be given by (3.5.1). According to Proposition 3.2.3, $DPP(C)$ exists and has intensity ρ if C verifies the condition $\mathcal{K}(\rho)$. By [1, Equation (9.1.7)], we have $C(0) = \rho$. It is immediate that C is a symmetric continuous real-valued function. Since Bessel functions are analytic and by the asymptotic form in [1, (9.2.1)], it is clear that C belongs to $L^2(\mathbb{R}^d)$. It remains to obtain $\mathcal{F}(C)$ and verify the condition $0 \leq \mathcal{F}(C) \leq 1$.

Define

$$p_\sigma(x) = \frac{J_{\frac{\sigma+d}{2}}(|x|)}{|x|^{\frac{\sigma+d}{2}}}, \quad \forall x \in \mathbb{R}^d. \tag{3.6.25}$$

As p_σ is radial, by [22, Appendix B.5],

$$\mathcal{F}(p_\sigma)(x) = \frac{2\pi}{|x|^{\frac{d-2}{2}}} \int_0^{+\infty} r^{\frac{d}{2}} p_\sigma(r) J_{\frac{d-2}{2}}(2\pi r|x|)dr.$$

By [21, Formula 6.575], we have for $\sigma > -2$

$$\mathcal{F}(p_\sigma)\left(x\right) = \frac{2\pi}{|2\pi x|^{\frac{d-2}{2}}} \frac{(1-|2\pi x|^2)_+^{\frac{\sigma}{2}} |2\pi x|^{\frac{d-2}{2}}}{2^{\frac{\sigma}{2}}\Gamma(\frac{\sigma}{2}+1)} = 2^{\frac{d}{2}-\frac{\sigma}{2}} \pi^{\frac{d}{2}} \frac{(1-|2\pi x|^2)_+^{\frac{\sigma}{2}}}{\Gamma(\frac{\sigma+2}{2})}.$$

Since $C(x) = \rho 2^{\frac{\sigma+d}{2}} \Gamma\left(\frac{\sigma+d+2}{2}\right) p_\sigma\left(2\frac{x}{\alpha}\sqrt{\frac{\sigma+d}{2}}\right)$, we obtain (3.5.2) by dilatation of the Fourier transform.

We have obviously $\mathcal{F}(C) \geq 0$. Since $\sigma \geq 0$, $\mathcal{F}(C)$ attains its maximum at 0. Thus $\mathcal{F}(C) \leq 1$ if and only if

$$\mathcal{F}(C)(0) = \frac{\rho(2\pi)^{\frac{d}{2}} \alpha^d \Gamma(\frac{\sigma+d+2}{2})}{(\sigma+d)^{\frac{d}{2}} \Gamma(\frac{\sigma+2}{2})} \leq 1,$$

which is equivalent to $\alpha^d \leq \frac{(\sigma+d)^{\frac{d}{2}}\Gamma(\frac{\sigma+2}{2})}{\rho(2\pi)^{\frac{d}{2}}\Gamma(\frac{\sigma+d+2}{2})}$.

Finally, when $\sigma = 0$ and $\alpha = \alpha_{\max}$, $DPP(C)$ exists and a straightforward calculation gives $C = C_B$. The convergence result (3.5.3) may be found in [17] and is a direct application of [53, Relation (1.8)].

3.6.6 Proof of Proposition 3.5.3

Define, for all $m \in \mathbb{N} \cap \{0\}$,

$$f_m(x) = L_m^{d/2}\left(|x|^2\right) e^{-|x|^2}, \quad \forall x \in \mathbb{R}^d. \tag{3.6.26}$$

This function is radial, thus by [22, Appendix B.5] we have

$$\mathcal{F}(f_m)(x) = \frac{2\pi}{|x|^{\frac{d-2}{2}}} \int_0^{+\infty} r^{\frac{d}{2}} L_m^{\frac{d}{2}}(r^2) e^{-r^2} J_{\frac{d-2}{2}}(2\pi r|x|)dr.$$

According to [38], we have

$$
\begin{aligned}
\mathcal{F}(f_m)(x) &= \frac{2\pi}{|x|^{\frac{d-2}{2}}}\frac{(-1)^m}{2}\left(\frac{|2\pi x|}{2}\right)^{\frac{d-2}{2}} e^{-\frac{|2\pi x|^2}{4}} L_m^{-1-m}\left(\frac{|2\pi x|^2}{4}\right) \\
&= \pi^{\frac{d}{2}}(-1)^m e^{-|\pi x|^2}\sum_{k=0}^{m}\binom{-1}{m-k}\frac{(-1)^k|\pi x|^{2k}}{k!} \\
&= \pi^{\frac{d}{2}}(-1)^m e^{-|\pi x|^2}\sum_{k=0}^{m}(-1)^{m-k}\frac{(-1)^k|\pi x|^{2k}}{k!}.
\end{aligned}
$$

Therefore,

$$
\mathcal{F}(f_m)(x) = \pi^{\frac{d}{2}} e^{-|\pi x|^2}\sum_{k=0}^{m}\frac{|\pi x|^{2k}}{k!}.
$$

As $C(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} f_{m-1}(\frac{1}{\sqrt{m}}\frac{x}{\alpha})$, we obtain (3.5.5) by dilatation and linearity of the Fourier transform.

Clearly, $\mathcal{F}(C) \geq 0$. Thus we investigate the condition $\mathcal{F}(C) \leq 1$ for the existence of $DPP(C)$. We notice from (3.5.5) that

$$
\mathcal{F}(C)(x) = ae^{-b|x|^2}\sum_{k=0}^{m-1}\frac{b^k|x|^{2k}}{k!}, \tag{3.6.27}
$$

where a and b are positive constants. Since $\mathcal{F}(C)$ depends on the variable x only through its norm, we consider the function h define for all $r \geq 0$ by $h(r) = \mathcal{F}(C)\left((r, 0, \cdots, 0)\right)$, so that for all $x \in \mathbb{R}^d$, $\mathcal{F}(C)(x) = h(|x|)$. For every $r > 0$, h is differentiable at r and a straightforward calculation leads to

$$
h'(r) = ae^{-br^2}\left(-2br\sum_{k=0}^{m-1}\frac{b^k r^{2k}}{k!} + \sum_{k=1}^{m-1} 2k\frac{b^k r^{2k-1}}{k!}\right) = -2ae^{-br^2}\frac{b^m r^{2m-1}}{(m-1)!}.
$$

Thus, the function h is decreasing on $(0, +\infty)$. Since h is continuous on $\mathbb{R}^+$, its maximum is attained at zero, so for every $x \in \mathbb{R}^d$,

$$
\mathcal{F}(C)(x) \leq \mathcal{F}(C)(0) = \frac{\rho\,(m\pi)^{\frac{d}{2}}}{\binom{m-1+\frac{d}{2}}{m-1}}\alpha^d.
$$

Hence, $\mathcal{F}(C) \leq 1$ if and only if $\alpha^d \leq \frac{\binom{m-1+\frac{d}{2}}{m-1}}{\rho(m\pi)^{\frac{d}{2}}}$. Moreover C is radial and since $L_{m-1}^{d/2}(0) = \binom{m-1+\frac{d}{2}}{m-1}$, see [1, Relation (22.4.7)], we have $C(0) = \rho$. Therefore, C verifies the condition $\mathcal{K}(\rho)$ and by Proposition 3.2.3, $DPP(C)$ exists and is stationary with intensity $\rho > 0$.

It remains to prove the convergence results (3.5.6) and (3.5.7). An immediate application of [62, Theorem 8.1.3] gives the convergence (3.5.6), see also [2, Proposition 1]. Moreover,

$$
\lim_{m\to+\infty}\alpha_{\max} = \frac{1}{\sqrt{\pi}\Gamma\left(\frac{d}{2}+1\right)^{\frac{1}{d}}\rho^{\frac{1}{d}}}. \tag{3.6.28}
$$

Hence, by (3.5.6) and (3.6.28), we obtain the convergence (3.5.7).

3.6.7 Proof of Proposition 3.5.4

By the discussion in Section 3.4, $DPP(C)$ exists and is an R-dependent DPP with intensity ρ if C verifies $\mathcal{K}_c(\rho, R)$. Since $u \in L^2(\mathbb{R}^d)$, the kernel C is continuous by [45, Theorem 2.20]. Moreover, $u(x) = 0$ for $|x| > \frac{R}{2}$, so by product convolution properties, $C(x) = 0$ for $|x| > R$. Hence C belongs to $L^2(\mathbb{R}^d)$. Since u is radial, so is C. It remains to verify that $0 \leq \mathcal{F}(C) \leq 1$ and $C(0) = \rho$.

By product convolution properties, we have $C(0) = \int_{\mathbb{R}^d} u(x)^2 dx$. From the definition of u in (3.5.9), we have

$$
\begin{aligned}
&\frac{\int_{\mathbb{R}^d} u^2(x)dx}{\rho\beta(R,\alpha)^2} \\
&= \int_{\mathbb{R}^d} \left(1 - 2\left(\frac{R}{2}\right)^{\frac{d}{2}-1} \frac{J_{\frac{d-2}{2}}(\left|\frac{x}{\alpha}\right|)}{J_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|^{\frac{d-2}{2}}} + \left(\frac{R}{2}\right)^{d-2} \frac{J^2_{\frac{d-2}{2}}(\left|\frac{x}{\alpha}\right|)}{J^2_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|^{d-2}}\right) \mathbf{1}_{\{|x|\leq \frac{R}{2}\}} dx \\
&= \frac{2\pi^{\frac{d}{2}}}{\Gamma(d/2)} \int_0^{\frac{R}{2}} \left(r^{d-1} - 2\left(\frac{R}{2}\right)^{\frac{d-2}{2}} \frac{J_{\frac{d-2}{2}}(\frac{r}{\alpha})}{J_{\frac{d-2}{2}}(\frac{R}{2\alpha})} r^{\frac{d}{2}} + \left(\frac{R}{2}\right)^{d-2} \frac{J^2_{\frac{d-2}{2}}(\frac{r}{\alpha})}{J^2_{\frac{d-2}{2}}(\frac{R}{2\alpha})} r\right) dr.
\end{aligned}
$$

By properties of Bessel functions, see [1], we notice that for all $b \in \mathbb{R}$, a primitive of $xJ^2_{\frac{d-2}{2}}(bx)$ is given by $\frac{x^2}{2}\left(J^2_{\frac{d-2}{2}}(xb) - J_{\frac{d}{2}-2}(xb)J_{\frac{d}{2}}(xb)\right)$. It follows from [22, Appendix B.3] that

$$
\begin{aligned}
\frac{\int_{\mathbb{R}^d} u^2(x)dx}{\rho\beta(R,\alpha)^2} = \frac{\pi^{d/2}R^d}{d\,\Gamma(\frac{d}{2})2^{d-1}} - 4\left(\frac{R}{2}\right)^{d-1} \frac{\pi^{\frac{d}{2}}\alpha}{\Gamma(\frac{d}{2})} \frac{J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J_{\frac{d}{2}-1}(\frac{R}{2\alpha})} \\
+ \left(\frac{R}{2}\right)^d \frac{\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})} \left(1 - \frac{J_{\frac{d}{2}-2}(\frac{R}{2\alpha})J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J^2_{\frac{d}{2}-1}(\frac{R}{2\alpha})}\right).
\end{aligned}
$$

Thus, by the definition of $\beta(R, \alpha)$, we obtain that $\int_{\mathbb{R}^d} u(x)^2 dx = \rho$.

We now calculate $\mathcal{F}(C)$. We have $\mathcal{F}(C) = \mathcal{F}(u)^2$. Since u is radial, $\mathcal{F}(u)$ is real valued and so $\mathcal{F}(C) \geq 0$. In addition, we have by [22, Appendix B.5] and (3.5.9),

$$
\begin{aligned}
\mathcal{F}(u)(x) = \sqrt{\rho}\beta(R,\alpha)\frac{2\pi}{|x|^{\frac{d-2}{2}}} \Bigg(\int_0^{\frac{R}{2}} r^{\frac{d}{2}} J_{\frac{d-2}{2}}(2\pi r|x|)dr \\
- \frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}-1}J_{\frac{d}{2}-1}(\frac{R}{2\alpha})} \int_0^{\frac{R}{2}} rJ_{\frac{d-2}{2}}\left(\frac{r}{\alpha}\right) J_{\frac{d-2}{2}}(2\pi r|x|)dr\Bigg).
\end{aligned}
$$

Since $\alpha > 0$, we have by [22, Appendix B.3] and [21, Formula 6.521],

$$
\begin{aligned}
\mathcal{F}(u)(x) = \sqrt{\rho}\beta(R,\alpha)\frac{2\pi}{|x|^{\frac{d-2}{2}}} \Bigg(\frac{R^{\frac{d}{2}}}{\pi 2^{\frac{d}{2}+1}} \frac{J_{\frac{d}{2}}(\pi R|x|)}{|x|} \\
+ \frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}}J_{\frac{d}{2}-1}(\frac{R}{2\alpha})} \frac{R\alpha J'_{\frac{d-2}{2}}(\frac{R}{2\alpha})J_{\frac{d-2}{2}}(\pi R|x|) - 2\pi\alpha^2 RJ_{\frac{d-2}{2}}(\frac{R}{2\alpha})|x|J'_{\frac{d-2}{2}}(\pi R|x|)}{1 - 4\pi^2|\alpha x|^2}\Bigg) \qquad (3.6.29)
\end{aligned}
$$

from which we deduce the Fourier transform of u in Proposition 3.5.4. Therefore, if α is such that $\mathcal{F}(u)^2 \leq 1$, then $\mathcal{F}(C) \leq 1$ and so C verifies $\mathcal{K}_c(\rho, R)$.

3.7 Supplementary Material for "Quantifying repulsiveness of determinantal point processes"

3.7.1 Non uniqueness of the most globally repulsive DPP

As seen in [7, Proposition 3.1], there exists an infinity of DPPs that are the most globally repulsive stationary DPPs. We give three examples in dimension $d = 2$. For $R > 0$, denote $B(0, R)$ the euclidean ball with radius R and for $0 < r < R$, denote $A(r, R)$ the annulus $B(0, R) \setminus B(0, r)$. We consider the DPPs with kernels

- $C_I(x) = C_B(x) = \mathcal{F}\left(\mathbf{1}_{B\left(0, \frac{1}{\sqrt{\pi}}\right)}\right)(x) = \frac{J_1(2\sqrt{\pi}x)}{\sqrt{\pi}x}$,
- $C_{II}(x) = \mathcal{F}\left(\mathbf{1}_{A(\frac{1}{2}, R) \cup B(0, \frac{1}{3})}\right)(x) = \frac{6RJ_1(2\pi Rx) - 3J_1(\pi x) + 2J_1\left(\frac{2\pi x}{3}\right)}{6x}$,
- $C_{III}(x) = \mathcal{F}\left(\mathbf{1}_{A\left(\frac{1}{\sqrt{\pi}}, \sqrt{\frac{2}{\pi}}\right)}\right)(x) = \frac{\sqrt{2}J_1(2\sqrt{2\pi}x) - J_1(2\sqrt{\pi}x)}{\sqrt{\pi}x}$,

where $R = \sqrt{\frac{1}{\pi} + \frac{1}{4} - \frac{1}{9}}$ and the formulas are deduced from [22, Appendix B.5]. Note that each kernel above is the Fourier transform of the indicator of a set with unit volume, so that each associated DPP has unit intensity and belongs to the most repulsive DPPs of [7, Proposition 3.1] when $\rho = 1$ and $d = 2$.

Figure 3.3 show the pcf of these three DPPs, as radial functions since all kernels above are isotropic. This representation confirms the natural choice $DPP(C_B)$ among all globally repulsive stationary DPPs, if we favor repulsiveness at small distances.

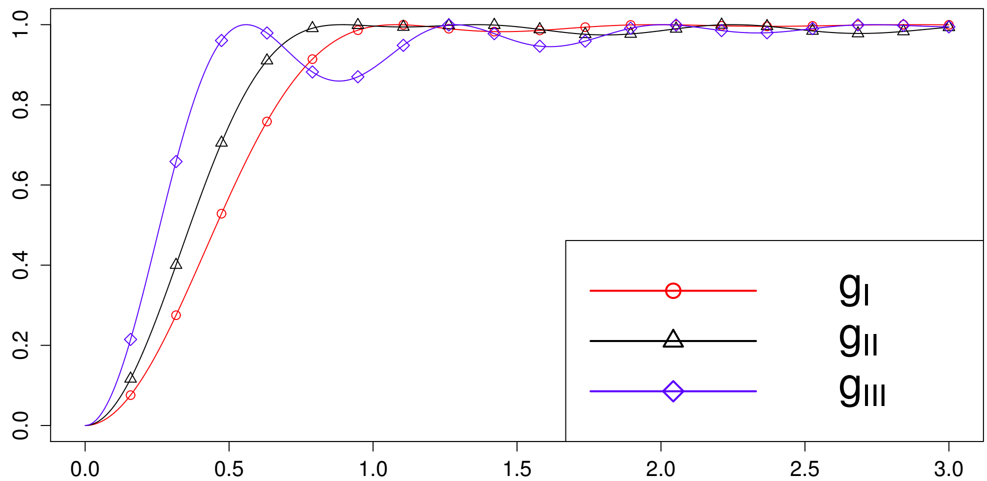


Figure 3.3 – Comparison between the pcfs of $DPP(C_I)$, $DPP(C_{II})$ and $DPP(C_{III})$ denoted g_I, g_{II} and g_{III} respectively

3.7.2 Bessel-type family of DPPs

In dimension d, for all $\sigma \geq 0,\ \alpha > 0,\ \rho > 0$, the Bessel-type kernel C is defined by

$$C(x) = \rho\, 2^{\frac{\sigma+d}{2}} \Gamma\left(\frac{\sigma + d + 2}{2}\right) \frac{J_{\frac{\sigma+d}{2}}\left(2|\frac{x}{\alpha}|\sqrt{\frac{\sigma+d}{2}}\right)}{\left(2|\frac{x}{\alpha}|\sqrt{\frac{\sigma+d}{2}}\right)^{\frac{\sigma+d}{2}}}, \quad \forall x \in \mathbb{R}^d, \tag{3.7.1}$$

where for a vector x, $|x|$ denotes its euclidean norm and $J_{\frac{\sigma+d}{2}}$ is the Bessel function of the first kind of order $\frac{\sigma+d}{2}$. According to [7, Proposition 5.1], $DPP(C)$ exists if and only if $\alpha \leq \alpha_{\max}$ where

$$\alpha^d_{\max} = \frac{(\sigma + d)^{\frac{d}{2}} \Gamma(\frac{\sigma+2}{2})}{\rho (2\pi)^{\frac{d}{2}} \Gamma(\frac{\sigma+d+2}{2})}.$$

Examples of pcfs for this family are represented as radial functions (the kernels being isotropic) in Figures 3.4-3.5. Figure 3.6 shows, when $d = 2$, six realizations of DPPs with Bessel-type kernel, illustrating that this family covers a large range of DPPs, from the Poisson point process when $\alpha \to 0$, which induces no interaction, to the most repulsive stationary DPP when $\sigma = 0$ and $\alpha = \alpha_{\max}$. The local and global repulsiveness are increasing on each row from left to right and each column from top to bottom. Note that Figure 3.6 (f) is a realization of $DPP(C_B)$ with intensity $\rho = 1$ which, according to [7, Corollary 3.3], is the most globally and the most locally repulsive DPP among all stationary DPPs with intensity $\rho = 1$ and kernel satisfying the condition $\mathcal{K}(\rho)$ in [7].

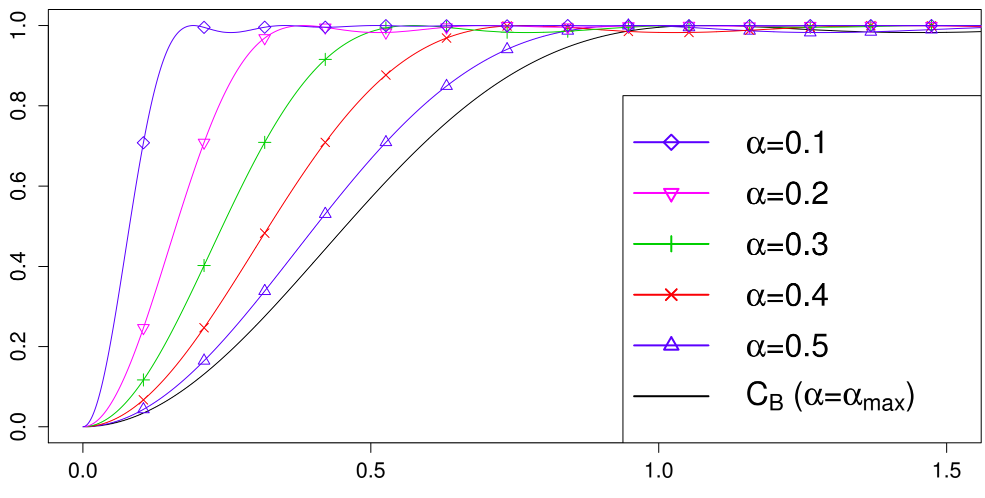


Figure 3.4 – Pcf's of $DPP(C)$ where C is a Bessel-type kernel given by (3.7.1), when $d = 2$, $\rho = 1$, $\sigma = 0$ and different values of α. The case $\alpha = \alpha_{\max} = 1/\sqrt{\pi} \approx 0.56$ corresponds to $C = C_B$.

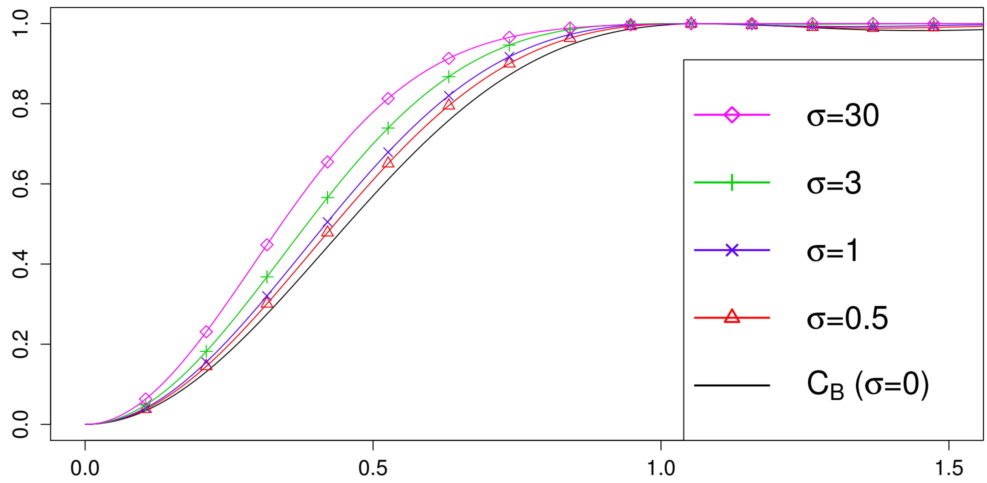


Figure 3.5 – Pcf's of $DPP(C)$ where C is a Bessel-type kernel given by (3.7.1), when $d = 2$, $\rho = 1$, $\alpha = \alpha_{\max}$, and different values of σ. The case $\sigma = 0$ corresponds to $C = C_B$.

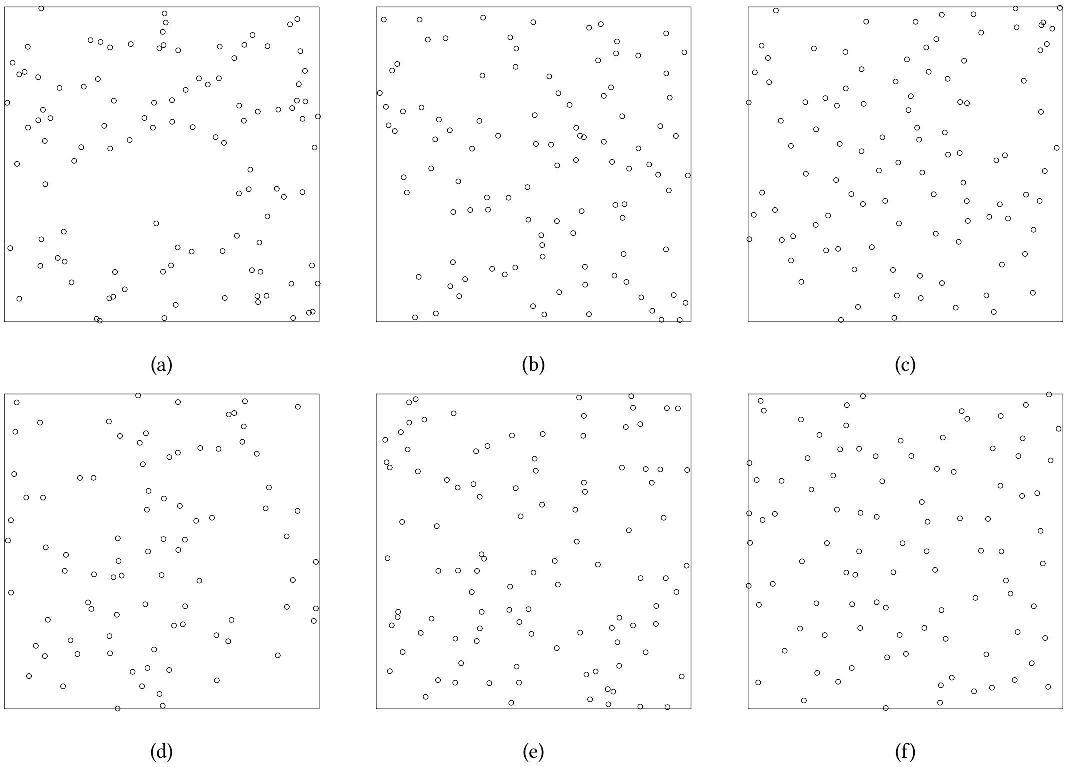


Figure 3.6 – Realizations on $[-5, 5]^2$ of DPPs with Bessel-type kernel (3.7.1) and parameters: (a)-(c) $\rho = 1,\ \sigma = 3$ and from left to right $\alpha = 0.1,\ 0.3,\ \alpha_{\max} = \frac{1}{\sqrt{\pi}} \approx 0.56$, (d)-(f) $\rho = 1, \sigma = 0$ and from left to right $\alpha = 0.1,\ 0.3,\ \alpha_{\max} = \frac{1}{\sqrt{\pi}} \approx 0.56$.

3.7.3 Laguerre-Gaussian family of DPPs

In dimension d, for all $m \in \mathbb{N},\ \alpha > 0,\ \rho > 0$ and $x \in \mathbb{R}^d$, the Laguerre-Gaussian kernel C is defined by

$$C(x) = \frac{\rho}{\binom{m-1+\frac{d}{2}}{m-1}} L_{m-1}^{\frac{d}{2}} \left(\frac{1}{m} \left| \frac{x}{\alpha} \right|^2 \right) e^{-\frac{1}{m}\left|\frac{x}{\alpha}\right|^2}, \quad \forall x \in \mathbb{R}^d, \tag{3.7.2}$$

where $L_{m-1}^{\frac{d}{2}}$ denotes the Laguerre polynomial. According to [7, Proposition 5.3], $DPP(C)$ exists if and only if $\alpha \leq \alpha_{\max}$ where

$$\alpha_{\max}^d = \frac{\binom{m-1+\frac{d}{2}}{m-1}}{\rho (m\pi)^{\frac{d}{2}}}.$$

The pcfs of DPPs with Laguerre-Gaussian kernels are represented in Figures 3.7 and 3.8 for different values of the parameters. Figure 3.7 illustrates the effect of the parameter α when m is fixed while Figure 3.8 illustrates the convergence result: for $\alpha = \alpha_{\max}$, $\lim_{m \to +\infty} C(x) = C_B(x)$. Six realizations of DPPs from this family are shown in Figure 3.9 when $d = 2$. In this plot, the local and global repulsiveness are increasing on each row from left to right and each column from top to bottom.

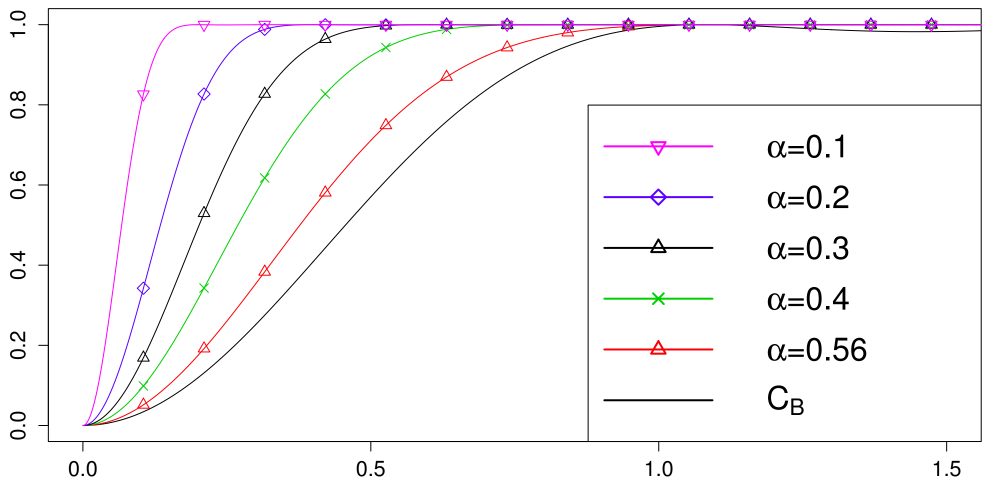


Figure 3.7 – Pcf's of $DPP(C)$ and $DPP(C_B)$ where C is a Laguerre-Gaussian kernel (3.7.2) with $d = 2$, $\rho = 1$, $m = 1$ and different values of α from 0.1 to $\alpha_{\max} \approx 0.56$.

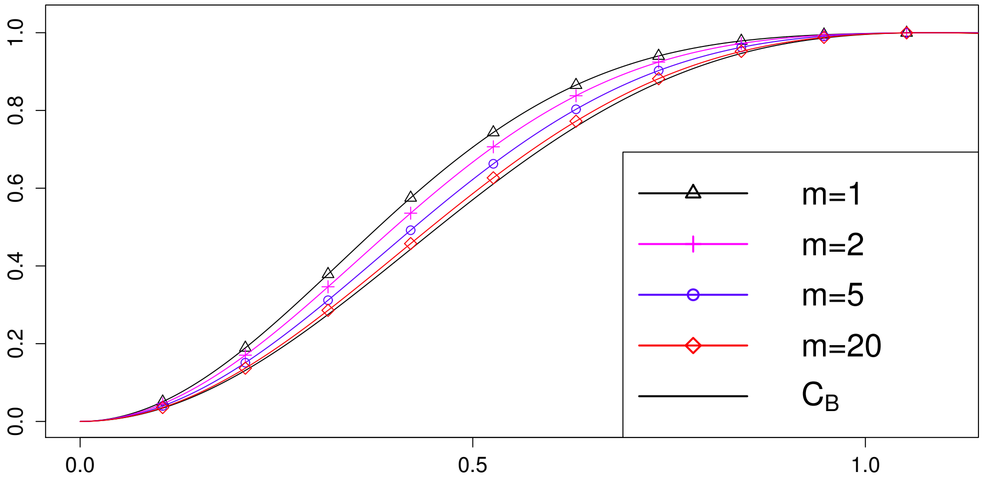


Figure 3.8 – Pcf's of $DPP(C)$ and $DPP(C_B)$ where C is given by (3.7.2) with $d = 2$, $\rho = 1$, $\alpha = \alpha_{\max} \approx 0.56$ and different values of m.

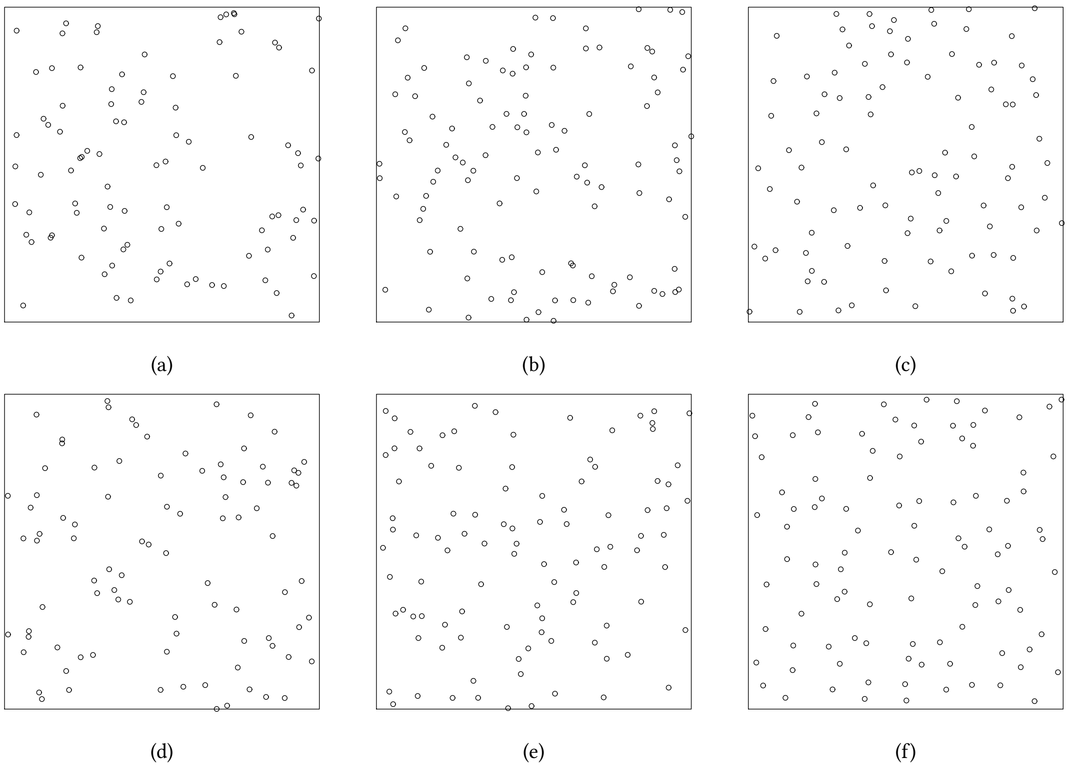


Figure 3.9 – Realizations on $[-5, 5]^2$ of DPPs with Laguerre-Gaussian kernel (3.7.2) and parameters: (a)-(c) $\rho = 1,\ m = 1$ and from left to right $\alpha = 0.1,\ 0.3,\ \alpha_{\max} = \frac{1}{\sqrt{\pi}} \approx 0.56$, (d)-(f) $\rho = 1,\ \sigma = 20$ and from left to right $\alpha = 0.1,\ 0.3,\ \alpha_{\max} = \frac{1}{\sqrt{\pi}} \approx 0.56$.

3.7.4 Families DPPs with compactly supported kernels

This section illustrates some properties of the two parametric families of DPPs with compactly supported kernels introduced in [7, Section 5.3].

The kernels of the first family are defined for $R > 0$ by

$$C_1(x) = \frac{1}{\|h\|^2}[h * h]\left(\frac{2x}{R}\right) C_B\left(x\right), \quad \forall x \in \mathbb{R}^d,$$

where

$$h(x) = \exp\left(\frac{1}{|x|^2 - 1}\right) \mathbf{1}_{\{|x|<1\}}, \quad \forall x \in \mathbb{R}^d.$$

According to [7, Proposition 4.1], C_1 defines a DPP for all $\rho > 0$ and $R > 0$.

The kernels of the second family take the form $C_2 = u * u$ where

$$u(x) = \sqrt{\rho}\,\beta(R,\alpha)\left(1 - \frac{R^{\frac{d}{2}-1}}{2^{\frac{d}{2}-1} J_{\frac{d}{2}-1}(\frac{R}{2\alpha})} \frac{J_{\frac{d}{2}-1}(\left|\frac{x}{\alpha}\right|)}{|x|^{\frac{d}{2}-1}}\right) \mathbf{1}_{\left\{|x| \leq \frac{R}{2}\right\}},$$

and

$$\beta(R,\alpha) = \left[\frac{R^{d-1}\pi^{d/2}}{2^{d-1}\Gamma\left(\frac{d}{2}\right)}\left(\frac{R}{d} - 4\alpha \frac{J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J_{\frac{d}{2}-1}\left(\frac{R}{2\alpha}\right)} + \frac{R}{2}\left(1 - \frac{J_{\frac{d}{2}-2}(\frac{R}{2\alpha}) J_{\frac{d}{2}}(\frac{R}{2\alpha})}{J^2_{\frac{d}{2}-1}\left(\frac{R}{2\alpha}\right)}\right)\right)\right]^{-\frac{1}{2}}.$$

According to [7, Proposition 5.4], C_2 is the kernel of a DPP if and only if α is such that $|\mathcal{F}(u)| \leq 1$ where $\mathcal{F}(u)$ is given in [7]. Given R and ρ, this condition must be checked numerically. No theoretical results are available to claim the existence of an admissible α, but from our experience, for any R and ρ, the set of all admissible α is infinite and bounded. We denote by $\alpha_{\max}$ the maximal value of α such that $|\mathcal{F}(u)| \leq 1$.

Figure 3.10 illustrates that, when the range R is fixed, $DPP(C_2)$ is more repulsive than $DPP(C_1)$. Figure 3.11 shows the pcfs of $DPP(C_2)$ for different values of R in comparison with the pcf of $DPP(C_B)$. Even if the behavior of the former when $R = 3$ seems close to the later, C_2 does not converge to C_B when R tends to infinity, contrary to C_1. Figure 3.12 illustrates the effect of α on C_2 when ρ and R are fixed. Finally we show in Figure 3.13 six realizations of $DPP(C_2)$.

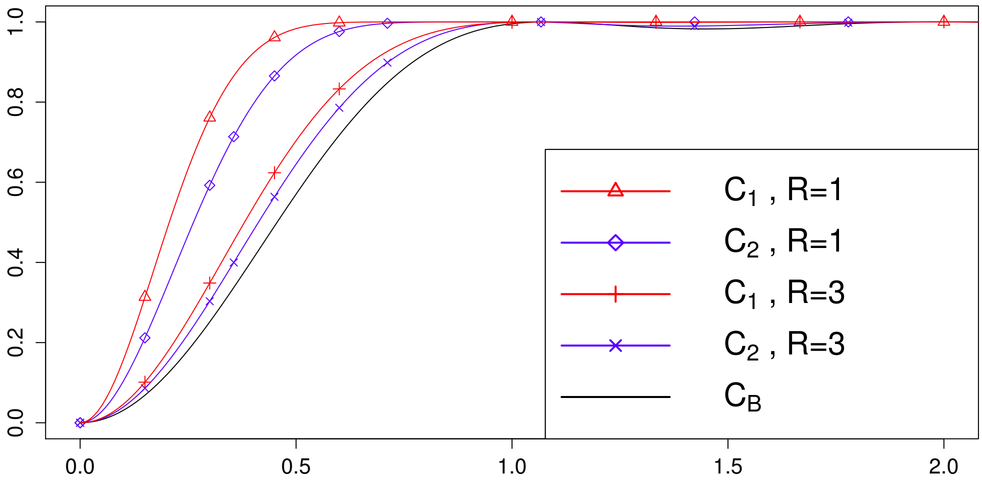


Figure 3.10 – Pcf's of $DPP(C_1)$, $DPP(C_2)$ and $DPP(C_B)$ when $d = 2$, $\rho = 1$, $\alpha = \alpha_{\max}$ for C_2 and different values of R.

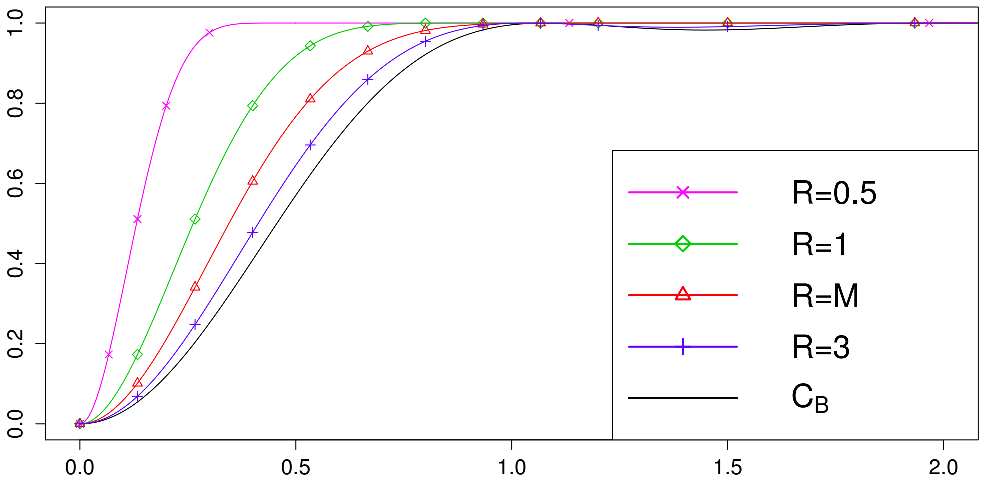


Figure 3.11 – Pcf's of $DPP(C_2)$ and $DPP(C_B)$ when $d = 2$, $\rho = 1$, $\alpha = \alpha_{\max}$ and different values of R (where $M \approx 1.357$ is defined in [7]).

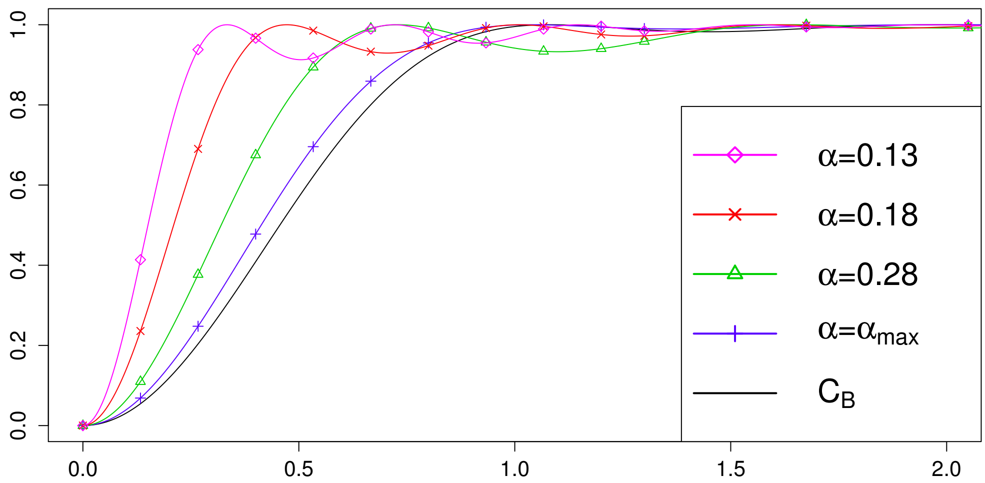


Figure 3.12 – Pcf's of $DPP(C_2)$ and $DPP(C_B)$ when $d = 2$, $\rho = 1$, $R = 3$ and different values of α.

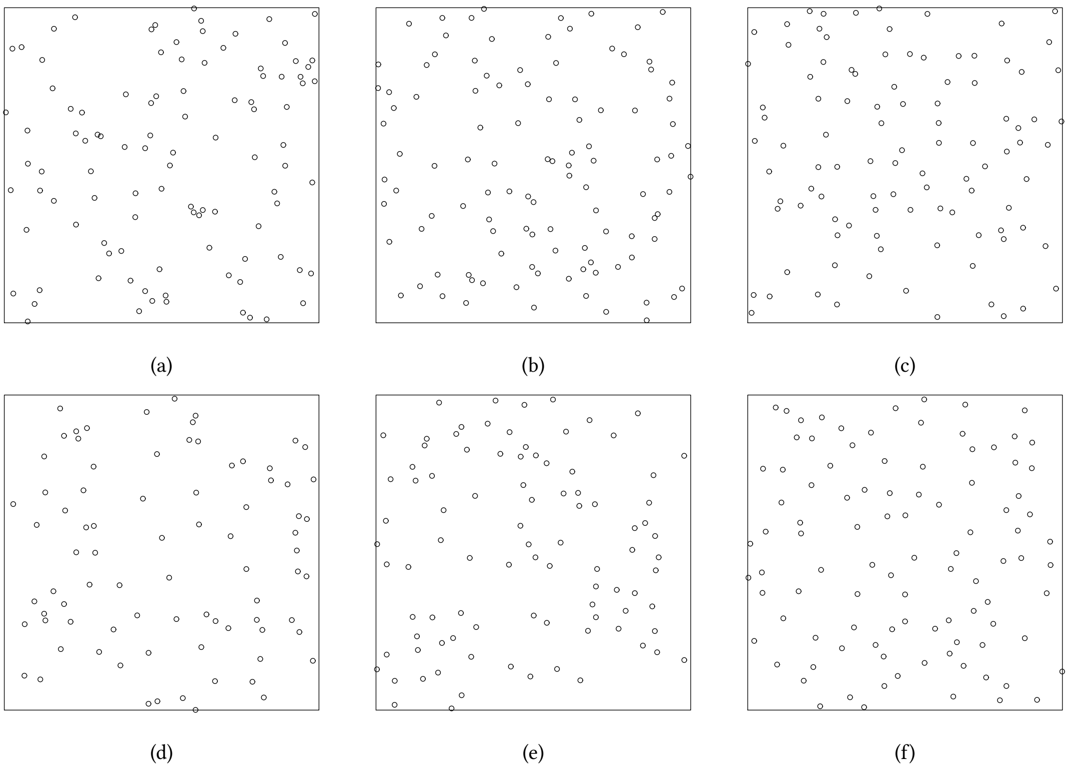


Figure 3.13 – Realizations on $[-5, 5]^2$ of $DPP(C_2)$ with $\rho = 1$ and parameters: (a) $R = 0.5,\ \alpha = \alpha_{\max}$, (b) $R = 1,\ \alpha = \alpha_{\max}$, (c) $R = M \approx 1.357,\ \alpha = \alpha_{\max}$, (d)-(f) $\rho = 1,\ R = 3$ and from left to right $\alpha = 0.13,\ 0.28,\ \alpha_{\max}$.

4

Brillinger mixing of determinantal point processes

The Brillinger mixing property is based on the moments of a stationary point process and is used to study the asymptotic behaviour of some functionals of the process including usual statistics, see for instance [25] or [40]. Several examples of Brillinger mixing point processes are known, for instance Poisson cluster processes and Matérn hardcore point processes (of type I, type II and some generalizations as in [60]), see [30] and [28]. To the best of our knowledge, the Matérn hardcore models are the only repulsive stationary point processes that have been proved to be Brillinger mixing. We show in this chapter that under the mild condition $\mathcal{K}(\rho)$, DPPs provide a flexible class of Brillinger mixing point processes. As an application, we derive the asymptotic behaviour of several quantities of interest as the estimators of the intensity and of the pcf.

4.1 Preliminaries

In this section, we review the definition of the cumulant and factorial cumulant moment measures of a point process $\mathbf{X}$ as well as their reduced version. These are at the basis of the Brillinger mixing property defined in the following. The relations with the Laplace and the probability generating functionals of $\mathbf{X}$ are also described. We assume that, for any bounded set A, the random variable $\mathbf{X}(A)$ has moment of any order. This ensures that the quantities introduced in this section are well defined. Further details may be found in [9, 10] and [40].

Definition 4.1.1. *For $k \in \mathbb{N}$, the cumulant (or semi-invariant) of the k random variables $X_1, \dots, X_k$ is, if it exists,*

$$\mathrm{Cum}(X_1, \dots, X_k) = \frac{\partial^k}{\partial t_1 \dots \partial t_k} \log \mathbb{E}\left[\exp\left(\sum_{i=1}^k t_i X_i\right)\right]\Bigg|_{t_1 = \dots = t_k = 0}.$$

The k-th order cumulant of the random variable X is $\mathrm{Cum}_k(X) := \mathrm{Cum}(X, \dots, X)$.

The notion of cumulant of random variables extends to point processes as follows.

Definition 4.1.2. *For $k \in \mathbb{N}$, the k-th order cumulant moment measure γ_k of a point process $\mathbf{X}$ is a locally finite signed measure on $\mathbb{R}^{dk}$ defined for any bounded measurable sets $A_1, \dots, A_k$ in $\mathbb{R}^d$ by*

$$\gamma_k\left(\prod_{i=1}^{k} A_i\right) = \mathrm{Cum}\left(\sum_{x\in\mathbf{X}} \mathbf{1}_{\{x\in A_1\}}, \dots, \sum_{x\in\mathbf{X}} \mathbf{1}_{\{x\in A_k\}}\right).$$

Definition 4.1.3. *For $k \in \mathbb{N}$, the k-th order factorial moment measure $\alpha^{(k)}$ of a point process $\mathbf{X}$ is a locally finite measure on $\mathbb{R}^{dk}$ defined by*

$$\alpha^{(k)}\left(\prod_{i=1}^{k} A_i\right) = \mathbb{E}\left(\sum_{(x_1,\dots,x_k)\in\mathbf{X}^k}^{\neq} \mathbf{1}_{\{x_1\in A_1,\dots,x_k\in A_k\}}\right)$$

where $\mathbb{E}$ is the expectation over the distribution of $\mathbf{X}$ and the symbol $\neq$ over the sum means that we consider only mutually disjoints k-tuples of points $x_1, \dots, x_k$.

Definition 4.1.4. *For $k \in \mathbb{N}$, the k-th order factorial cumulant moment measure $\gamma_{[k]}$ of a point process with factorial moment measure $\alpha^{(r)}$, for $r \leq k$, is a locally finite signed measure on $\mathbb{R}^{dk}$ defined for any bounded measurable sets $A_1, \dots, A_k$ in $\mathbb{R}^d$ by*

$$\gamma_{[k]}\left(\prod_{i=1}^{k} A_i\right) = \sum_{j=1}^{k}(-1)^{j-1}(j-1)! \sum_{K_1,\dots,K_j\in\mathcal{P}_j^k} \prod_{i=1}^{j} \alpha^{(|K_i|)}\left(\prod_{k_i\in K_i} A_{k_i}\right),$$

where for all $j \leq k$, $\mathcal{P}_j^k$ denote the set of all partitions of $\{1, \dots, k\}$ into j non empty sets $K_1, \dots, K_j$.

Notice that the factorial cumulant measure is symmetric. In other words, in the last definition, $\gamma_{[k]}\left(\prod_{i=1}^{k} A_i\right) = \gamma_{[k]}\left(\prod_{i=1}^{k} A_{\sigma(i)}\right)$ for any permutation σ of $\{1, \dots, k\}$. Since we only consider in this study stationary point processes, we can define the so-called reduced version of the previous measures.

Definition 4.1.5. *For any $k \geq 2$, the reduced k-th order factorial cumulant moment measure $\gamma_{[k]}^{red}$ of a stationary point process is a locally finite signed measure on $\mathbb{R}^{d(k-1)}$ defined for any bounded measurable sets $A_1, \dots, A_k$ in $\mathbb{R}^d$ by*

$$\gamma_{[k]}\left(\prod_{i=1}^{k} A_i\right) = \int_{A_k} \gamma_{[k]}^{red}\left(\prod_{i=1}^{k-1}(A_i - x)\right) dx$$

where for $i = 1, \dots, k-1$, $A_i - x$ is the translation of the set A_i by $-x$.

We may define similarly for any $k \geq 2$, γ_k^{red} the reduced k-th order cumulant moment measure. An important property of signed measures is given by the following theorem allowing the definition of the total variation of a signed measure.

Theorem 4.1.6 (Hahn-Jordan decomposition, see [13, Theorem 5.6.1]). *For any signed measure ν, there exist two measures ν^+ and ν^- uniquely determined by ν such that at least one of them is finite and*

$$\nu = \nu^+ - \nu^-.$$

Definition 4.1.7 ([13, p.179]). *Let ν be a signed measure with Hahn-Jordan decomposition $\nu = \nu^+ - \nu^-$. The total variation measure $|\nu|$ of ν is defined by*

$$|\nu| = \nu^+ + \nu^-.$$

Following Theorem 4.1.6, for $k \geq 2$, we denote the Hahn-Jordan decomposition of the reduced k-th order moment factorial cumulant measure $\gamma_{[k]}^{red} = \gamma_{[k]}^{+red} - \gamma_{[k]}^{-red}$.

Definition 4.1.8. *A point process is Brillinger mixing if, for $k \geq 2$, we have*

$$\left|\gamma_{[k]}^{red}\right| \left(\mathbb{R}^{d(k-1)}\right) < +\infty.$$

Notice that the last definition remains the same if, for $k \geq 2$, $\gamma_{[k]}^{red}$ is replaced by γ_k^{red}. This is an immediate consequence of Proposition 4.1.14 presented later, the factorial cumulant moment measure being a polynomial function of the cumulant moment measures of lower order and reciprocally.

The different moment measures are related to the power series expansion of the Laplace and the probability generating functionals of $\mathbf{X}$.

Definition 4.1.9. *The Laplace functional $L_{\mathbf{X}}$ of a point process $\mathbf{X}$ is defined for any bounded measurable function f that vanishes outside a bounded set of $\mathbb{R}^d$ by*

$$L_{\mathbf{X}}(f) = \mathbb{E}\left(e^{-\sum_{x\in\mathbf{X}} f(x)}\right).$$

Definition 4.1.10. *The probability generating functional of a point process $\mathbf{X}$ is defined for any function h from $\mathbb{R}^d$ into $[0, 1]$, such that $1 - h$ vanishes outside a bounded set, by*

$$G_{\mathbf{X}}(h) = \mathbb{E}\left(\exp\left(\sum_{x\in\mathbf{X}} \log h(x)\right)\right).$$

Notice that for any function h defined as in Definition 4.1.10 and taking values within a closed subset of $(0, 1]$, we have

$$G_{\mathbf{X}}(h) = L_{\mathbf{X}}(-\log(h)).$$

Proposition 4.1.11 ([10, Section 9.5]). *Let $\mathbf{X}$ be a point process with cumulant and factorial cumulant moment measures, of order $k \geq 1$, γ_k and $\gamma_{[k]}$, respectively. Let f and η be bounded measurable functions on $\mathbb{R}^d$ that vanish outside a bounded set. Assume further that η takes value in $[0, 1]$. Then, for all $N \in \mathbb{N}$ and for all $s \in \mathbb{R}^+$, we have the following power series expansion*

$$\log L_{\mathbf{X}}(sf) = \sum_{j=1}^{N} \frac{(-s)^j}{j!} \int f(x_1)f(x_2)\ldots f(x_j)\gamma_j(dx_1 \times dx_2 \times \ldots \times dx_j) + o(s^N),$$

$$\log G_{\mathbf{X}}(1 - s\eta) = \sum_{j=1}^{N} \frac{(-s)^j}{j!} \int \eta(x_1)\eta(x_2)\ldots\eta(x_j)\gamma_{[j]}(dx_1 \times dx_2 \times \ldots \times dx_j) + o(s^N).$$

We conclude this section by giving the relations between the factorial cumulant moment measure and the ordinary cumulant measure in Proposition 4.1.14. To this end, we recall the definition of the Stirling numbers of the first and second kind and refer to [9, Section 5.2] for a detailed presentation.

Definition 4.1.12. *For any integers n and k, the factorial powers of n, denoted by $n^{[k]}$ is defined by*

$$n^{[k]} = n(n-1)\dots(n-k+1)\mathbf{1}_{\{0\le k\le n\}}.$$

Definition 4.1.13. *For $n \ge k \ge j$, the Stirling numbers of the first kind $D_{j,k}$ and of the second kinds $\Delta_{j,k}$ are defined by the relations*

$$\begin{aligned} n^{[k]} &= \sum_{j=1}^{k}(-1)^{k-j}D_{j,k}n^j, \\ n^k &= \sum_{j=1}^{k}\Delta_{j,k}n^{[j]}. \end{aligned}$$

Proposition 4.1.14. *Let A be a bounded set of $\mathbb{R}^d$. For $k \in \mathbb{N}$, we have the relations*

$$\begin{aligned} \gamma_{[k]}(A^k) &= \sum_{j=1}^{k}(-1)^{k-j}D_{j,k}\,\gamma_j(A^j), \\ \gamma_k(A^k) &= \sum_{j=1}^{k}\Delta_{j,k}\,\gamma_{[j]}(A^j). \end{aligned}$$

Proof. We denote by L and G the Laplace and probability generating functionals of a point process with factorial cumulant moment and the cumulant moment measures, of order $k \ge 1$, $\gamma_{[k]}$ and γ_k, respectively. Let A be a bounded set. By Proposition 4.1.11, for all $N \in \mathbb{N}$ and any non-negative real s, we have

$$\log G(1 - s\mathbf{1}_A) = \sum_{k=1}^{N}\frac{(-s)^k}{k!}\gamma_{[k]}(A^k) + o(s^N) \tag{4.1.1}$$

As noticed after Definition 4.1.10, we have

$$\begin{aligned} \log G(1 - s\mathbf{1}_A) &= \log L(-\log(1 - s\mathbf{1}_A)) \\ &= \log L(-\log(1-s)\mathbf{1}_A)). \end{aligned}$$

Note that $s \sim -\log(1-s)$ as $s \to 0$. Thus, by Proposition 4.1.11, we have

$$\log G(1 - s\mathbf{1}_A) = \sum_{j=1}^{N}\frac{[\log(1-s)]^j}{j!}\gamma_j(A^j) + o(s^N).$$

Then, we deduce by [1, (24.1.3.I.B)] that

$$\begin{aligned}\log G(1-s\mathbf{1}_A) &= \sum_{j=1}^{N}\frac{\gamma_j(A^j)}{j!}j!\sum_{k=j}^{N}(-1)^{k-j}D_{j,k}\frac{(-s)^k}{k!}+o(s^N)\\ &= \sum_{k=1}^{N}\frac{(-s)^k}{k!}\sum_{j=1}^{k}(-1)^{k-j}D_{j,k}\gamma_j(A^j)+o(s^N).\end{aligned} \tag{4.1.2}$$

We conclude by identifying the coefficients in (4.1.1) and (4.1.2). The proof of the second formula is done in the same way by starting with the other powers expansion in Proposition 4.1.11 and use [1, (24.1.4.I.B)] instead of [1, (24.1.3.I.B)] □

4.2 Main result

In this section, we prove in Theorem 4.2.2 below that a DPP with kernel verifying the condition $\mathcal{K}(\rho)$ is Brillinger mixing. We recall that this mixing property involves the factorial cumulant moments of the DPP. By definition, if $\mathbf{X}$ is a stationary DPP then for any bounded set A and $k \geq 1$, $\mathbb{E}(\mathbf{X}(A)^k)$ is finite. Hence, in this case, all moment measures introduced in the last section are well defined. It is not easy to deduce the factorial cumulant measures from the initial Definition 4.1.4. However, the power series expansion of the log-Laplace functional in Proposition 4.1.11, which is known for a DPP, allows us to derive a closed form expression as stated in the following lemma.

Lemma 4.2.1. *Consider a DPP with kernel C verifying condition $\mathcal{K}(\rho)$ and, for $k \in \mathbb{N}$, denote its k-th factorial cumulant moment measure by $\gamma_{[k]}$. For every measurable bounded set A in $\mathbb{R}^d$ and $k \geq 2$, we have*

$$\gamma_{[k]}(A^k) = (-1)^{k+1}(k-1)!\int_{A^k} C(x_2-x_1)\dots C(x_1-x_k)dx_1\dots dx_k.$$

Proof. By [54, Proposition 3.9], we deduce that for any bounded set $A \subset \mathbb{R}^d$ and s small enough,

$$\log\left(L_{\mathbf{X}}(s\mathbf{1}_A)\right) = \sum_{j=1}^{\infty}\frac{(-s)^j}{j!}\sum_{n=1}^{j}(-1)^{n+1}\sum_{\substack{j_1+\dots+j_n=j\\ j_1,\dots,j_n\geq 1}}\frac{j!}{n\cdot j_1!j_2!\cdots j_n!}\int_{A^n} C(x_2-x_1)\dots C(x_1-x_n)dx_1\dots dx_n.$$

Then, by Proposition 4.1.11, we have by the last equation that for all $j \in \mathbb{N}$ and any bounded set $A \subset \mathbb{R}^d$,

$$\gamma_j(A^j) = \sum_{n=1}^{j}(-1)^{n+1}\sum_{\substack{j_1+\dots+j_n=j\\ j_1,\dots,j_n\geq 1}}\frac{j!}{n\cdot j_1!j_2!\cdots j_n!}\int_{A^n} C(x_2-x_1)\dots C(x_1-x_n)dx_1\dots dx_n.$$

Thus, by Proposition 4.1.14, we have for $k \geq 2$,

$$\gamma_{[k]}(A^k) = \sum_{j=1}^{k} (-1)^{k-j} D_{j,k} \sum_{n=1}^{j} (-1)^{n+1} \sum_{\substack{j_1+\ldots+j_n=j \\ j_1,\ldots,j_n \geq 1}} \frac{j!}{n \cdot j_1! j_2! \cdots j_n!} \int_{A^n} C(x_2 - x_1) \ldots C(x_1 - x_n) dx_1 \ldots dx_n. \tag{4.2.1}$$

By [1, (24.1.2.I.B)], it is easily seen that for a given $n = 1, \ldots, k$, we have

$$\begin{aligned} \sum_{\substack{j_1+\ldots+j_n=j \\ j_1,\ldots,j_n \geq 1}} \frac{j!}{j_1! j_2! \cdots j_n!} &= \sum_{j_1+\ldots+j_n=j} \frac{j!}{j_1! j_2! \cdots j_n!} - \sum_{j_1+\ldots+j_{n-1}=j} \frac{j!}{j_1! j_2! \cdots j_{n-1}!} \\ &= n^j - (n-1)^j. \end{aligned} \tag{4.2.2}$$

By Definition 4.1.12,

$$\sum_{j=1}^{k} (-1)^{k-j} D_{j,k} (n^j - (n-1)^j) = n^{[k]} - (n-1)^{[k]} \tag{4.2.3}$$

which is null for every $n < k$. Therefore, by (4.2.2) and (4.2.3), only the terms $n = k$ is non null in the sum (4.2.1). $\square$

We are now in position to prove the main result of this chapter that gives conditions on the DPP's kernel to ensure that the associated DPP is Brillinger mixing.

Theorem 4.2.2. *A DPP with kernel verifying the condition $\mathcal{K}(\rho)$, for a given $\rho > 0$, is Brillinger mixing.*

Proof. For all $t \in \mathbb{R}^+$, we have by taking $f = \mathbf{1}_{\{[-t,t]^d\}}$ in Definition 4.1.5,

$$\gamma_{[k]}([-t,t]^{dk}) = \int_{\mathbb{R}^d} \mathbf{1}_{[-t,t]^d}(x) \gamma_{[k]}^{red} \left(\left([-t,t]^d - x \right)^{k-1} \right) dx$$

where for a given $A \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$, $A - x$ is the translation of the set A by x. Then, by Lemma 4.2.1 we have

$$\int_{\mathbb{R}^d} \mathbf{1}_{[-t,t]^d}(x) \gamma_{[k]}^{red} \left(\left([-t,t]^d - x \right)^{k-1} \right) dx = (-1)^{k+1} (k-1)! \, I_k(t) \tag{4.2.4}$$

where for all $k \geq 1$ and $t \in \mathbb{R}^+$, $I_k(t) := \int_{[-t,t]^{dk}} C(x_2 - x_1) \ldots C(x_1 - x_k) dx_1 \ldots dx_k$. Since C verifies the condition $\mathcal{K}(\rho)$, by Mercer's theorem, see also [43, Section 2.3], we have for all $t \in \mathbb{R}^+$,

$$C(x - y) = \sum_{n \in \mathbb{N}} \lambda_n(t) \phi_n(x) \phi_n(y), \quad \forall (x, y) \in [-t,t]^d,$$

where for all $n \in \mathbb{N}$, $\{\phi_n\}_{n \in \mathbb{N}}$ is an orthonormal basis of $L^2\left([-t,t]^d\right)$ and $\lambda_n(t)$ belongs to $[0,1]$ by [46, Theorem 4.5.5]. Then, by orthogonality of the basis $\{\phi_n\}_{n \in \mathbb{N}}$, we have for all $t \in \mathbb{R}^+$ and $k \geq 1$,

$$I_k(t) = \sum_{n \in \mathbb{N}} \lambda_n^k(t) \leq \sum_{n \in \mathbb{N}} \lambda_n(t) = I_1(t). \tag{4.2.5}$$

Notice that $I_1(t) = \int_{[-t,t]^d} \rho dx = O(t^d)$. Thus, by Theorem 4.1.6, (4.2.4) and (4.2.5), there exist a constant $\kappa > 0$ and $T \in \mathbb{R}^+$ such that for all $t \geq T$,

$$\left| \int_{\mathbb{R}^d} \mathbf{1}_{\{[-t,t]^d\}}(x) \left[\gamma_{[k]}^{+red} \left(\left([-t,t]^d - x \right)^{k-1} \right) - \gamma_{[k]}^{-red} \left(\left([-t,t]^d - x \right)^{k-1} \right) \right] dx \right| \leq \kappa t^d. \tag{4.2.6}$$

Henceforth, we assume $t \geq T$. By Theorem 4.1.6, at least one of the measure $\gamma_{[k]}^{+red}$ or $\gamma_{[k]}^{-red}$ is finite. In a first time, assume that $\gamma_{[k]}^{-red}$ is finite. Thus, by (4.2.6) and the monotonicity of the measure $\gamma_{[k]}^{-red}$, we have

$$\int_{[-t,t]^d} \gamma_{[k]}^{+red} \left(\left([-t,t]^d - x \right)^{k-1} \right) dx \leq t^d \left(\kappa + 2^d \gamma_{[k]}^{-red} (\mathbb{R}^{d(k-1)}) \right). \tag{4.2.7}$$

By positivity of $\gamma_{[k]}^{+red}$,

$$\int_{\left[\frac{-t}{2},\frac{t}{2}\right]^d} \gamma_{[k]}^{+red} \left(\left([-t,t]^d - x \right)^{k-1} \right) dx \leq \int_{[-t,t]^d} \gamma_{[k]}^{+red} \left(\left([-t,t]^d - x \right)^{k-1} \right) dx. \tag{4.2.8}$$

Further, for all $(x,y) \in \left[\frac{-t}{2},\frac{t}{2}\right]^{2d}$, $y + x \in [-t,t]^d$, so for all $x \in \left[\frac{-t}{2},\frac{t}{2}\right]^d$ we have $\left[\frac{-t}{2},\frac{t}{2}\right]^d \subset [-t,t]^d - x$. It follows by (4.2.8) and the monotonicity of $\gamma_{[k]}^{+red}$ that

$$\int_{\left[\frac{-t}{2},\frac{t}{2}\right]^d} \gamma_{[k]}^{+red} \left(\left[\frac{-t}{2},\frac{t}{2}\right]^{d(k-1)} \right) dx \leq \int_{\left[\frac{-t}{2},\frac{t}{2}\right]^d} \gamma_{[k]}^{+red} \left(\left([-t,t]^d - x \right)^{k-1} \right) dx. \tag{4.2.9}$$

Hence by (4.2.7)- (4.2.9), we have

$$\gamma_{[k]}^{+red} \left(\left[\frac{-t}{2},\frac{t}{2}\right]^{d(k-1)} \right) \leq \left(\kappa + 2^d \gamma_{[k]}^{-red} (\mathbb{R}^{d(k-1)}) \right).$$

By letting t tend to infinity in the last equation, we prove that $\gamma_{[k]}^{+red}$ is finite and so is $\left|\gamma_{[k]}^{red}\right|$ by Definition 4.1.7. Finally, the proof is done in the same ways if we assume that $\gamma_{[k]}^{+red}$ is finite. □

4.3 Statistical applications

Many applications of the Brillinger mixing property may be found in [26], [27],[29], [36] and [37]. We present in this section two kinds of these applications to the DPPs. First, we prove a general central limit theorem that is of main interest for the study of the asymptotic properties of some estimators. For instance, we give in this section an application to the estimator of the intensity and in the next chapter on some minimum contrast estimators. Second, we present several properties of an estimator of the pcf. In particular, we state a central limit theorem for a pointwise estimator of the pcf and for the integrated squared error of this estimator.

4.3.1 Asymptotic behaviour of functionals of order p

We present an important consequence of the Brillinger mixing property, namely the asymptotic properties of a wide class of functionals of the process including a central limit theorem and the convergence of its moments. A first theorem was mentioned in [40] and proved in [35]. We present here a more general version that yields in particular the asymptotic normality of standard statistics as the natural estimator of the intensity of the process.

For a given set D of $\mathbb{R}^d$, we denote by ∂D and $|D|$ the boundary and the volume of the set D, respectively.

Definition 4.3.1. *A sequence of subsets $\{D_n\}_{n\in\mathbb{N}}$ of $\mathbb{R}^d$ is called regular if for all $n \in \mathbb{N}$, $D_n \subset D_{n+1}$, D_n is compact, convex and there exist constants α_1 and α_2 such that*

$$\alpha_1 n^d \le |D_n| \le \alpha_2 n^d,$$
$$\alpha_1 n^{d-1} \le \mathcal{H}_{d-1}(\partial D_n) \le \alpha_2 n^{d-1}$$

where $\mathcal{H}_{d-1}$ is the $(d-1)$-dimensional Hausdorff measure.

Note that any sequence of subsets as above grows to $\mathbb{R}^d$ in all directions. For $p \ge 1$, let f_D be a function from $\mathbb{R}^{dp}$ into $\mathbb{R}$ that depends on a given set $D \subset \mathbb{R}^d$ and define for a stationary point process $\mathbf{X}$,

$$N_p(f_D) := \sum_{(x_1,\dots,x_p)\in\mathbf{X}^p} f_D(x_1,\dots,x_p).$$

By letting the set D in the last equation be a sequence of regular subsets $\{D_n\}_{n\in\mathbb{N}}$, we have under some suitable conditions on the function f_{D_n}, the following central limit theorem on the sequence $\{N_p(f_{D_n})\}_{n\in\mathbb{N}}$.

Theorem 4.3.2. *Let $\{D_n\}_{n\in\mathbb{N}}$ and $\{\widetilde{D}_n\}_{n\in\mathbb{N}}$ be two sequences of regular sets in the sense of Definition 4.3.1 such that $\frac{|\widetilde{D}_n|}{|D_n|} \xrightarrow{n\to+\infty} \kappa$ for a given $\kappa > 0$. For all $n \in \mathbb{N}$, let $\{f_{D_n}\}_{n\in\mathbb{N}}$ be a family of functions from $\mathbb{R}^{dp}$ into $\mathbb{R}$. Assume that there exists a bounded function F from $\mathbb{R}^{d(p-1)}$ into $\mathbb{R}^+$ with compact support such that for all $n \in \mathbb{N}$ and $(x_1,\dots,x_p) \in \mathbb{R}^{dp}$,*

$$|f_{D_n}(x_1,\dots,x_p)| \le \frac{1}{|\widetilde{D}_n|}\mathbf{1}_{\{x_1\in D_n\}}F(x_2-x_1,\dots,x_p-x_1). \tag{4.3.1}$$

Assume further that the point process $\mathbf{X}$ is ergodic, admits moment of any order and is Brillinger mixing in the sense of Definition 4.1.8. Then, for all $k \ge 2$, we have

$$\mathrm{Cum}_k\left(\sqrt{|D_n|}N_p(f_{D_n})\right) = O\left(|D_n|^{1-\frac{k}{2}}\right). \tag{4.3.2}$$

Moreover, if there exists a real σ such that

$$\mathrm{Var}\left(\sqrt{|D_n|}N_p(f_{D_n})\right) \xrightarrow[n\to+\infty]{} \sigma^2, \tag{4.3.3}$$

we have the convergence

$$\sqrt{|D_n|}\left[N_p(f_{D_n}) - \mathbb{E}(N_p(f_{D_n}))\right] \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(0,\sigma^2) \tag{4.3.4}$$

and the convergence of all moments to the corresponding moments of $\mathcal{N}(0,\sigma^2)$.

By (4.3.2), the variance given in (4.3.3) is uniformly bounded with respect to $n \in \mathbb{N}$. If D_n and f_{D_n} in (4.3.3) are sufficiently generic, the convergence of the variance holds true. However, in the general case, it has to be checked. For this purpose, it is convenient to express the variance in (4.3.3) in terms of the factorial cumulant moment measures of $\mathbf{X}$. However, the calculus becomes fastidious for large values of p. In appendix, we detail this expression for the important situations $p = 1$ and $p = 2$ with $f_{D_n}(x_1, x_2) = 0$ for $x_1 \neq x_2$, see Lemmas 4.6.1, 4.6.2 and 4.6.3.

Soshnikov in [56] proves that a stationary DPP is ergodic. Moreover, a DPP with kernel verifying the condition $\mathcal{K}(\rho)$ admits, by definition, moments of any order and is Brillinger mixing by Theorem 4.2.2. Hence, Theorem 4.3.2 holds for a DPP with kernel verifying the condition $\mathcal{K}(\rho)$ provided the condition on the functions f_{D_n} is verified. As a direct application when $p = 1$, we retrieve with the following corollary, a result of [57] giving the asymptotic normality of the estimator of the intensity of a stationary DPP.

Corollary 4.3.3. *Let $\mathbf{X}$ be a DPP with kernel C verifying the condition $\mathcal{K}(\rho)$ for a given $\rho > 0$ and $\{D_n\}_{n\in\mathbb{N}}$ be a family of regular sets. Define for all $n \in \mathbb{N}$,*

$$\widehat{\rho}_n = \frac{1}{|D_n|} \sum_{x\in\mathbf{X}} \mathbf{1}_{\{x\in D_n\}}.$$

We have the convergence

$$\sqrt{|D_n|}\,(\widehat{\rho}_n - \rho) \xrightarrow[n\to+\infty]{distr.} N(0, \sigma^2)$$

where $\sigma^2 = \lim_{n\to+\infty} Var\left(\sqrt{|D_n|}\widehat{\rho}_n\right) = \rho - \int_{\mathbb{R}^d} C(x)^2 dx$.

The proof of the last corollary follows by taking $p = 1$ and, for $x \in \mathbb{R}^d$ and $n \in \mathbb{N}$, $f_{D_n}(x) = \frac{1}{|D_n|}\mathbf{1}_{\{x\in D_n\}}$ in Theorem 4.3.2. In this case, the convergence (4.3.3) holds by Lemma 4.6.1 and a straightforward calculus.

4.3.2 Applications to statistics of order two, the pair correlation function.

We adapt in this section some results of [26] and [27] on a kernel estimator of the pcf of an isotropic point process. Note that in this case, the pcf g is a radial function. Thus, there exists a function $\widetilde{g}$ such that for all $x \in \mathbb{R}^d$, $g(x) = \widetilde{g}(|x|)$. To shorten, we denote in this section g in place of $\widetilde{g}$ so the pcf is considered as a function from $\mathbb{R}^+$ into $\mathbb{R}^+$. Let $\{D_n\}_{n\in\mathbb{N}}$ be a sequence of regular subsets of $\mathbb{R}^d$ in the sense of Definition 4.3.1, $\{b_n\}_{n\in\mathbb{N}}$ be a positive sequence of real, and k be a function from $\mathbb{R}$ into $\mathbb{R}$. Further, define for all $n \in \mathbb{N}$ and $z \in \mathbb{R}^d$, $D_n^z := \{u, u + z \in D_n\} = D_n - z$. For an isotropic point process $\mathbf{X}$ with intensity ρ on $\mathbb{R}^d$, define for all $n \in \mathbb{N}$ and $t > 0$, the following kernel estimator of the pcf,

$$\widehat{g}_n(t) = \frac{1}{\sigma_d t^{d-1}\widehat{\rho}_n^2} \sum_{\substack{(x,y)\in\mathbf{X}^2 \\ x\neq y}} \mathbf{1}_{\{x\in D_n,\, y\in D_n\}} \frac{1}{b_n|D_n \cap D_n^{x-y}|} k\left(\frac{t - |x-y|}{b_n}\right) \tag{4.3.5}$$

where $\widehat{\rho}_n$ is as in Corollary 4.3.3 and for $d \geq 2$, $\sigma_d = \frac{2\pi^{d/2}}{\Gamma(d/2)}$ denotes the surface-area of the d-dimensional unit sphere. Further details about this estimator may be found in [49, Section 4.3.5] or [25] for instance. The following proposition gives the asymptotic normality of the pointwise estimator $\widehat{g}_n(t)$ for $t > 0$. Its proof is given in Section 4.5.

Proposition 4.3.4. *Assume that $\{D_n\}_{n\in\mathbb{N}}$ is a regular sequence of subsets of $\mathbb{R}^d$ such that $b_n^3|D_n| \to +\infty$ and $b_n^5|D_n| \to 0$. Let k be a symmetric and bounded function with compact support included in $[-T,T]$, for a given $T>0$, and $\int_{\mathbb{R}} k(x)dx = 1$. Let C be a twice differentiable kernel on $\mathbb{R}^d \setminus \{0\}$ such that it verifies the condition $\mathcal{K}(\rho)$ for a given $\rho > 0$ and $DPP(C)$ is isotropic. Denote g the pcf of $DPP(C)$ and define for all $n \in \mathbb{N}$, $\widehat{g}_n$ as in* (4.3.5)*. Then, for all $t > 0$, we have the convergence*

$$\sqrt{b_n|D_n|}\left(\widehat{g}_n(t) - g(t)\right) \xrightarrow[n\to+\infty]{distr.} N(0,\tau^2)$$

where $\tau^2 = 2\frac{g(t)}{\rho^2\sigma_d t^{d-1}}\sqrt{\int_{\mathbb{R}} k^2(x)dx}$.

In addition to the previous result, we state the asymptotic normality of the integrated squared error of the estimator $\widehat{\rho}_n^2\widehat{g}_n$ where $\widehat{g}_n$ is defined in (4.3.5). This is the basis of an asymptotic goodness-of-fit test for stationary DPPs as presented in [26]. For all segment $I \subset \mathbb{R}^+ \setminus \{0\}$ and $n \in \mathbb{N}$, denote $\mathrm{ISE}_n(I) = \int_I \left(\widehat{\rho}_n^2\widehat{g}_n(t) - \rho^2 g(t)\right)^2 dt$.

Proposition 4.3.5. *Assume that $b_n \to 0$ and $\{D_n\}_{n\in\mathbb{N}}$ is a regular sequence of subsets of $\mathbb{R}^d$ such that $b_n|D_n| \to +\infty$. Let k be a symmetric and bounded function with compact support included in $[-T,T]$, for a given $T>0$, and $\int_{\mathbb{R}} k(x)dx = 1$. Let C be a twice differentiable kernel on $\mathbb{R}^d \setminus \{0\}$ such that it verifies the condition $\mathcal{K}(\rho)$ for a given $\rho>0$ and $DPP(C)$ is isotropic. Denote g the pcf of $DPP(C)$ and define for all $n \in \mathbb{N}$, $\widehat{g}_n$ as in* (4.3.5)*. Then, for all segment $I \subset \mathbb{R}^+ \setminus \{0\}$, we have as n tends to infinity,*

$$b_n|D_n|\mathbb{E}\left(\mathrm{ISE}_n(I)\right) = 2\rho^2\int_I \frac{g(t)}{\sigma_d t^{d-1}}dt\int_{\mathbb{R}} k(x)^2dx + O\left(b_n\right) + O(|D_n|b_n^5).$$

Proposition 4.3.6. *Assume that in addition to the assumptions of Proposition 4.3.5, $b_n^5|D_n|$ tends to 0 as n tends to infinity, then we have for all segment $I \subset \mathbb{R}^+ \setminus \{0\}$ the convergence*

$$\sqrt{b_n}|D_n|\left(\mathrm{ISE}_n(I) - \mathbb{E}\left(\mathrm{ISE}_n(I)\right)\right) \xrightarrow[n\to+\infty]{distr.} N(0,\tau^2)$$

where $\tau^2 = 8\rho^4\int_I\left(\frac{g(t)}{\sigma_d t^{d-1}}\right)^2 dt\int_{\mathbb{R}}(k*k)^2(s)ds.$

Proposition 4.3.5 and 4.3.6 are direct applications in the DPP's case of the results given in [26]. In addition to the Brillinger mixing and the properties of the sequences $\{D_n\}_{n\in\mathbb{N}}$, the authors need two additional assumptions. Namely, these assumptions are the locally uniform Lipschitz continuity of the first derivative of the pcf and a second assumption related to the densities of the reduced factorial cumulant measure. As C is twice differentiable on $\mathbb{R}^d \setminus \{0\}$, the first derivative of the pcf is uniformly Lipschitz continuous on every compact sets in $\mathbb{R}^+ \setminus \{0\}$ so the first assumption holds. The second assumption is verified by Lemma 4.3.7 below. Consequently, Propositions 4.3.5 and 4.3.6 are proved by [26, Lemma 3.4] and [26, Theorem 3.5].

Lemma 4.3.7. *Let be an isotropic DPP with kernel C verifying the condition $\mathcal{K}(\rho)$ and densities of the reduced factorial cumulant moment measure of order 3 and 4, $c_{[3]}^{red}$ and $c_{[4]}^{red}$, respectively. For all compact set $K \subset \mathbb{R}^d$ and $\epsilon > 0$, we have*

$$\sup_{\substack{(u,v)\in\mathbb{R}^{2d}\\(u,v)\in(K^{\oplus\epsilon})^2}}\left|c_{[3]}^{red}(u,v)\right| < +\infty \tag{4.3.6}$$

and

$$\sup_{\substack{(u,v)\in\mathbb{R}^{2d}\\(u,v)\in(K^{\oplus\epsilon})^2}} \int_{\mathbb{R}^d} \left| c_{[4]}^{red}(u,w,v+w) \right| dw < +\infty, \tag{4.3.7}$$

where $K^{\oplus\epsilon} := K + B(0,\epsilon)$ *and* $B(0,\epsilon)$ *is the Euclidean ball centred at* 0 *with radius* ϵ *.*

Proof. By (4.6.2)-(4.6.3) in Section 4.6, we have for all $(u,v,w) \in \mathbb{R}^{3d}$,

$$c_{[3]}^{red}(u,v) = 2C(u)C(v)C(v-u)$$

and

$$\begin{aligned} c_{[4]}^{red}(u,v,w) = -2\,&[C(u)C(v)C(u-w)C(v-w) \\ &+C(u)C(w)C(u-v)C(v-w) + C(v)C(w)C(u-v)C(u-w)]\,. \end{aligned}$$

Notice that $K^{\oplus\epsilon}$ is compact and since C verifies the condition $\mathcal{K}(\rho)$, it is continuous. Therefore, by (4.6.2), (4.3.6) holds immediately. Finally, (4.3.7) is verified by Cauchy-Schwarz inequality and (4.6.3). □

4.4 Proofs of Theorem 4.3.2

4.4.1 Complement on the moments and cumulants of a point process

We present here the necessary background to prove Theorem 4.3.2. Let p and k be two integers and $\mathbf{X}$ a point process that admits moments of any order. Consider, for $1 \leq i \leq k$, the random variables

$$N_p(\phi_i) = \sum_{(x_1,\ldots,x_p)\in\mathbf{X}^p} \phi_i(x_1,\ldots,x_p) \tag{4.4.1}$$

where for $i = 1,\ldots,p$, ϕ_i is a function from $\mathbb{R}^{dp}$ to $\mathbb{R}$.

For $l, s \leq kp$, denote $\mathcal{P}_l^{kp}$ (resp. $\mathcal{Q}_s^l$) the set of all partitions of $\{1,\ldots,kp\}$ (resp. $\{1,\ldots,l\}$) into l (resp. s) non empty sets $p_1,\ldots,p_l$ (resp. $q_1,\ldots,q_s$). For $r = 1,\ldots,s$, denote $\beta_1,\ldots,\beta_{|q_r|}$ the elements of the set q_r and $|q|$ the cardinal of a given set q. Then, as proved by Jolivet in [35, p121-122], we have

$$E\left(N_p(\phi_1)\ldots N_p(\phi_k)\right) = \sum_{l=1}^{kp} \sum_{\Pi_l\in\mathcal{P}_l^{kp}} \sum_{s=1}^{l} \sum_{\chi_s^l\in\mathcal{Q}_s^l} I_l(\Pi_l,\chi_s^l) \tag{4.4.2}$$

where for all $l, s \leq kp$,

$$\begin{aligned} I_l(\Pi_l,\chi_s^l) = \int_{\mathbb{R}^{dl}} \prod_{m=1}^{l} \prod_{j\in p_m} \mathbf{1}_{\{x_m=\theta_j\}} \times \ldots \\ \times \prod_{i=1}^{k} \phi_i(\theta_{(i-1)p+1},\ldots,\theta_{ip}) \prod_{r=1}^{s} \gamma_{|q_r|}(dx_{\beta_1}\ldots dx_{\beta_{|q_r|}}). \end{aligned} \tag{4.4.3}$$

The introduction of the term θ is not easy to understand at first sight. For the sake of clarity, we give an example for $p = k = 2$ and $\Pi_2 := \{p_1, p_2\}$ a given partition of the set $\{1,2,3,4\}$ into 2 non empty sets, namely $p_1 = \{1,4\}$ and $p_2 = \{2,3\}$. In this case, we have

$$\prod_{i=1}^{k} \phi_i(\theta_{(i-1)p+1}, \ldots, \theta_{ip}) = \phi_1(\theta_1, \theta_2)\phi_2(\theta_3, \theta_4).$$

Thus, by the last equation, we have

$$\prod_{m=1}^{2} \prod_{j \in p_m} \mathbf{1}(x_m = \theta_j) \prod_{i=1}^{k} \phi_i(\theta_{(i-1)p+1}, \ldots, \theta_{ip}) = \phi_1(x_1, x_2)\phi_2(x_2, x_1)$$

and a similar calculus is done if, for $l = 1, \ldots, 4$, we choose another partition Π_l of $\{1,2,3,4\}$. We can now describe completely $\mathrm{Cum}(N_p(\phi_1), \ldots, N_p(\phi_k))$.

Theorem 4.4.1 ([35]). *The cumulant moment* $\mathrm{Cum}(N_p(\phi_1), \ldots, N_p(\phi_k))$ *is equal to the sum of integrals* $I_l(\Pi_l, \chi_s^l)$ *in Formula* (4.4.2) *that are indecomposable, i.e. that can not be decomposed as a product of at least two integrals.*

4.4.2 Proof of Theorem 4.3.2

It follows immediately from Definition 4.1.1 that for all $n \in \mathbb{N}$ and $k \geq 1$,

$$\mathrm{Cum}_k\left(\sqrt{|D_n|} N_p(f_{D_n})\right) = |D_n|^{\frac{k}{2}} \mathrm{Cum}_k(N_p(f_{D_n})). \tag{4.4.4}$$

Further, by [55, Chapter II, Section 12, Equation (37)], for X and Y two independent random variables, we have $\mathrm{Cum}_k(X+Y) = \mathrm{Cum}_k(X) + \mathrm{Cum}_k(Y)$ and a straightforward calculation proves that the cumulant of order k of a constant is null for $k \geq 2$. Consequently, by (4.4.4), for all $k \geq 2$, we have

$$\mathrm{Cum}_k\left(\sqrt{|D_n|}\left[N_p(f_{D_n}) - \mathbb{E}(N_p(f_{D_n}))\right]\right) = |D_n|^{\frac{k}{2}} \mathrm{Cum}_k(N_p(f_{D_n})).$$

We prove that $\mathrm{Cum}_k(N_p(f_{D_n})) = O\left(|D_n|^{1-k}\right)$. In this case, (4.3.2) holds by the last equation. Thus, for all $n \in \mathbb{N}$, the cumulant of order 2 is finite and for all $k > 2$ we have

$$\mathrm{Cum}_k\left(\sqrt{|D_n|}\left[N_p(f_{D_n}) - \mathbb{E}(N_p(f_{D_n}))\right]\right) \xrightarrow[n \to +\infty]{} 0.$$

Further, if (4.3.3) holds, the convergence (4.3.4) is obtained by [34, Theorem 1]. By Theorem 4.4.1, for every $k \in \mathbb{N}$, $\mathrm{Cum}_k(N_p(f_{D_n}))$ is a finite sum of indecomposable integrals $I_l(\Pi_l, \chi_s^l)$ in Formula (4.4.2) where we add the exponent t to indicate that these integrals depend of t. Thus, it is sufficient to prove that for a given $k \geq 2$, each integral $\left|I_l(\Pi_l, \chi_s^l)\right| = O\left(|D_n|^{1-k}\right)$. By (4.4.3), we have

$$I_l(\Pi_l, \chi_s^l)$$
$$= \int_{\mathbb{R}^{dl}} \prod_{m=1}^{l} \prod_{j \in p_m} \mathbf{1}_{\{x_m = \theta_j\}} \prod_{i=1}^{k} f_{D_n}(\theta_{(i-1)p+1}, \theta_{(i-1)p+2}, \ldots, \theta_{ip}) \prod_{r=1}^{s} \gamma_{q_r}(dx_{\beta_1} \ldots dx_{\beta_{|q_r|}}). \tag{4.4.5}$$

Then, by Definition 4.1.7, we obtain from the last equation that

$$|I_l(\Pi_l, \chi_s^l)| \leq \int_{\mathbb{R}^{dl}} \prod_{m=1}^{l} \prod_{j \in p_m} \mathbf{1}_{\{x_m = \theta_j\}} \prod_{i=1}^{k} \left| f_{D_n}(\theta_{(i-1)p+1}, \theta_{(i-1)p+2}, \dots, \theta_{ip}) \right| \prod_{r=1}^{s} |\gamma_{q_r}| \, (dx_{\beta_1} \dots dx_{\beta_{|q_r|}}).$$

Thus, by (4.3.1), we have

$$|I_l(\Pi_l, \chi_s^l)| \leq \frac{1}{|\widetilde{D}_n|^k} \int_{\mathbb{R}^{dl}} \prod_{m=1}^{l} \prod_{j \in p_m} \mathbf{1}_{\{x_m = \theta_j\}} \prod_{i=1}^{k} \mathbf{1}_{D_n}(\theta_{(i-1)p+1}) F\left(\theta_{(i-1)p+2} - \theta_{(i-1)p+1}, \dots, \theta_{ip} - \theta_{(i-1)p+1}\right) \prod_{r=1}^{s} |\gamma_{q_r}| \, (dx_{\beta_1} \dots dx_{\beta_{|q_r|}}). \quad (4.4.6)$$

Let $||F||_\infty$ denotes the supremum of F on $\mathbb{R}^{d(p-1)}$. Since the function F is bounded and compactly supported, there exist compacts $K_1, \dots, K_{p-1}$ such that

$$\forall (x_1, \dots, x_{p-1}) \in \mathbb{R}^{d(p-1)}, \quad F(x_1, \dots, x_{p-1}) \leq ||F||_\infty \mathbf{1}_{\{x_1 \in K_1\}} \cdots \mathbf{1}_{\{x_{p-1} \in K_{p-1}\}}.$$

Then, we deduce from (4.4.6) that

$$|I_l(\Pi_l, \chi_s^l)| \leq \left(\frac{||F||_\infty}{|\widetilde{D}_n|} \right)^k \int_{\mathbb{R}^{dl}} \prod_{m=1}^{l} \prod_{j \in p_m} \mathbf{1}_{\{x_m = \theta_j\}} \prod_{i=1}^{k} \mathbf{1}_{\{\theta_{(i-1)p+1} \in D_n\}} \prod_{\eta=1}^{p-1} \mathbf{1}_{\{\theta_{(i-1)p+\eta+1} - \theta_{(i-1)p+1} \in K_\eta\}} \prod_{r=1}^{s} |\gamma_{q_r}| (dx_{\beta_1} \dots dx_{\beta_{|q_r|}}). \quad (4.4.7)$$

Moreover, as already proved in [35, Section 4, Theorem 3], we have as n tends to infinity,

$$\int_{\mathbb{R}^{dl}} \prod_{m=1}^{l} \prod_{j \in p_m} \mathbf{1}_{\{x_m = \theta_j\}} \prod_{i=1}^{k} \mathbf{1}_{\{\theta_{(i-1)p+1} \in D_n\}} \prod_{\eta=1}^{p-1} \mathbf{1}_{\{\theta_{(i-1)p+\eta+1} - \theta_{(i-1)p+1} \in K_\eta\}} \prod_{r=1}^{s} |\gamma_{q_r}| (dx_{\beta_1} \dots dx_{\beta_{|q_r|}}) = O\left(|D_n|\right). \quad (4.4.8)$$

Since $\frac{|\widetilde{D}_n|}{|D_n|} \xrightarrow[n \to +\infty]{} \kappa$, the right hand term of (4.4.7) is, by (4.4.8), asymptotically of order $|D_n|^{1-k}$. Therefore

$$\left| I_l(\Pi_l, \chi_s^l) \right| = O\left(|D_n|^{1-k} \right).$$

4.5 Proof of Proposition 4.3.4

The proof is based on the following lemmas.

Lemma 4.5.1. *Assume that $b_n \to 0$ and $\{D_n\}_{n\in\mathbb{N}}$ is a regular sequence of subsets of $\mathbb{R}^d$ such that $b_n^3|D_n| \to +\infty$. Let k be a symmetric and bounded function with compact support included in $[-T,T]$, for a given $T>0$, and $\int_{\mathbb{R}} k(x)dx = 1$. Let C be a twice differentiable kernel on $\mathbb{R}^d \setminus \{0\}$ such that C verifies the condition $\mathcal{K}(\rho)$ for a given $\rho>0$ and $DPP(C)$ is isotropic. Denote g the pcf of $DPP(C)$ and define for all $n\in\mathbb{N}$, $\widehat{g}_n$ as in* (4.3.5)*. Then, for all $t>0$, we have the convergence*

$$\sqrt{b_n|D_n|}\left(\widehat{\rho}_n^2\widehat{g}_n(t) - \mathbb{E}(\widehat{\rho}_n^2\widehat{g}_n(t))\right) \xrightarrow[n\to+\infty]{distr.} N(0,\kappa^2)$$

where $\kappa^2 = 2\rho^2\frac{g(t)}{\sigma_d t^{d-1}}\sqrt{\int_{\mathbb{R}} k^2(x)dx}$.

Lemma 4.5.2. *Under the same assumptions as in Proposition 4.3.5, for all segment $I\subset\mathbb{R}^+\setminus\{0\}$, there exists a constant $M\geq 0$ such that*

$$\sup_{t\in I}\left|\mathbb{E}\left(\widehat{\rho}_n^2\widehat{g}_n(t) - \rho^2 g(t)\right)\right| \leq b_n^2 M\rho^2\int_{\mathbb{R}} u^2|k(u)|du.$$

The proofs of Lemmas 4.5.1-4.5.2 are postponed to the end of the section. Let us now prove Proposition 4.3.4. For all $n\in\mathbb{N}$ and $t>0$, we have

$$\widehat{\rho}_n^2\sqrt{b_n|D_n|}(\widehat{g}_n(t) - g(t)) = A_n + B_n + C_n \tag{4.5.1}$$

where

$$\begin{aligned} A_n &= \sqrt{b_n|D_n|}\left[\widehat{\rho}_n^2\widehat{g}_n(t) - \mathbb{E}(\widehat{\rho}_n^2\widehat{g}_n(t))\right] \\ B_n &= \sqrt{b_n|D_n|}\left[\mathbb{E}(\widehat{\rho}_n^2\widehat{g}_n(t)) - \rho^2 g(t)\right] \\ C_n &= \sqrt{b_n|D_n|}g(t)\left[\rho^2 - \widehat{\rho}_n^2\right]. \end{aligned}$$

By Lemma 4.5.1 we have the convergences

$$A_n \xrightarrow[n\to+\infty]{distr.} N(0,\kappa^2) \tag{4.5.2}$$

and since $b_n^5|D_n|$ tends to 0 as n tends to infinity, we have by Lemma 4.5.2,

$$B_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0. \tag{4.5.3}$$

By Corollary 4.3.3 and the delta method, we know that $\sqrt{|D_n|}(\hat{\rho}_n^2 - \rho^2)$ converges in distribution. Since $b_n\to 0$, we deduce that

$$C_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0. \tag{4.5.4}$$

Finally, by inserting (4.5.2)-(4.5.4) in (4.5.1), the proposition is proved by Slutsky theorem and by the almost sure convergence of $\widehat{\rho}_n^2$ to ρ^2.

4.5.1 Proof of Lemma 4.5.1

We need the following lemma.

Lemma 4.5.3. *Let $r > 0$ and D a subset of $\mathbb{R}^d$ such that the Euclidean ball $B(0,r)$ is included in D. Then, for all $x \in B(0,r)$, we have $D^{\ominus r} \subset D \cap D^x$.*

Proof of Lemma 4.5.3. By definition, $D \cap D^x = \{u \in D,\ u + x \in D\}$ and

$$D^{\ominus r} = \{u \in D,\ \forall v \in B(0,r),\ u + v \in D\}.$$

As $x \in B(0,r)$, we have the inclusion $D^{\ominus r} \subset D \cap D^x$. □

Define for all $n \in \mathbb{N}$ and $(x_1, x_2) \in \mathbb{R}^{2d}$,

$$f_{D_n}(x_1, x_2) = \mathbf{1}_{\{x_1 \in D_n,\ x_2 \in D_n\}} \frac{1}{|D_n \cap D_n^{x_1 - x_2}|} k\left(\frac{t - |x_1 - x_2|}{b_n}\right).$$

Notice by (4.3.5) that

$$b_n \sigma_d t^{d-1} \widehat{\rho}_n^2 \widehat{g}_n(t) = \sum_{\substack{(x_1,x_2) \in \mathbf{X}^2 \\ x_1 \neq x_2}} f_{D_n}(x_1, x_2). \tag{4.5.5}$$

The support of k is included in $[-T, T]$ so for any $(x_1, x_2) \in \mathbb{R}^{2d}$,

$$\begin{aligned}
\left| k\left(\frac{t - |x_1 - x_2|}{b_n}\right) \right| \mathbf{1}_{\{x_2 \in D_n\}} &\leq \left| k\left(\frac{t - |x_1 - x_2|}{b_n}\right) \right| \mathbf{1}_{\{|x_1 - x_2| < t + Tb_n\}} \\
&\leq \left| k\left(\frac{t - |x_1 - x_2|}{b_n}\right) \right| \mathbf{1}_{\{|x_1 - x_2| < t + T\}}
\end{aligned}$$

as soon as $b_n < 1$ which we assume without loss of generality since $\{b_n\}_{n \in \mathbb{N}}$ tends to 0. Then, by Lemma 4.5.3 and since k is bounded, there exists $M > 0$ such that for all $n \in \mathbb{N}$,

$$|f_{D_n}(x_1, x_2)| \leq \frac{M \mathbf{1}_{\{x_1 \in D_n\}}}{|D_n^{\ominus t+T}|} \mathbf{1}_{\{|x_1 - x_2| \leq t + T\}}.$$

Thus, by (4.3.2) in Theorem 4.3.2 and (4.5.5), we have for all $k \geq 3$,

$$\mathrm{Cum}_k\left(\sqrt{|D_n|} b_n \sigma_d t^{d-1} \widehat{\rho}_n^2 \widehat{g}_n(t)\right) = O(|D_n|^{1 - \frac{k}{2}}).$$

Hence, for all $k \geq 3$,

$$\mathrm{Cum}_k\left(\sqrt{b_n |D_n|} \widehat{\rho}_n^2 \widehat{g}_n(t)\right) = O\left(\frac{1}{b_n^{k/2} |D_n|^{\frac{k}{2} - 1}}\right)$$

which tends to 0 when n goes to infinity since $b_n^3 |D_n| \to \infty$. Further, the convergences of $\mathrm{Cum}_k\big(\sqrt{b_n |D_n|} \widehat{\rho}_n^2 \widehat{g}_n(t)\big)$ for $k = 1, 2$ are proved in [27] under conditions that we have already verified after Proposition 4.3.6. Finally, Lemma 4.5.1 is proved by [34, Theorem 1].

4.5.2 Proof of Lemma 4.5.2

By (4.3.5), we have for all $n \in \mathbb{N}$ and $t \in I$,

$$\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t)\right) = \frac{\rho^2}{\sigma_d t^{d-1} b_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\mathbf{1}_{\{x \in D_n, y \in D_n\}}}{\left|D_n \cap D_n^{x-y}\right|} k\left(\frac{t-|x-y|}{b_n}\right) \frac{\rho^{(2)}(x,y)}{\rho^2} dx dy$$

where $\rho^{(2)}$ is the second order factorial product densities of $DPP(C)$. Since we have assumed that $DPP(C)$ is stationary and isotropic, we have

$$\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t)\right) = \frac{\rho^2}{\sigma_d t^{d-1} b_n} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\mathbf{1}_{\{x \in D_n, y \in D_n\}}}{\left|D_n \cap D_n^{x-y}\right|} k\left(\frac{t-|x-y|}{b_n}\right) g(|x-y|) dx dy.$$

To shorten, denote $k_{b_n}(.) = \frac{1}{b_n} k\left(\frac{.}{b_n}\right)$. By the substitution $z = x - y$ and as $y \in D_n^z \Leftrightarrow z \in D_n^y$, we obtain from the last equation that

$$\begin{aligned}
\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t)\right) &= \frac{\rho^2}{\sigma_d t^{d-1}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \frac{\mathbf{1}_{\{y \in D_n^z \cap D_n\}}}{|D_n \cap D_n^z|} k_{b_n}(t-|z|) g(|z|) dz dy \\
&= \frac{\rho^2}{\sigma_d t^{d-1}} \int_{\mathbb{R}^d} k_{b_n}(t-|z|) g(|z|) dz.
\end{aligned}$$

Then, as k is symmetric, we have

$$\begin{aligned}
\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t)\right) &= \rho^2 \int_0^{+\infty} \left(\frac{r}{t}\right)^{d-1} k_{b_n}(t-r) g(r) dr \\
&= \rho^2 \int_{-\frac{t}{b_n}}^{+\infty} k(u) \left(\frac{t+ub_n}{t}\right)^{d-1} g(t+ub_n) du.
\end{aligned}$$

For all n large enough, we have for all $t \in I$, $\frac{t}{b_n} \geq T$ where T is such that $[-T, T]$ contains the support of k. Hence, for all n large enough, we have

$$\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t)\right) = \rho^2 \int_{-T}^{T} k(u) \left(\frac{t+ub_n}{t}\right)^{d-1} g(t+ub_n) du.$$

Assume that I writes $[t_{min}, t_{max}]$ for $t_{max} > t_{min} > 0$ and define for $s \in \mathbb{R}^+$, $f(s) := \left(\frac{s}{t}\right)^{d-1} g(s)$. Notice that $I^{\oplus T b_n} := [t_{min} - Tb_n, t_{max} + Tb_n] \subset \mathbb{R}^+ \setminus \{0\}$ as soon as n is large enough which we assume without loss of generality. Then $g(.)$ is of class $\mathcal{C}^2$ on $I^{\oplus T b_n}$, so is $f(.)$. Thus, by Taylor Lagrange expansion, we have

$$\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t)\right) = \rho^2 \int_{-T}^{T} k(u) \left(f(t) + f'(t) u b_n + \int_t^{t+ub_n} f''(s)(ub_n + t - s) ds\right) du. \tag{4.5.6}$$

As k is symmetric, we have $\int_{-T}^{T} uk(u)du = 0$. Moreover,

$$\sup_{s \in I^{\oplus T b_n}} \left|f''(s)\right| \leq \frac{1}{t_{min}^{d-1}} \sup_{s \in I^{\oplus T b_n}} \left|\left(s^{d-1} g(s)\right)''\right|.$$

Since $(.)^{d-1} g(.)$ is of class $\mathcal{C}^2$ on $I^{\oplus T b_n}$, $f''(.)$ is, by the last inequality, uniformly bounded on $I^{\oplus T b_n}$ by a constant M . Further, for all $n \in \mathbb{N}$ and $s \in [t, t+ub_n]$, $|ub_n + t - s| \leq |ub_n|$. Finally, since $\int_{-T}^{T} k(u)du = 1$, by (4.5.6), we have for n large enough,

$$\left|\mathbb{E}\left(\widehat{\rho}_n^2 \widehat{g}_n(t) - \rho^2 g(t)\right)\right| \leq b_n^2 M \rho^2 \int_{\mathbb{R}} u^2 |k(u)| du, \quad \forall t \in [t_{min}, t_{max}].$$

4.6 Appendix

We gather here some results useful to compute the asymptotic variance in Theorem 4.3.2 and Corollary 4.3.3.

For $\mathbf{X}$ a stationary point process on $\mathbb{R}^d$, we denote $c_{[2]}^{red}$, $c_{[3]}^{red}$ and $c_{[4]}^{red}$ the densities of its reduced factorial cumulant moment measures of order 2, 3 and 4, respectively, assuming they exist. If $\mathbf{X}$ is a DPP with kernel C verifying the condition $\mathcal{K}(\rho)$, for a given $\rho > 0$, then we deduce from Definitions 3.2.2 and 4.1.4 that for all $(u, v, w) \in \mathbb{R}^{3d}$,

$$c_{[2]}^{red}(u) = - C^2(u) \tag{4.6.1}$$

$$c_{[3]}^{red}(u, v) = 2\, C(u)C(v)C(v - u) \tag{4.6.2}$$

$$\begin{aligned} c_{[4]}^{red}(u, v, w) = & - 2\big[C(u)C(v)C(w - u)C(w - v) \\ & + C(u)C(w)C(v - u)C(v - w) + C(v)C(w)C(u - v)C(u - w)\big] \end{aligned} \tag{4.6.3}$$

Lemma 4.6.1. *Let f be a function from $\mathbb{R}^d$ into $\mathbb{R}$ that is bounded, measurable and compactly supported. Then we have*

$$\mathrm{Var}\left(\sum_{x\in\mathbf{X}} f(x)\right) = \int_{\mathbb{R}^{2d}} f(x)f(x+y)c_{[2]}^{red}(y)dxdy + \rho \int_{\mathbb{R}^d} f^2(x)dx.$$

Proof. Notice that

$$\left(\sum_{x\in\mathbf{X}} f(x)\right)^2 = \sum_{(x,y)\in\mathbf{X}^2}^{\neq} f(x)f(y) + \sum_{x\in\mathbf{X}} f^2(x).$$

Then, denoting $\rho^{(2)}$ the density of the second order factorial moment measure, we have by Definitions 4.1.3 and 4.1.4,

$$\begin{aligned} \mathrm{Var}\left(\sum_{x\in\mathbf{X}} f(x)\right) &= \int_{\mathbb{R}^{2d}} f(x)f(y)\left(\rho^{(2)}(x,y) - \rho^2\right) dxdy + \rho \int_{\mathbb{R}^d} f^2(x)dx \\ &= \int_{\mathbb{R}^{2d}} f(x)f(y)c_{[2]}(x,y)dxdy + \rho \int_{\mathbb{R}^d} f^2(x)dx. \end{aligned}$$

Finally, by Definition 4.1.5, we have

$$\mathrm{Var}\left(\sum_{x\in\mathbf{X}} f(x)\right) = \int_{\mathbb{R}^{2d}} f(x)f(x+y)c_{[2]}^{red}(y)dxdy + \rho \int_{\mathbb{R}^d} f^2(x)dx.$$

□

Lemma 4.6.2. *Let f be a function from $\mathbb{R}^{2d}$ into $\mathbb{R}$ that is bounded, measurable and compactly* *supported. Then, we have*

$$
\begin{aligned}
&\operatorname{Var}\left(\sum_{(x,y)\in\mathbf{X}^2}^{\neq} f(x,y)\right)\\
&= \int_{\mathbb{R}^{2d}} \left(f^2(x,x+y)+f(x,x+y)f(x+y,x)\right) c_{[2]}^{red}(y)dxdy\\
&+ \rho^2 \int_{\mathbb{R}^{2d}} \left(f^2(x,y)+f(x,y)f(y,x)\right) dxdy\\
&+ \int_{\mathbb{R}^{3d}} \left[f(x,x+y)+f(x+y,x)\right]\left[f(x+y,x+u)+f(x+u,x+y)\right] c_{[3]}^{red}(y,u)dxdydu\\
&+ 2\rho \int_{\mathbb{R}^{3d}} \left[f(x,y)+f(y,x)\right]\left[f(y,y+u)+f(y+u,y)\right] c_{[2]}^{red}(u)dxdydu\\
&+ \rho \int_{\mathbb{R}^{3d}} \left[f(x,y)+f(y,x)\right]\left[f(y,x+u)+f(x+u,y)\right] c_{[2]}^{red}(u)dxdydu\\
&+ \rho^3 \int_{\mathbb{R}^{3d}} \left[f(x,y)+f(y,x)\right]\left[f(y,u)+f(u,y)\right] dxdydu\\
&+ \int_{\mathbb{R}^{4d}} f(x,x+y)f(x+u,x+v)c_{[4]}^{red}(y,u,v)dxdydudv\\
&+ 4\rho \int_{\mathbb{R}^{4d}} f(x,y)f(y+u,y+v)c_{[3]}^{red}(u,v)dxdydudv\\
&+ 2 \int_{\mathbb{R}^{4d}} f(x,y)f(x+u,y+v)c_{[2]}^{red}(u)c_{[2]}^{red}(v)dxdydudv\\
&+ 4\rho^2 \int_{\mathbb{R}^{4d}} f(x,y)f(x+u,v)c_{[2]}^{red}(u)dxdydudv.
\end{aligned}
$$

Proof. This lemma is a generalization of [29, Lemma 5] for a function f non necessary symmetric. The variance is first computed with respect to the factorial moment measure by Definition 4.1.3. Then, the factorial moment measure is written in terms of the factorial cumulant moment measure by [9, Corollary 5.2 VII] and the result is found by using Definition 4.1.5. We refer to the proof of [29, Lemma 5] for the detailed calculus, the only change being the use of the following decomposition in place of the original one,

$$
\begin{aligned}
\left(\sum_{(x,y)\in\mathbf{X}^2}^{\neq} f(x,y)\right)^2 &= \sum_{(x,y)\in\mathbf{X}^2}^{\neq} f^2(x,y)+f(x,y)f(y,x)\\
&+ \sum_{(x,y,u)\in\mathbf{X}^3}^{\neq} \left[f(x,y)+f(y,x)\right]\left[f(y,u)+f(u,y)\right]\\
&+ \sum_{(x,y,u,v)\in\mathbf{X}^4}^{\neq} f(x,y)f(u,v).
\end{aligned}
$$

□

Lemma 4.6.3. *Let f be a function from $\mathbb{R}^{2d}$ into $\mathbb{R}$ that is bounded, measurable and compactly supported. Let h be a function from $\mathbb{R}^d$ into $\mathbb{R}$ that is bounded, measurable and compactly supported.*

Then, we have

$$
\begin{aligned}
\operatorname{Cov}\left(\sum_{(x,y)\in \mathbf{X}^2}^{\neq} f(x,y), \sum_{u\in\mathbf{X}} h(u) \right) &= \int_{\mathbb{R}^{3d}} f(x,x+y)h(u+x)c_{[3]}^{red}(y,u)dxdydu \\
&+ \rho \int_{\mathbb{R}^{3d}} f(x,y)h(u+x)c_{[2]}^{red}(u)dxdydu \\
&+ \rho \int_{\mathbb{R}^{3d}} f(x,y)h(u+y)c_{[2]}^{red}(u)dxdydu \\
&+ \int_{\mathbb{R}^{2d}} f(x,y+x)\left[h(x)+h(y+x)\right] c_{[2]}^{red}(y)dxdy \\
&+ \rho^2 \int_{\mathbb{R}^{2d}} f(x,y)\left[h(x)+h(y)\right]dxdy.
\end{aligned}
$$

Proof. Notice that

$$
\left(\sum_{(x,y)\in \mathbf{X}^2}^{\neq} f(x,y) \right)\left(\sum_{u\in\mathbf{X}} h(u) \right) = \sum_{(x,y,u)\in \mathbf{X}^3}^{\neq} f(x,y)h(u) + \sum_{(x,y)\in \mathbf{X}^2}^{\neq} f(x,y)(h(x)+h(y)).
$$

Then, by the last equation and Definition 4.1.3, we have

$$
\begin{aligned}
\operatorname{Cov}\left(\sum_{(x,y)\in \mathbf{X}^2}^{\neq} f(x,y), \sum_{u\in\mathbf{X}} h(u) \right) &= \iiint f(x,y)h(u)\left(\rho^{(3)}(x,y,u)-\rho\rho^{(2)}(x,y)\right)dxdydu \\
&+ \iint f(x,y)(h(x)+h(y))\rho^{(2)}(x,y)dxdy.
\end{aligned}
$$

Finally, the proof is concluded by [9, Corollary 5.2 VII] and Definition 4.1.5. □

5

Asymptotic properties of minimum contrast estimators

In this chapter, we study minimum contrast estimation for DPPs, based on the Ripley's K-function and the pcf g. This method is already widely used for inference of some Cox processes. It is implemented for these models in the `spatstat` library of R, see [4], and asymptotic results have been obtained in the stationary and non-stationary cases, see [23] and [66]. For DPPs, the methodology is similar, as detailed in Section 5.1, and routines are also available in the `spatstat` library but no theoretical results were available so far. We prove in Section 5.2 the consistency and asymptotic normality of the minimum contrast estimator based on K and g for stationary DPPs. In particular we obtain a closed form expression for the asymptotic variance. These results are in fact derived from a more general theorem, dealing with asymptotic properties of minimum contrast estimation of a stationary point process, that we state and prove in Section 5.3, generalizing a result in [23].

5.1 Framework

We consider a parametric family of DPPs with kernel $C_{\rho,\theta}$ where $\rho = C_{\rho,\theta}(0) > 0$ and θ belongs to a subset Θ_ρ of $\mathbb{R}^p$, for a given $p \geq 1$. To ensure the existence of $DPP(C_{\rho,\theta})$, we assume that for all $\rho > 0$ and any $\theta \in \Theta_\rho$, the kernel $C_{\rho,\theta}$ verifies $\mathcal{K}(\rho)$, which explains the indexation of Θ_ρ by ρ. We assume further that for a given $\rho_0 > 0$ and θ_0 in the interior of Θ_{ρ_0} (provided this interior is non-empty) we observe a point process $\mathbf{X} \sim DPP(C_{\rho_0,\theta_0})$ on a regular sequence $\{D_n\}_{n\in\mathbb{N}}$ of subsets of $\mathbb{R}^d$ in the sense of Definition 4.3.1. Subsequently, we are interested in the asymptotic properties of the estimators of ρ_0 and θ_0 defined as follows.

For all $n \in \mathbb{N}$, ρ_0 is estimated on the set D_n by

$$\widehat{\rho}_n = \frac{1}{|D_n|} \sum_{x\in\mathbf{X}} \mathbf{1}_{\{x\in D_n\}}. \tag{5.1.1}$$

By the ergodic theorem, this estimator is strongly consistent, see for instance [25], and by Corollary 4.3.3, it is asymptotically normal. Hence, in the following, we focus our attention on θ_0. As explained in [43], likelihood inference is in theory feasible if we know a spectral representation of $C_{\rho,\theta}$ on D_n. Unfortunately no spectral representations are known in the general case and some approximations are necessary, see [43]. To the best of our knowledge, the asymptotic properties of likelihood inference for DPPs is still an open question, even when a spectral representation is assumed to be known. In this chapter we consider another option to estimate θ_0, namely a minimum contrast method. This procedure, described below, does not require to know any spectral representation of the kernel.

For $\rho > 0$ and $\theta \in \Theta_\rho$, let $J(.,\theta)$ be a function from $\mathbb{R}^d$ into $\mathbb{R}^+$ which is a summary statistic of $DPP(C_{\rho,\theta})$ that does not depend on ρ. In the DPP's case, the most important and natural examples are the K-function and the pcf g. Then, for all $n \in \mathbb{N}$, consider $\widehat{J}_n$ an estimator of J from the observation of $\mathbf{X}$ on D_n. Further, let $c \in \mathbb{R}$ be a parameter such that $\widehat{J}_n(t)^c$ and $J(t,\theta)^c$ are well defined for all $n \in \mathbb{N}$, $t \in \mathbb{R}$ and $\theta \in \Theta_{\rho_0}$. Finally, define for $r_{max} > r_{min} \geq 0$ and $c \neq 0$, the discrepancy measure

$$U_n(\theta) = \int_{r_{min}}^{r_{max}} w(t) \left\{ \widehat{J}_n(t)^c - J(t,\theta)^c \right\}^2 dt \tag{5.1.2}$$

where w is a weight function. We present the necessary assumptions on w in the following but in most cases, w is chosen to be smooth on (r_{min}, r_{max}) or at least integrable. The case $c = 0$ is not of interest and is not considered. We denote

$$\widehat{\theta}_n = \underset{\theta \in \Theta_{\hat{\rho}_n}}{\arg\min}\, U_n(\theta) \tag{5.1.3}$$

and study the consistency and asymptotic normality of the estimator $\widehat{\theta}_n$ of θ_0. For all $\rho > 0$ and $\theta \in \Theta_\rho$, define the function R_θ as $R_\theta = C_{\rho,\theta}/\rho$ that we assume does not depend on ρ but only on θ. Note that this is the case for all parametric families considered in Section 3.5 as well as for the Whittle-Matèrn and the Generalized Cauchy families, see [44]. For $r > 0$, we denote by $B(0,r)$ (resp. $\overline{B}(0,r)$) the open (resp. closed) Euclidean ball centred at 0 with radius r and by $\Theta_{\rho_0}^{\oplus r} := \Theta_{\rho_0} + \overline{B}(0,r)$ the r-dilation of Θ_{ρ_0}. Further, for all $x \in \mathbb{R}^d$, denote $R_\theta^{(i)}(x)$, the i-th derivative with respect to θ. We make the following assumptions.

($\mathcal{H}1$) For all $\rho > 0$, Θ_ρ is a compact convex set with non-empty interior and the mapping $\rho \to \Theta_\rho$ is continuous with respect to the Haussdorff distance on the compact sets. Further $0 \leq r_{min} < r_{max}$, $c \neq 0$ and $\{D_n\}_{n\in\mathbb{N}}$ is a sequence of regular subsets of $\mathbb{R}^d$ in the sense of Definition 4.3.1.

($\mathcal{H}2$) For all $\theta \in \Theta_{\rho_0}$, $C_{\rho_0,\theta}$ verifies the condition $\mathcal{K}(\rho_0)$ and there exists $\epsilon > 0$ such that $C_{\rho_0,\theta} \in L^2(\mathbb{R}^d)$ and $\mathcal{F}(C_{\rho_0,\theta}) \geq 0$ for all $\theta \in \Theta_{\rho_0}^{\oplus\epsilon}$.

($\mathcal{H}3$) There exists $\epsilon > 0$ such that for all $x \in B(0, r_{max})$, the function $\theta \mapsto R_\theta(x)$ is of class $\mathcal{C}^2$ on $\Theta_{\rho_0}^{\oplus\epsilon}$. Further, for $i \in \{1,2\}$, there exists $M > 0$ such that for all $x \in B(0, r_{max})$ and $\theta \in \Theta_{\rho_0}^{\oplus\epsilon}$, $\left|R_\theta^{(i)}(x)\right| \leq M$.

The first assumption is needed to handle the fact that the minimisation (5.1.3) is done over the random set $\Theta_{\hat{\rho}_n}$ in place of Θ_{ρ_0}. These three assumptions are verified by most usual parametric families of DPPs.

5.2 Minimum contrast estimation based on the K and g functions

In this section, we present the minimum contrast method introduced in Section 5.1 in its two most common versions. In Section 5.2.1 the statistic J in (5.1.2) corresponds to the K-function and in Section 5.2.2, J is the pcf g. The proofs of this section are postponed to Sections 5.4-5.5.

5.2.1 The Ripley's K-function

Definition 5.2.1. *The Ripley's K-function of a stationary point process with pair correlation function g is defined for all $t \geq 0$ by*

$$K(t) = \int_{B(0,t)} g(x)dx.$$

Note that for a given $\rho > 0$ and $\theta \in \Theta_\rho$, the K-function of $DPP(C_{\rho,\theta})$ depends only on R_θ. In particular, it does not depend on ρ. Consequently, we denote for all $\theta \in \Theta_{\rho_0}$, $K(.,\theta)$ the K-function of $DPP(C_{\rho_0,\theta})$. For a subset D_n of $\mathbb{R}^d$, we define for any $t \geq 0$, $D_n^{\ominus t} := \{x,\ B(x,t) \in D_n\}$. We consider the estimator of the K-function defined for all $t \geq 0$ and $n \in \mathbb{N}$ by

$$\widehat{K}_n(t) := \frac{1}{\widehat{\rho}_n^2} \sum_{(x,y)\in\mathbf{X}^2}^{\neq} \mathbf{1}_{\{x\in D_n\}} \mathbf{1}_{\{y\in D_n^{\ominus t}\}} \frac{\mathbf{1}_{\{|x-y|\leq t\}}}{|D_n^{\ominus t}|} \tag{5.2.1}$$

where $\widehat{\rho}_n$ is as in (5.1.1). For all $t \in [r_{min}, r_{max}]$, denote $K^{(i)}(t,\theta)$ the i-th derivative with respect to θ. We consider the following assumptions.

($\mathcal{H}_K1$) w is a positive and integrable function in $[r_{min}, r_{max}]$.

($\mathcal{H}_K2$) If $r_{min} = 0$, then $c \geq 2$.

($\mathcal{H}_K3$) For $\theta_1 \neq \theta_2$, there exists a set $A \in [r_{min}, r_{max}]$ of positive Lebesgue measure such that

$$\int_{x\in B(0,t)} R_{\theta_1}(x)^2 dx \neq \int_{x\in B(0,t)} R_{\theta_2}(x)^2 dx, \quad \forall t \in A.$$

($\mathcal{H}_K4$) The matrix $\int_{r_{min}}^{r_{max}} w(t)K(t,\theta_0)^{2c-2}K^{(1)}(t,\theta_0)K^{(1)}(t,\theta_0)^T dt$ is invertible.

Assumption ($\mathcal{H}_K1$) is not restrictive. The constraint on c implied by ($\mathcal{H}_K2$) in the case $r_{min} = 0$ tends to confirm the practice, which consists in the choice $r_{min} > 0$. ($\mathcal{H}_K3$) is an identifiability assumption and ($\mathcal{H}_K4$) turns out to be the main assumption. Define for all $t \in [r_{min}, r_{max}]$,

$$j_K(t) := w(t)K(t,\theta_0)^{2c-2}K^{(1)}(t,\theta_0).$$

The following theorem is proved in Section 5.4.

Theorem 5.2.2. *Let $\mathbf{X}$ be a DPP with kernel $C_{\rho_0,\theta_0} = \rho_0 R_{\theta_0}$ for a given $\rho_0 > 0$ and θ_0 an interior point of Θ_{ρ_0}. For all $n \in \mathbb{N}$, let U_n be defined as in* (5.1.2) *with $J = K$ and $\widehat{J}_n = \widehat{K}_n$.*

Assume that ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K 1$)-($\mathcal{H}_K 4$) hold. Then, the minimum contrast estimator $\widehat{\theta}_n$ defined by (5.1.3) *exists and is strongly consistent for θ_0. Moreover, it satisfies*

$$\sqrt{|D_n|}(\widehat{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} \mathcal{N}\left[0, B_{\theta_0}^{-1}\Sigma_{\rho_0,\theta_0}\{B_{\theta_0}^{-1}\}^T\right]$$

with

$$B_{\theta_0} := \int_{r_{min}}^{r_{max}} w(t)K(t,\theta_0)^{2c-2}K^{(1)}(t,\theta_0)K^{(1)}(t,\theta_0)^T dt \tag{5.2.2}$$

and

$$\Sigma_{\rho_0,\theta_0} = \int_{r_{min}}^{r_{max}}\int_{r_{min}}^{r_{max}} h(t_1,t_2)j_K(t_1)j_K(t_2)dt_1dt_2$$

where h can be expressed in terms of C_{ρ_0,θ_0}. Specifically, for all $(t_1,t_2)\in[r_{min},r_{max}]^2$,

$$\begin{aligned}
h(t_1,t_2) := {}& 2\int_{\mathbb{R}^d} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|x|\le t_2\}}\left(c_{[2]}^{red}(x)+\rho_0^2\right)dx\\
&+4\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|y-x|\le t_2\}}\left(c_{[3]}^{red}(x,y)+\rho_0 c_{[2]}^{red}(y)\right)dxdy\\
&+4\rho_0\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|y|\le t_2\}}\left(2c_{[2]}^{red}(y)+\rho_0^2\right)dxdy\\
&+\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|z-y|\le t_2\}}c_{[4]}^{red}(x,y,z)dxdydz\\
&+4\rho_0\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|z-y|\le t_2\}}c_{[3]}^{red}(y,z)dxdydz\\
&+2\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|x+z-y|\le t_2\}}c_{[2]}^{red}(y)c_{[2]}^{red}(z)dxdydz\\
&+4\rho_0^2\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{0<|x|\le t_1\}}\mathbf{1}_{\{0<|z-y|\le t_2\}}c_{[2]}^{red}(y)dxdydz\\
&-4\rho_0\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{0<|x|\le t_1\}}K(t_2,\theta_0)\left(c_{[3]}^{red}(x,y)+2\rho_0 c_{[2]}^{red}(y)\right)dxdy\\
&-8\rho_0\int_{\mathbb{R}^{d}} \mathbf{1}_{\{0<|x|\le t_1\}}K(t_2,\theta_0)\left(c_{[2]}^{red}(x)+\rho_0^2\right)dx\\
&+4\rho_0^2K(t_1,\theta_0)K(t_2,\theta_0)\left(\rho_0-\int_{\mathbb{R}^d}C_{\rho_0,\theta_0}(x)^2dx\right)
\end{aligned}$$

where $c_{[2]}^{red}$, $c_{[3]}^{red}$ and $c_{[4]}^{red}$ are given with respect to C_{ρ_0,θ_0} in (4.6.1)-(4.6.3) *in Section 4.6.*

Notice that the finiteness of the integrals involved in the last expression follows from the Brillinger mixing property of the DPPs with kernel verifying the condition $\mathcal{K}(\rho_0)$, see Theorem 4.2.2.

5.2.2 The pair correlation function

We assume in this section that all DPPs of the parametric family are isotropic. In this case, for all $\rho > 0$ and $\theta \in \Theta_\rho$, there exists $\widetilde{R}_\theta$ such that $R_\theta(x) = \widetilde{R}_\theta(|x|)$ for all $x \in \mathbb{R}^d$ so the pcf of $DPP(C_{\rho,\theta})$ writes, for any $x\in\mathbb{R}^d$,

$$g(x,\theta) = 1 - \widetilde{R}_\theta(|x|)^2 \tag{5.2.3}$$

and does not depend on ρ. Hence, there exists a function $\tilde{g}$ such that for all $(x,\theta) \in \mathbb{R}^d \times \Theta_\rho$, $g(x,\theta) = \tilde{g}(|x|,\theta)$. In the following, to alleviate the notation, we omit the symbol tilde and for all $\theta \in \Theta_\rho$, we view $R_\theta(.)$ as a function from $\mathbb{R}^+$ into $\mathbb{R}$ and $g(.,\theta)$ as a function from $\mathbb{R}^+$ into $\mathbb{R}^+$. Moreover, for all $\theta \in \Theta_\rho$, we extend by symmetry the domain of definition of $R_\theta(.)$ on $\mathbb{R}$. Denote, for all $d \geq 2$, the surface area of the d-dimensional unit ball,

$$\sigma_d := \frac{2\pi^{d/2}}{\Gamma(d/2)}.$$

The pcf is usually estimated via kernel estimation method. For $n \in \mathbb{N}$ and $t \in \mathbb{R}^+ \setminus \{0\}$, we consider the estimator, see [49, Section 4.3.5],

$$\widehat{g_n}(t) := \frac{1}{\sigma_d t^{d-1} \widehat{\rho}_n^2} \sum_{(x,y)\in\mathbf{X}^2}^{\neq} \mathbf{1}_{\{x \in D_n,\, y \in D_n\}} \frac{1}{b_n |D_n \cap D_n^{x-y}|} k\left(\frac{t-|x-y|}{b_n}\right) \tag{5.2.4}$$

where $D_n^z := \{u, u+z \in D_n\}$ for any $z \in \mathbb{R}^d$, $\widehat{\rho}_n$ is as in (5.1.1) and b_n and k are the bandwidth and the kernel to be chosen according to the assumptions below. For all $t \in [r_{min}, r_{max}]$, denote by $g^{(i)}(t,\theta)$ the i-th derivative of g with respect to θ. We consider the assumptions:

($\mathcal{H}_g1$) $r_{min} > 0$.

($\mathcal{H}_g2$) The kernel k is positive, symmetric and bounded with compact support included in $[-T,T]$ for a given $T > 0$. Further, $\int_{\mathbb{R}} k(x)dx = 1$.

($\mathcal{H}_g3$) $\{b_n\}_{n\in\mathbb{N}}$ is a positive sequence, $b_n \to 0$, $b_n|D_n| \to +\infty$ and $b_n^4|D_n| \to 0$.

($\mathcal{H}_g4$) There exists $\epsilon > 0$ such that for all $\theta \in \Theta_{\rho_0}^{\oplus\epsilon}$, $R_\theta(.)$ is of class $\mathcal{C}^2$ on $\mathbb{R} \setminus \{0\}$.

($\mathcal{H}_g5$) w is a positive and continuous function on $[r_{min}, r_{max}]$.

($\mathcal{H}_g6$) For $\theta_1 \neq \theta_2$, there exists a set $A \in [r_{min}, r_{max}]$ of positive Lebesgue measure such that

$$|R_{\theta_1}(t)| \neq |R_{\theta_2}(t)|, \quad \forall t \in A.$$

($\mathcal{H}_g7$) The matrix $\int_{r_{min}}^{r_{max}} w(t) g(t,\theta_0)^{2c-2} g^{(1)}(t,\theta_0) g^{(1)}(t,\theta_0)^T dt$ is invertible.

The first four assumptions are easy to satisfy by appropriate choices of r_{min}, b_n and k. ($\mathcal{H}_g6$) is an identifiability assumption and as in the previous section, the main assumption is in fact ($\mathcal{H}_g7$). The proof of the following theorem is postponed to Section 5.5. Let

$$j_g(t) := w(t) g(t,\theta_0)^{2c-2} g^{(1)}(t,\theta_0), \quad t \in [r_{min}, r_{max}].$$

Theorem 5.2.3. *Let $\mathbf{X}$ be an isotropic DPP with kernel $C_{\rho_0,\theta_0} = \rho_0 R_{\theta_0}$ for a given $\rho_0 > 0$ and θ_0 an interior point of Θ_{ρ_0}. For all $n \in \mathbb{N}$, let U_n be defined as in* (5.1.2) *with $J = g$ and $\widehat{J}_n = \widehat{g}_n$. Assume that ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_g1$)-($\mathcal{H}_g7$) hold. Assume further that for all $\theta \in \Theta_{\rho_0}$, $R_\theta(.)$ is isotropic. Then, the minimum contrast estimator $\widehat{\theta}_n$ defined by* (5.1.3) *exists and is consistent for θ_0. Moreover, it satisfies*

$$\sqrt{|D_n|}(\widehat{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} \mathcal{N}\left[0, B_{\theta_0}^{-1} \Sigma_{\rho_0,\theta_0} \{B_{\theta_0}^{-1}\}^T\right]$$

with

$$B_{\theta_0} := \int_{r_{min}}^{r_{max}} w(t) g(t,\theta_0)^{2c-2} g^{(1)}(t,\theta_0) g^{(1)}(t,\theta_0)^T dt$$

and

$$\begin{aligned}
\Sigma_{\rho_0,\theta_0} &= 2\int_{\mathbb{R}^d} \mathbf{1}_{\{r_{min}\le|x|\le r_{max}\}} \frac{j_g(|x|)j_g(|x|)}{\sigma_d^2|x|^{2(d-1)}} \left(c_{[2]}^{red}(x)+\rho_0^2\right) dx \\
&+ 4\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{r_{min}\le|x|,|y-x|\le r_{max}\}} \frac{j_g(|x|)j_g(|y-x|)}{\sigma_d^2|x|^{d-1}|y-x|^{d-1}} \left(c_{[3]}^{red}(x,y)+\rho_0 c_{[2]}^{red}(y)\right) dxdy \\
&+ 4\rho_0\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{r_{min}\le|x|,|y|\le r_{max}\}} \frac{j_g(|x|)j_g(|y|)}{\sigma_d^2|x|^{d-1}|y|^{d-1}} \left(2c_{[2]}^{red}(x)+\rho_0^2\right) dxdy \\
&+ \int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y|)}{\sigma_d^2|x|^{d-1}|z-y|^{d-1}} c_{[4]}^{red}(x,y,z)\, dxdydz \\
&+ 4\rho_0\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y|)}{\sigma_d^2|x|^{d-1}|z-y|^{d-1}} c_{[3]}^{red}(y,z)\, dxdydz \\
&+ 2\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y+x|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y+x|)}{\sigma_d^2|x|^{d-1}|z-y+x|^{d-1}} c_{[2]}^{red}(y)c_{[2]}^{red}(z)\, dxdydz \\
&+ 4\rho_0^2\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y|)}{\sigma_d^2|x|^{d-1}|z-y|^{d-1}} c_{[2]}^{red}(y)\, dxdydz \\
&- 4\rho_0\left(\int_{r_{min}}^{r_{max}} g(t,\theta_0)j_g(t)dt\right) \\
&\qquad\qquad \int_{\mathbb{R}^{2d}} \mathbf{1}_{\{r_{min}\le|x|\le r_{max}\}} \frac{j_g(|x|)}{\sigma_d|x|^{d-1}} \left(c_{[3]}^{red}(x,y)+2\rho_0 c_{[2]}^{red}(y)\right) dxdy \\
&- 8\rho_0\left(\int_{r_{min}}^{r_{max}} g(t,\theta_0)j_g(t)dt\right) \int_{\mathbb{R}^d} \mathbf{1}_{\{r_{min}\le|x|\le r_{max}\}} \frac{j_g(|x|)}{\sigma_d|x|^{d-1}} \left(c_{[2]}^{red}(x)+\rho_0^2\right) dx \\
&+ 4\rho_0^2\left(\int_{r_{min}}^{r_{max}} g(t,\theta_0)j_g(t)dt\right)^2 \left(\rho_0 - \int_{\mathbb{R}^d} C_{\rho_0,\theta_0}(x)^2 dx\right)
\end{aligned}$$

where $c_{[2]}^{red}$, $c_{[3]}^{red}$ and $c_{[4]}^{red}$ are given with respect to C_{ρ_0,θ_0} in (4.6.1)-(4.6.3) *in Section 4.6.*

5.3 A general result for minimum contrast estimation method of stationary point processes

We present in this section two theorems about consistency and asymptotic normality of the estimator defined by (5.1.3). Contrary to the results in Section 5.2.1-5.2.2, these theorems hold for an arbitrary stationary point process and an arbitrary statistic J, generalizing a study by Sherman and Guan in [23].

5.3.1 Statements

Let $\mathbf{X}$ be a stationary point process belonging to a parametric family indexed by, among possibly other parameters, $\theta \in \Theta$ where $\Theta \subset \mathbb{R}^p$, for a given $p \ge 1$. For any $t \in [r_{min}, r_{max}]$, let $J(t,\theta)$ be any real valued summary statistic of $\mathbf{X}$ that depends on θ (specific assumptions on J are listed below). For any $t \in [r_{min}, r_{max}]$, let $\widehat{J}_n(t)$ be an estimator of $J(t,\theta_0)$ where θ_0 is the true parameter ruling the distribution of $\mathbf{X}$.

We denote by $J^{(i)}(t,\theta)$, the i-th derivative with respect to θ. Define for all $\theta \in \Theta$,

$$B(\theta) := \int_{r_{min}}^{r_{max}} w(t) J(t,\theta)^{2c-2} J^{(1)}(t,\theta) J^{(1)}(t,\theta)^T dt, \tag{5.3.1}$$

and for all $t \in [r_{min}, r_{max}]$,

$$j(t) = w(t) J(t,\theta_0)^{2c-2} J^{(1)}(t,\theta_0).$$

We consider the following assumptions.

($\mathcal{A}1$) Θ is a compact set with non-empty interior, $0 \leq r_{min} < r_{max}$, $c \neq 0$ and $\{D_n\}_{n\in\mathbb{N}}$ is a regular sequence of subsets of $\mathbb{R}^d$ in the sense of Definition 4.3.1.

($\mathcal{A}2$) w is a positive and integrable function in $[r_{min}, r_{max}]$.

($\mathcal{A}3$) $J(.,.)$ and $J(.,.)^c$ are well defined continuous functions on $[r_{min}, r_{max}] \times \Theta$. Moreover, there exists a set $A \in [r_{min}, r_{max}]$ such that $[r_{min}, r_{max}] \setminus A$ is of Lebesgue measure null and for all $t \in A$, $\theta \in \Theta$, we have $J(t,\theta) > 0$.

($\mathcal{A}4$) For all $n \in \mathbb{N}$, $\widehat{J}_n(.)$ is almost surely bounded on $[r_{min}, r_{max}]$ and so is $\widehat{J}_n(.)^c$ for all n large enough.

($\mathcal{A}5$) There exists a set $A \in [r_{min}, r_{max}]$ such that $[r_{min}, r_{max}] \setminus A$ is of Lebesgue measure null and

$$\sup_{t\in A} \left| \widehat{J}_n(t) - J(t,\theta_0) \right| \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

($\mathcal{A}6$) For $\theta_1 \neq \theta_2$, there exists a set A of positive Lebesgue measure such that

$$J(t,\theta_1) \neq J(t,\theta_2), \quad \forall t \in A.$$

($\mathcal{A}7$) For all $t \in [r_{min}, r_{max}]$, $J^{(1)}(t,\theta)$ and $J^{(2)}(t,\theta)$ exist, are continuous with respect to θ and uniformly bounded with respect to $t \in [r_{min}, r_{max}]$ and $\theta \in \Theta$.

($\mathcal{A}8$) There exists $M > 0$ such that for all $(t,\theta) \in [r_{min}, r_{max}] \times \Theta$ and $a \in \{c-2, 2c-2\}$, $\left|J(t,\theta)\right|^a \leq M$.

($\mathcal{A}9$) The matrix $B(\theta_0)$ is invertible.

($\mathcal{TCL}$) There exists $m \in \mathbb{R}$ and a covariance matrix Σ such that

$$\sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\widehat{J}_n(t) - J(t,\theta_0) \right] j(t) dt \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(m, \Sigma).$$

Further, define $(\mathcal{A}5)'$ as the assumption $(\mathcal{A}5)$ with the convergence in probability replaced by the almost sure convergence.

Theorem 5.3.1. *Let $\mathbf{X}$ be a stationary point process with distribution ruled by a given θ_0, an interior point of Θ. For all $n \in \mathbb{N}$, let U_n be defined as in* (5.1.2). *Assume that($\mathcal{A}1$)-($\mathcal{A}6$) hold. Then, the minimum contrast estimator $\widetilde{\theta}_n$ defined by*

$$\widetilde{\theta}_n = \arg\min_{\theta\in\Theta} U_n(\theta)$$

exists almost surely, is consistent for θ_0 and strongly consistent if $(\mathcal{A}5)'$ holds.

Theorem 5.3.2. *Let $\mathbf{X}$ be a stationary point process with distribution ruled by a given θ_0, an interior point of Θ. For all $n \in \mathbb{N}$, let U_n be defined as in* (5.1.2). *Assume that ($\mathcal{A}$1)-($\mathcal{A}$9) and $(\mathcal{TCL})$ hold. Then, the minimum contrast estimator $\widetilde{\theta}_n$ defined as in Theorem 5.3.1 verifies*

$$\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} \mathcal{N}\left(m, B(\theta_0)^{-1}\Sigma\left(B(\theta_0)^{-1}\right)^T\right)$$

where B is defined as in (5.3.1) *and Σ comes from $(\mathcal{TCL})$.*

5.3.2 Proof of Theorem 5.3.1

For a sequence $\{\theta_m\}_{m\in\mathbb{N}}$ belonging to Θ, we have for all $n \in \mathbb{N}$,

$$|U_n(\theta_m) - U_n(\theta)| \leq \int_{r_{min}}^{r_{max}} |w(t)| \left(\left|2\widehat{J}_n(t)^c\right| |J(t,\theta_m)^c - J(t,\theta)^c| + \left|J(t,\theta_m)^{2c} - J(t,\theta)^{2c}\right|\right) dt. \quad (5.3.2)$$

Denote A the intersection of the sets defined in ($\mathcal{A}$3) and ($\mathcal{A}$5). By ($\mathcal{A}$3), $J(.,.)^c$ is continuous on $[r_{min}, r_{max}] \times \Theta$ which is compact by ($\mathcal{A}$1). We deduce that

$$\sup_{t\in[r_{min},r_{max}]} |J(t,\theta_m)^c - J(t,\theta)^c| \leq K.$$

By ($\mathcal{A}$3)-($\mathcal{A}$4), for all $\theta \in \Theta$, $J(t,\theta)^c$ and $\widehat{J}_n(t)^c$ are almost surely bounded on $[r_{min}, r_{max}]$, for all n large enough. Further, by ($\mathcal{A}$2), w is integrable on $[r_{min}, r_{max}]$ thus, by (5.3.2) and the dominated convergence theorem, we have the convergence

$$|U_n(\theta_m) - U_n(\theta)| \xrightarrow[\theta_m\to\theta]{a.s.} 0.$$

Therefore, for all $n \in \mathbb{N}$, U_n is almost surely continuous so the almost sure existence of $\widetilde{\theta}_n$ follows by ($\mathcal{A}$1). Define for all $\theta \in \Theta$,

$$U_n^*(\theta) = U_n(\theta) - U_n(\theta_0). \quad (5.3.3)$$

By (5.1.2) and (5.3.3),

$$U_n^*(\theta) = 2\int_{r_{min}}^{r_{max}} w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_0)^c\right]\left[J(t,\theta_0)^c - J(t,\theta)^c\right]dt + \int_{r_{min}}^{r_{max}} w(t)\left[J(t,\theta_0)^c - J(t,\theta)^c\right]^2 dt.$$

Note that from (5.3.3) $U_n^*(\widetilde{\theta}_n) \leq U_n^*(\theta_0) = 0$, so

$$\int_{r_{min}}^{r_{max}} w(t)\left[J(t,\theta_0)^c - J(t,\widetilde{\theta}_n)^c\right]^2 dt \leq 2\int_{r_{min}}^{r_{max}} w(t)\left|\widehat{J}_n(t)^c - J(t,\theta_0)^c\right|\left|J(t,\theta_0)^c - J(t,\widetilde{\theta}_n)^c\right| dt. \quad (5.3.4)$$

By ($\mathcal{A}3$)-($\mathcal{A}4$), $J(.,.)^c$ is continuous on $[r_{min}, r_{max}] \times \Theta$ and for n large enough, $\widehat{J}_n(.)^c$ is almost surely bounded on $[r_{min}, r_{max}]$ so by ($\mathcal{A}5$), the right-hand term in (5.3.4) tends in probability to 0. Hence, we have

$$\int_{r_{min}}^{r_{max}} w(t)\left[J(t,\theta_0)^c - J(t,\widetilde{\theta}_n)^c\right]^2 dt \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

It follows by ($\mathcal{A}2$) and ($\mathcal{A}6$) that $\widetilde{\theta}_n$ converges in probability to θ_0. Finally, by a similar argument, we prove by (5.3.4) that this last convergence is almost sure if $(\mathcal{A}5)'$ holds.

5.3.3 Proof of Theorem 5.3.2

Denote A the intersection of the sets defined in ($\mathcal{A}3$) and ($\mathcal{A}5$). Then, by ($\mathcal{A}3$), ($\mathcal{A}7$) and ($\mathcal{A}8$), we see that U_n is almost surely twice differentiable on Θ and that we can differentiate twice under the integral sign. Thus, by the mean value theorem, for all $j = 1, \ldots, p$, there exists $s \in (0,1)$ such that $c_j = \theta_0 + s(\widetilde{\theta}_n - \theta_0)$ and

$$\partial_j U_n(\widetilde{\theta}_n) - \partial_j U_n(\theta_0) = \left(\partial^2_{ij} U_n(c_j)\right)_{i=1,\ldots,p} \left(\widetilde{\theta}_n - \theta_0\right).$$

To shorten, denote $\partial_\theta U_n$ the gradient of U_n and $\partial^2_{\theta\theta^T} U_n(\theta_n^*)$ the Hessian of U_n at θ_n^* with θ_n^* taking the value c_j on the line $j = 1, \ldots, p$. We have then

$$\partial_\theta U_n(\widetilde{\theta}_n) - \partial_\theta U_n(\theta_0) = \partial^2_{\theta\theta^T} U_n(\theta_n^*)(\widetilde{\theta}_n - \theta_0), \tag{5.3.5}$$

$$\partial^2_{\theta\theta^T} U_n(\theta_n^*) = 2c^2 B(\theta_n^*) - E_n \tag{5.3.6}$$

where B is as in (5.3.1) and

$$\begin{aligned} E_n := 2c(c-1) \int_{r_{min}}^{r_{max}} & w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_n^*)^c\right] J(t,\theta_n^*)^{c-2} J^{(1)}(t,\theta_n^*) J^{(1)}(t,\theta_n^*)^T dt \\ &+ 2c \int_{r_{min}}^{r_{max}} w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_n^*)^c\right] J(t,\theta_n^*)^{c-1} J^{(2)}(t,\theta_n^*) dt. \end{aligned}$$

Since U_n is minimal at $\widetilde{\theta}_n$, $\partial_\theta U_n(\widetilde{\theta}_n) = 0$ and by (5.3.5), we have

$$\partial^2_{\theta\theta^T} U_n(\theta_n^*)(\widetilde{\theta}_n - \theta_0) = 2c \int_{r_{min}}^{r_{max}} w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_0)^c\right] J(t,\theta_0)^{c-1} J^{(1)}(t,\theta_0) dt \tag{5.3.7}$$

Note that by ($\mathcal{A}3$) and ($\mathcal{A}8$), $J(.,\theta_0)^{c-1}$ is bounded on A and strictly positive. Thus, by ($\mathcal{A}4$), we can use the Taylor expansion of the function $x \mapsto x^c$ so, for all $t \in A$,

$$\widehat{J}_n(t)^c - J(t,\theta_0)^c = cJ(t,\theta_0)^{c-1}\left(\widehat{J}_n(t) - J(t,\theta_0)\right) + o\left(\widehat{J}_n(t) - J(t,\theta_0)\right).$$

Then, by ($\mathcal{A}5$), (5.3.7) and the last equation,

$$\sqrt{|D_n|}\,\partial^2_{\theta\theta^T} U_n(\theta_n^*)(\widetilde{\theta}_n - \theta_0) = 2c^2 A_n(\theta_0) + o(A_n(\theta_0)) \tag{5.3.8}$$

where

$$A_n(\theta_0) = \sqrt{|D_n|} \int_A \left[\widehat{J}_n(t) - J(t, \theta_0)\right] j(t)dt.$$

By $(\mathcal{TCL})$, we have $2c^2 A_n(\theta_0) \xrightarrow[n\to+\infty]{distr.} 2c^2 N(m, \Sigma)$. Hence, by Slutsky's theorem and (5.3.8), we have

$$\sqrt{|D_n|}\partial^2_{\theta\theta^T} U_n(\theta_n^*)(\widetilde{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} 2c^2 N(m, \Sigma). \tag{5.3.9}$$

Moreover, by Theorem 5.3.1, $\widetilde{\theta}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \theta_0$ so, $\theta_n^* \xrightarrow[n\to+\infty]{\mathbb{P}} \theta_0$. Then, by $(\mathcal{A}7)$-$(\mathcal{A}8)$, E_n tends in probability to 0. Note that by continuity of $J(., \theta)$ for all $\theta \in \Theta$, the integrability on $[r_{min}, r_{max}]$ of $w(.)J(., \theta)^{c-2}$ implies the one of $w(.)J(., \theta)^{c-1}$. Further, we deduce by $(\mathcal{A}3)$, $(\mathcal{A}7)$ and $(\mathcal{A}8)$ that $(t, \theta) \mapsto J(t, \theta)^{2c-2} J^{(1)}(t, \theta) J^{(1)}(t, \theta)^T$ is continuous with respect to $\theta \in \Theta$ and uniformly bounded for $t \in [r_{min}, r_{max}]$. Thus, by the dominated convergence theorem,

$$B(\theta_n^*) \xrightarrow[n\to+\infty]{\mathbb{P}} B(\theta_0).$$

By $(\mathcal{A}9)$, $B(\theta_0)$ is invertible so by (5.3.6) and Slutsky's theorem,

$$\partial^2_{\theta\theta^T} U_n(\theta_n^*) B(\theta_0)^{-1} \xrightarrow[n\to+\infty]{\mathbb{P}} 2c^2. \tag{5.3.10}$$

Since the group of invertible matrix is an open set, it follows from the last convergence that, for n large enough, $\partial^2_{\theta\theta^T} U_n(\theta_n^*)$ is invertible so we have

$$B(\theta_0)\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0) = B(\theta_0)\left[\partial^2_{\theta\theta^T} U_n(\theta_n^*)\right]^{-1} \partial^2_{\theta\theta^T} U_n(\theta_n^*)\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0).$$

Hence, by (5.3.9)-(5.3.10) and Slutsky's theorem,

$$B(\theta_0)\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} N(m, \Sigma).$$

Finally, it follows from the last equation that

$$\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} \mathcal{N}\left(m, B(\theta_0)^{-1}\Sigma\left(B(\theta_0)^{-1}\right)^T\right).$$

5.4 Proof of Theorem 5.2.2

Since C_{ρ_0,θ_0} verifies $\mathcal{K}(\rho_0)$, $\widehat{\rho}_n$ converges almost surely to ρ_0, so by $(\mathcal{H}1)$, for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $n \geq N$, $\Theta_{\hat{\rho}_n} \subset \Theta_{\rho_0}^{\oplus\epsilon}$ almost surely. Henceforth, without loss of generality, we let $\epsilon > 0$ and assume that $\Theta_{\hat{\rho}_n} \subset \Theta_{\rho_0}^{\oplus\epsilon}$ for all $n \in \mathbb{N}$. We prove below the strong consistency and asymptotic normality of the estimator $\widetilde{\theta}_n$ defined in Theorems 5.3.1 and 5.3.2 with $\Theta = \Theta_{\rho_0}^{\oplus\epsilon}$, $J = K$ and $\widehat{J}_n = \widehat{K}_n$. As a consequence, almost surely, there exist $r > 0$ such that $B(\theta_0, r) \subset \Theta_{\rho_0}$ and $N_r \in \mathbb{N}$ such that for all $n \geq N_r$, $\widetilde{\theta}_n \in B(\theta_0, r)$. The next lemma shows that for n sufficiently large, $B(\theta_0, r) \subset \Theta_{\hat{\rho}_n}$. Hence, almost surely, for n large enough, the minimum of U_n is attained in $\Theta_{\hat{\rho}_n} \subset \Theta_{\rho_0}^{\oplus\epsilon}$ so $\widetilde{\theta}_n = \widehat{\theta}_n$ and the proof is completed.

Lemma 5.4.1. *For $p \geq 1$, let Θ be a convex compact set in $\mathbb{R}^p$ and $\{\Theta_n\}_{n\in\mathbb{N}}$ be a sequence of convex compact sets in $\mathbb{R}^p$ that converges to Θ with respect to the Hausdorff distance. Let $r \geq 0$ and x be an interior point of Θ such that $B(x,r)$ belongs to the interior of Θ. Then, there exists $N \in \mathbb{N}$ such that for all $n \geq N$,*

$$B(x,r) \subset \Theta_n.$$

Let us now prove the strong consistency and asymptotic normality of $\widetilde{\theta}_n$. To that end, we verify all the assumptions of Theorems 5.3.1-5.3.2 using a series of lemmas stated below. The assumptions ($\mathcal{H}1$) and ($\mathcal{H}_K1$) imply directly ($\mathcal{A}1$)-($\mathcal{A}2$). For all $\theta \in \Theta$, we have

$$K(t,\theta) = \sigma_d t^d - \int_{x\in B(0,t)} R_\theta(x)^2 dx \tag{5.4.1}$$

where $\mathcal{F}(R_\theta) \geq 0$ by ($\mathcal{H}2$). Further, by [51, Corollary 1.4.13], for all $\theta \in \Theta$, if for a given $x \neq 0$, $|R_\theta(x)| = 1$, then R_θ is invariant by translation of x. Since for all $\theta \in \Theta$, $R_\theta(.) \in L^2(\mathbb{R}^d)$, this is impossible so, for all $x \neq 0$ and $\theta \in \Theta$, $|R_\theta(x)| < 1$. Hence, by (5.4.1), $K(t,\theta) > 0$ on $(r_{min}, r_{max}] \times \Theta$ and $K(.,.)$ is continuous on $[r_{min}, r_{max}] \times \Theta$. Consequently, $K(.,.)^c$ is continuous for all $c \in \mathbb{R}$ if $r_{min} > 0$ and for all $c > 0$ if $r_{min} = 0$. Therefore, under ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K2$), ($\mathcal{A}3$) holds. By the same arguments, $K(.,.)^{c-2}$ and $K(.,.)^{2c-2}$ are continuous for all $c \in \mathbb{R}$ if $r_{min} > 0$ and for all $c \geq 2$ if $r_{min} = 0$. Thus, by ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K2$), ($\mathcal{A}8$) holds. For all $t \in [r_{min}, r_{max}]$, $\widehat{K}_n(t)$ is bounded by $\widehat{K}_n(r_{max})$ and it follows from the ergodic theorem that $\widehat{K}_n(r_{max})$ is almost surely finite as soon as n and so D_n is large enough. Moreover, by Lemma 5.4.2, $\widehat{K}_n(t)$ is almost surely strictly positive for $t > 0$ and n large enough. Hence, under ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K2$), ($\mathcal{A}4$) holds. We have for all $\theta \in \Theta$ and $t \in (0, r_{max})$

$$K^{(1)}(t,\theta) = -\frac{\partial}{\partial\theta} \int_{x\in B(0,t)} R_\theta(x)^2 dx.$$

By ($\mathcal{H}3$), the function $(x,\theta) \mapsto R_\theta^{(1)}(x)$ is continuous with respect to θ and bounded for all $x \in B(0, r_{max})$ and $\theta \in \Theta$. Thus by the dominated convergence theorem,

$$K^{(1)}(t,\theta) = -2 \int_{x\in B(0,t)} R_\theta(x) R_\theta^{(1)}(x) dx. \tag{5.4.2}$$

We obtain similarly

$$K^{(2)}(t,\theta) = -2 \int_{x\in B(0,t)} \left(R_\theta^{(1)}(x) R_\theta^{(1)}(x)^T + R_\theta^{(2)}(x) R_\theta(x) \right) dx.$$

By ($\mathcal{H}3$), the terms inside the integral in the last equation are bounded uniformly with respect to $(x,\theta) \in B(0, r_{max}) \times \Theta$. Therefore, $K^{(1)}(t,\theta)$ and $K^{(2)}(t,\theta)$ are continuous with respect to θ and uniformly bounded with respect to $t \in [r_{min}, r_{max}]$ and $\theta \in \Theta$ so ($\mathcal{A}7$) holds. The assumption $(\mathcal{A}5)'$ is proved by Lemma 5.4.2 while ($\mathcal{A}6$) and ($\mathcal{A}9$) are directly implied by ($\mathcal{H}_K3$) and ($\mathcal{H}_K4$), respectively. Finally, $(\mathcal{TCL})$ follows from Lemma 5.4.5.

Lemma 5.4.2. *Let K be the Ripley's function of a DPP with kernel C verifying $\mathcal{K}(\rho_0)$ and $\widehat{K}_n$ the estimator given by* (5.2.1). *Then, for all $r_{max} > r_{min} \geq 0$,*

$$\sup_{t\in[r_{min}, r_{max}]} \left| \widehat{K}_n(t) - K(t) \right| \xrightarrow[n\to+\infty]{a.s.} 0,$$

Lemma 5.4.3. *If ($\mathcal{H}_K$1)-($\mathcal{H}_K$2) and ($\mathcal{H}$3) hold, then for all $r_{max} > r_{min} \geq 0$,*

$$\int_{r_{min}}^{r_{max}} |j_K(t)|\, dt < +\infty.$$

To shorten, define for all $n \in \mathbb{N}$ and $t \in [r_{min}, r_{max}]$,

$$H_n(t) := \widehat{\rho}_n^2 \widehat{K}_n(t) - 2\rho_0 K(t, \theta_0)\widehat{\rho}_n.$$

Lemma 5.4.4. *If ($\mathcal{H}$1)-($\mathcal{H}$3) and ($\mathcal{H}_K$1) hold, for all $s \in \mathbb{R}^d$, we have*

$$\lim_{n\to+\infty} |D_n| \mathrm{Var}\Big(\int_{r_{min}}^{r_{max}} H_n(t) s^T j_K(t) dt \Big) = \int_{[r_{min}, r_{max}]^2} h(t_1, t_2) s^T j_K(t_1) s^T j_K(t_2) dt_1 dt_2$$

where h is defined as in Theorem 5.2.2.

Lemma 5.4.5. *If ($\mathcal{H}$1)-($\mathcal{H}$3) and ($\mathcal{H}_K$1)-($\mathcal{H}_K$2) hold, then*

$$\sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \Big[\widehat{K}_n(t) - K(t, \theta_0) \Big] j_K(t) dt \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(0, \Sigma_{\rho_0, \theta_0})$$

where $\Sigma_{\rho_0, \theta_0}$ is defined as in Theorem 5.2.2.

5.4.1 Proof of Lemma 5.4.1

Since $B(x, r)$ belongs to the interior of Θ, there exists $\delta > 0$ such that

$$B(x, r + \delta) \subset \Theta. \tag{5.4.3}$$

Assume that the lemma is wrong, then for all $N \in \mathbb{N}$, there exists $n \geq N$ such that $B(x, r) \not\subseteq \Theta_n$. Denote y a point in $B(x, r)$ that does not belong to Θ_n. By Lemma 5.4.6 stated below, $B(y, \delta) \not\subseteq \Theta_n^{\oplus\delta}$. But by (5.4.3), $B(y, \delta) \subset \Theta$ so $\Theta \not\subseteq \Theta_n^{\oplus\delta}$ which contradicts the convergence of the sequence $\{\Theta_n\}_{n\in\mathbb{N}}$ to Θ.

Lemma 5.4.6. *For all $p \geq 1$, let Ξ be a compact convex set in $\mathbb{R}^p$. Then, for all $y \in \mathbb{R}^p \setminus \Xi$ and $\delta \geq 0$, $B(y, \delta) \not\subseteq \Xi^{\oplus\delta}$.*

Proof. Since Ξ is a closed convex set, the projection of y onto Ξ, denoted by $p_\Xi(y)$, is the unique element belonging to Ξ that, for all $u \in \Xi$, verifies

$$(y - p_\Xi(y)).(u - p_\Xi(y)) \leq 0. \tag{5.4.4}$$

For all $\delta \geq 0$, the line $(y, p_\Xi(y))$ intersects $\partial B(y, \delta)$ at two points, one inside the segment $[y, p_\Xi(y)]$ and the other, that we denote by v, outside the segment. Thus, there exists $t > 1$ such that $v = p_\Xi(y) + t(y - p_\Xi(y))$. Notice that for all $u \in \Xi$,

$$(v - p_\Xi(y)).(u - p_\Xi(y)) = t(y - p_\Xi(y)).(u - p_\Xi(y)).$$

Thus, as $t > 1$, we deduce from (5.4.4) and the last equation that $p_\Xi(y)$ is the projection of v onto Ξ. It follows that $d(v, \Xi) = d(y, \Xi) + \delta$ and as $y \notin \Xi$ and Ξ is closed, $d(v, \Xi) > \delta$. Therefore, $v \notin \Xi^{\oplus\delta}$ but $v \in \partial B(y, \delta)$ by construction so $B(y, \delta) \not\subseteq \Xi^{\oplus\delta}$. □

5.4.2 Proof of Lemma 5.4.2

Since a stationary DPP is ergodic by [56, Theorem 7], we have

$$\widehat{\rho}_n \xrightarrow[n\to+\infty]{a.s.} \rho_0 \tag{5.4.5}$$

and

$$\sup_{t\in[r_{min},r_{max}]} \left|\widehat{\rho}_n^2 \widehat{K}_n(t) - \rho_0^2 K(t)\right| \xrightarrow[n\to+\infty]{a.s.} 0, \tag{5.4.6}$$

see for instance [25, Section 4.2.2]. Further, as K is an increasing function, we have

$$\begin{aligned}\widehat{\rho}_n^2 & \sup_{t\in[r_{min},r_{max}]} \left|\widehat{K}_n(t) - K(t)\right| \\ & \leq \sup_{t\in[r_{min},r_{max}]} \left|\widehat{\rho}_n^2 \widehat{K}_n(t) - \rho_0^2 K(t)\right| + K(r_{max}) \sup_{t\in[r_{min},r_{max}]} \left|\widehat{\rho}_n^2 - \rho_0^2\right| .\end{aligned}$$

Hence, by (5.4.5)-(5.4.6) and the last equation, we have the convergence

$$\sup_{t\in[r_{min},r_{max}]} \left|\widehat{K}_n(t) - K(t)\right| \xrightarrow[n\to+\infty]{a.s.} 0.$$

5.4.3 Proof of Lemma 5.4.3

By (5.4.2), we have

$$\int_{r_{min}}^{r_{max}} |j_K(t)|\, dt = 2\int_{r_{min}}^{r_{max}} \left| w(t)K(t,\theta_0)^{2c-2} \int_{x\in B(0,t)} R_{\theta_0}(x) R_{\theta_0}^{(1)}(x) dx\right| dt. \tag{5.4.7}$$

By $(\mathcal{H}3)$, the function defined for all $t \in \mathbb{R}^+$ by

$$t \mapsto \int_{x\in B(0,t)} R_{\theta_0}(x) R_{\theta_0}^{(1)}(x) dx$$

is continuous so bounded on $[r_{min}, r_{max}]$ for $r_{max} > r_{min} \geq 0$. As already noticed below (5.4.1), $K(t,\theta) > 0$ on $(r_{min}, r_{max}] \times \Theta$. Consequently, if $r_{min} > 0$, the lemma is proved since w is integrable on $[r_{min}, r_{max}]$ by $(\mathcal{H}_K 1)$. Finally, if $r_{min} = 0$, the integrability at 0 of the function $t \mapsto |j_K(t)|$ follows from $(\mathcal{H}_K 2)$.

5.4.4 Proof of Lemma 5.4.4

By (5.2.1), we have

$$\int_{r_{min}}^{r_{max}} H_n(t) s^T j_K(t) dt = \sum_{(x,y)\in\mathbf{X}^2} f_n(x,y) - \sum_{x\in\mathbf{X}} h_n(x)$$

where for all $n \in \mathbb{N}$,

$$f_n(x,y) := \mathbf{1}_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} \frac{1}{|D_n^{\ominus t}|} \mathbf{1}_{\{y\in D_n^{\ominus t}\}} \mathbf{1}_{\{0<|x-y|\leq t\}} s^T j_K(t) dt.$$

and

$$h_n(x) = \frac{2\rho_0}{|D_n|}\mathbf{1}_{\{x\in D_n\}}\int_{r_{min}}^{r_{max}} K(t,\theta_0)s^T j_K(t)dt.$$

Notice that for all $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$, $f_n(x,x) = 0$. Thus, we have from the last equation,

$$\begin{aligned}
&\mathrm{Var}\left(\int_{r_{min}}^{r_{max}} H_n(t)s^T j_K(t)dt\right)\\
&= \mathrm{Var}\left(\sum_{(x,y)\in\mathbf{X}^2}^{\neq} f_n(x,y)\right) + \mathrm{Var}\left(\sum_{x\in\mathbf{X}} h_n(x)\right) - 2\mathrm{Cov}\left(\sum_{(x,y)\in\mathbf{X}^2}^{\neq} f_n(x,y), \sum_{x\in\mathbf{X}} h_n(x)\right).
\end{aligned}$$

Lastly, by Lemmas 4.6.2-4.6.3, the limit is obtained by a standard but fastidious calculus.

5.4.5 Proof of Lemma 5.4.5

For all $n \in \mathbb{N}$, we have

$$\begin{aligned}
&\rho_0^2\sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t)dt\\
&\qquad = \sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\rho_0^2 - \widehat{\rho}_n^2\right] \widehat{K}_n(t)j_K(t)dt\\
&\qquad\quad + \sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\widehat{\rho}_n^2\widehat{K}_n(t) - \rho_0^2 K(t,\theta_0)\right] j_K(t)dt. \qquad (5.4.8)
\end{aligned}$$

Since $\mathbf{X}$ is ergodic by [56, Theorem 7], $\widehat{\rho}_n$ converges almost surely to ρ_0, see [25]. Then, by Taylor expansion of the function $x \to x^2$ at ρ_0, we have almost surely

$$\left[\rho_0^2 - \widehat{\rho}_n^2\right] = 2\rho_0\left[\rho_0 - \widehat{\rho}_n\right] + o\left(\rho_0 - \widehat{\rho}_n\right). \qquad (5.4.9)$$

Further,

$$\begin{aligned}
&2\rho_0\sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\rho_0 - \widehat{\rho}_n\right]\widehat{K}_n(t)j_K(t)dt\\
&\qquad = 2\rho_0\sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\rho_0 - \widehat{\rho}_n\right]\left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t)dt\\
&\qquad\quad + 2\rho_0\sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\rho_0 - \widehat{\rho}_n\right] K(t,\theta_0)j_K(t)dt. \qquad (5.4.10)
\end{aligned}$$

Using the notation

$$\begin{aligned}
A_n &= 2\rho_0\sqrt{|D_n|}\left[\rho_0 - \widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}} \widehat{K}_n(t)j_K(t)dt,\\
B_n &= 2\rho_0\sqrt{|D_n|}\left[\rho_0 - \widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}} \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t)dt,\\
C_n &= \sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left(\left[\rho_0 - \widehat{\rho}_n\right]2\rho_0 K(t,\theta_0) + \left[\widehat{\rho}_n^2\widehat{K}_n(t) - \rho_0^2 K(t,\theta_0)\right]\right) j_K(t)dt,
\end{aligned}$$

we have by (5.4.8)-(5.4.10),

$$\rho_0^2\sqrt{|D_n|}\int_{r_{min}}^{r_{max}}\left[\widehat{K}_n(t)-K(t,\theta_0)\right]j_K(t)dt = B_n + C_n + o\left(A_n\right). \tag{5.4.11}$$

We prove that $B_n + o(A_n)$ tends in probability to 0 and C_n tends in distribution to a Gaussian variable. Then, the proof is concluded by Slutsky's theorem and (5.4.11). By ($\mathcal{H}1$)-($\mathcal{H}2$) and Corollary 4.3.3, we have the convergence

$$\sqrt{|D_n|}\left(\rho_0-\widehat{\rho}_n\right)\xrightarrow[n\to+\infty]{distr.} N(0,\sigma^2) \tag{5.4.12}$$

where $\sigma^2=\rho_0-\int_{\mathbb{R}^d}C_{\theta_0}(x)^2dx$. Further, by Lemma 5.4.2,

$$\sup_{t\in[r_{min},r_{max}]}\left|\widehat{K}_n(t)-K(t,\theta_0)\right|\xrightarrow[n\to+\infty]{a.s.}0$$

so

$$\int_{r_{min}}^{r_{max}}\widehat{K}_n(t)j_K(t)dt\xrightarrow[n\to+\infty]{a.s.}\int_{r_{min}}^{r_{max}}K(t,\theta_0)j_K(t)dt. \tag{5.4.13}$$

Further $K(.,\theta_0)$ is continuous on $[r_{min},r_{max}]$ so $\int_{r_{min}}^{r_{max}}K(t,\theta_0)j_K(t)dt$ is finite by Lemma 5.4.3. Hence, by (5.4.12)-(5.4.13) and Slutsky's theorem, $B_n\xrightarrow[n\to+\infty]{\mathbb{P}}0$ and $o(A_n)\xrightarrow[n\to+\infty]{\mathbb{P}}0$.

Notice that

$$C_n=\sqrt{|D_n|}\left(\int_{r_{min}}^{r_{max}}H_n(t)j_K(t)dt-\left[-\int_{r_{min}}^{r_{max}}\rho_0^2K(t,\theta_0)j_K(t)dt\right]\right). \tag{5.4.14}$$

We prove the normal convergence of C_n by Cramer-Wold device, see for instance [6, Theorem 29.4]. For all $t\in[r_{min},r_{max}]$ and $s\in\mathbb{R}^p$, we have

$$s^TC_n=\sqrt{|D_n|}\left(\int_{r_{min}}^{r_{max}}H_n(t)s^Tj_K(t)dt-\left[-\int_{r_{min}}^{r_{max}}\rho_0^2K(t,\theta_0)s^Tj_K(t)dt\right]\right).$$

By (5.2.1), we have

$$\int_{r_{min}}^{r_{max}}H_n(t)s^Tj_K(t)dt=\sum_{(x,y)\in\mathbf{X}^2}f_{D_n}(x,y) \tag{5.4.15}$$

where

$$f_{D_n}(x,y):=\mathbf{1}_{\{x\in D_n\}}\int_{r_{min}}^{r_{max}}\left(\frac{\mathbf{1}_{\{y\in D_n^{\ominus t}\}}}{|D_n^{\ominus t}|}\mathbf{1}_{\{0<|x-y|\le t\}}-2\rho_0\frac{K(t,\theta_0)}{|D_n|}\mathbf{1}_{\{x-y=0\}}\right)s^Tj_K(t)dt.$$

Notice that for $t\in[r_{min},r_{max}]$, $s^Tj_K(t)\le|j_K(t)||s|$ and $K(t,\theta_0)\le K(r_{max},\theta_0)$ so we have

$$\begin{aligned}&|f_{D_n}(x,y)|\\&\quad\le\frac{|s|}{|D_n^{\ominus r_{max}}|}\mathbf{1}_{D_n}(x)\left(\mathbf{1}_{\{0<|x-y|\le r_{max}\}}+\mathbf{1}_{\{x-y=0\}}2\rho_0K(r_{max},\theta_0)\right)\int_{r_{min}}^{r_{max}}|j_K(t)|\,dt.\end{aligned} \tag{5.4.16}$$

The right-hand term in (5.4.16) is compactly supported and is bounded by Lemma 5.4.3. Moreover,

$$\mathbb{E}\left(\int_{r_{min}}^{r_{max}} \left|H_n(t)s^T j_K(t)\right| dt\right) \leq |s|\left[\mathbb{E}\left(\left|\widehat{\rho}_n^2\widehat{K}_n(t)\right|\right) + 2\rho_0 K(r_{max},\theta_0)\mathbb{E}\left(|\widehat{\rho}_n|\right)\right]\int_{r_{min}}^{r_{max}} |j_K(t)|dt.$$

Further, for $n \in \mathbb{N}$ and $t \in [r_{min}, r_{max}]$, $\widehat{\rho}_n^2\widehat{K}_n(t)$ and $\widehat{\rho}_n$ are positive and unbiased estimator of $\rho_0^2 K(t,\theta_0)$ and ρ_0, respectively, see for instance [25, Section 4.2.2]. Thus,

$$\mathbb{E}\left(\int_{r_{min}}^{r_{max}} \left|H_n(t)s^T j_K(t)\right| dt\right) \leq 3|s|\rho_0^2 K(r_{max},\theta_0)\int_{r_{min}}^{r_{max}} |j_K(t)|dt,$$

which is finite by Lemma 5.4.3. Then, by Fubini's theorem, (5.4.15) and the last equation, we have

$$\mathbb{E}\left(\sum_{(x,y)\in\mathbf{X}^2} f_{D_n}(x,y)\right) = -\int_{r_{min}}^{r_{max}} \rho_0^2 K(t,\theta_0)s^T j_K(t)dt.$$

Moreover, by (5.4.15) and Lemma 5.4.4,

$$\lim_{n\to+\infty} Var\left(\sqrt{|D_n|}\sum_{(x,y)\in\mathbf{X}^2} f_{D_n}(x,y)\right) = s^T\Sigma_{\rho_0,\theta_0}s.$$

Therefore, by (5.4.14)-(5.4.16), the last two equations and Theorem 4.3.2, we have

$$s^T C_n \xrightarrow[n\to+\infty]{distr.} N(0, s^T\Sigma_{\rho_0,\theta_0}s).$$

Finally, by Cramer-Wold device,

$$C_n \xrightarrow[n\to+\infty]{distr.} N(0, \Sigma_{\rho_0,\theta_0}).$$

5.5 Proof of Theorem 5.2.3

As in the proof of Theorem 5.2.2, we consider, without loss of generality, a given $\epsilon > 0$ such that $\Theta_{\hat{\rho}_n} \subset \Theta_{\rho_0}^{\oplus\epsilon}$, for all $n \in \mathbb{N}$. We prove below the consistency and asymptotic normality of $\widetilde{\theta}_n$ defined in Theorems 5.3.1 and 5.3.2 with $\Theta = \Theta_{\rho_0}^{\oplus\epsilon}$, $J = g$ and $\widehat{J}_n = \widehat{g}_n$. Then, for $r \geq 0$ such that $B(\theta_0, r) \subset \Theta_{\rho_0}$, we have

$$P(\widetilde{\theta}_n \in B(\theta_0, r)) \xrightarrow[n\to+\infty]{} 1.$$

Thus, by Lemma 5.4.6, with probability tending to one $\widetilde{\theta}_n \in \Theta_{\hat{\rho}_n}$ so

$$P(\widetilde{\theta}_n = \widehat{\theta}_n) \xrightarrow[n\to+\infty]{} 1.$$

Therefore, $\widehat{\theta}_n$ has the same asymptotic behaviour than $\widetilde{\theta}_n$.

Let us now determine the asymptotic properties of $\widetilde{\theta}_n$ by application of Theorems 5.3.1 and 5.3.2. The assumptions $(\mathcal{A}1)$, $(\mathcal{A}2)$, $(\mathcal{A}6)$, $(\mathcal{A}7)$ and $(\mathcal{A}9)$ are directly implied by $(\mathcal{H}1)$-$(\mathcal{H}3)$, $(\mathcal{H}_g1)$, $(\mathcal{H}_g5)$, $(\mathcal{H}_g6)$ and $(\mathcal{H}_g7)$. Moreover, $r_{min} > 0$ by $(\mathcal{H}_g1)$ so, $(\mathcal{A}4)$ is directly implied by (5.2.4), $(\mathcal{H}_g2)$, $(\mathcal{H}_g3)$ and the ergodic theorem, see [50] or [25]. By $(\mathcal{H}2)$, $R_{\theta_0}(.)$ is continuous on $[r_{min}, r_{max}]$ so is g. By [51, Corollary 1.4.14], for all $\theta \in \Theta$, if for a given $t > 0$, $|R_\theta(t)| = 1$, then R_θ is periodic of period t. This is incompatible with $(\mathcal{H}2)$ so, for all $t > 0$ and $\theta \in \Theta$, $|R_\theta(t)| < 1$. Consequently, by (5.2.3) and $(\mathcal{H}_g1)$, $g(t, \theta)$ is strictly positive for all $(t, \theta) \in [r_{min}, r_{max}] \times \Theta$. Thus, for all $c \in \mathbb{R}$, $g(., .)^c$ is well defined and strictly positive on $[r_{min}, r_{max}] \times \Theta$ so $(\mathcal{A}3)$ holds. By the same arguments, it follows that $(\mathcal{A}8)$ holds. Finally, the assumptions $(\mathcal{A}5)$ and $(\mathcal{TCL})$ are proved by Lemmas 5.5.1 and 5.5.5, respectively while the other lemmas are auxiliary results used in the following.

Lemma 5.5.1. *If $(\mathcal{H}1)$-$(\mathcal{H}3)$ and $(\mathcal{H}_g1)$-$(\mathcal{H}_g3)$ hold then, for all $r_{max} > r_{min} > 0$, there exists a set A verifying $|[r_{min}, r_{max}] \setminus A| = 0$ such that*

$$\sup_{t \in A} |\widehat{g}_n(t) - g(t, \theta_0)| \xrightarrow[n \to +\infty]{\mathbb{P}} 0.$$

Lemma 5.5.2. *If $(\mathcal{H}1)$-$(\mathcal{H}3)$, $(\mathcal{H}_g1)$ and $(\mathcal{H}_g5)$ hold then $j_g(.)$ is continuous on $[r_{min}, r_{max}]$.*

To abbreviate, we define for all $n \in \mathbb{N}$ and $t \in [r_{min}, r_{max}]$,

$$H_n^g(t) := \widehat{\rho}_n^2 \widehat{g}_n(t) - 2\rho_0 \widehat{\rho}_n g(t, \theta_0).$$

Lemma 5.5.3. *If $(\mathcal{H}1)$-$(\mathcal{H}3)$ and $(\mathcal{H}_g1)$-$(\mathcal{H}_g5)$ hold, we have for all $s \in \mathbb{R}^d$,*

$$\lim_{n \to +\infty} |D_n| \mathrm{Var}\left(\int_{r_{min}}^{r_{max}} H_n^g(t) s^T j_g(t) dt \right) = s^T \Sigma_{\rho_0, \theta_0} s$$

with $\Sigma_{\rho_0, \theta_0}$ defined as in Theorem 5.2.3.

Lemma 5.5.4. *Assume that $(\mathcal{H}1)$-$(\mathcal{H}3)$, $(\mathcal{H}_g1)$-$(\mathcal{H}_g3)$ and $(\mathcal{H}_g5)$ hold. For a given $s \in \mathbb{R}^d$ and all $n \in \mathbb{N}$, let f_{D_n} be a function from $\mathbb{R}^{2d}$ into $\mathbb{R}$ such that for all $(x, y) \in \mathbb{R}^{2d}$,*

$$\begin{aligned} &f_{D_n}(x, y) \\ &:= \mathbf{1}_{\{x \in D_n\}} \int_{r_{min}}^{r_{max}} \left(\frac{k\left(\frac{t - |x-y|}{b_n}\right) \mathbf{1}_{\{|x-y| > 0, y \in D_n\}}}{\sigma_d t^{d-1} b_n |D_n \cap D_n^{x-y}|} - \frac{2\rho_0 g(t, \theta_0)}{|D_n|} \mathbf{1}_{\{x-y=0\}} \right) s^T j_g(t) dt. \end{aligned}$$

Then, there exists $M > 0$ such that for all $(x, y) \in \mathbb{R}^{2d}$,

$$|f_{D_n}(x, y)| \leq \frac{|s| M \mathbf{1}_{\{x \in D_n\}}}{|D_n^{\ominus r_{max}+T}|} \left(\frac{1}{\sigma_d r_{min}^{d-1}} \mathbf{1}_{\{0 < |x-y| \leq r_{max}+T\}} + 2\rho_0 ||g||_\infty \mathbf{1}_{\{x-y=0\}} \right).$$

Lemma 5.5.5. *If $(\mathcal{H}1)$-$(\mathcal{H}3)$ and $(\mathcal{H}_g1)$-$(\mathcal{H}_g5)$ hold, then*

$$\sqrt{|D_n|} \int_{r_{min}}^{r_{max}} [\widehat{g}_n(t) - g(t, \theta_0)] \, j_g(t) dt \xrightarrow[n \to +\infty]{distr.} \mathcal{N}(0, \Sigma_{\rho_0, \theta_0})$$

with $\Sigma_{\rho_0, \theta_0}$ defined as in Theorem 5.2.3.

5.5.1 Proof of Lemma 5.5.1

By ($\mathcal{H}2$) and Theorem 4.2.2, $DPP(C_{\rho_0,\theta_0})$ is Brillinger mixing. Then, as ($\mathcal{H}_g2$), ($\mathcal{H}_g3$) and ($\mathcal{H}3$) hold, we have by Proposition 4.3.5,

$$\mathbb{E}\left[\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2\widehat{g}_n(t) - \rho_0^2 g(t,\theta_0)\right)^2 dt\right] \\ = \frac{2\rho_0^2}{b_n|D_n|}\int_{r_{min}}^{r_{max}} \frac{g(t,\theta_0)}{\sigma_d t^{d-1}}dt \int_{\mathbb{R}} k(x)^2 dx + O\left(\frac{1}{|D_n|}\right) + O(b_n^4). \quad (5.5.1)$$

By ($\mathcal{H}_g1$), ($\mathcal{H}_g2$) and ($\mathcal{H}3$) we have $\int_{r_{min}}^{r_{max}} \frac{g(t,\theta_0)}{\sigma_d t^{d-1}}dt \int_{\mathbb{R}} k(x)^2 dx < +\infty$. Hence, with ($\mathcal{H}_g3$), the right-hand term in (5.5.1) tends to 0 as n tends to infinity. Moreover, the term inside the expectation in (5.5.1) is positive so there exists a set A as in Lemma 5.5.1, such that

$$\sup_{t\in A}\left|\widehat{\rho}_n^2\widehat{g}_n(t) - \rho_0^2 g(t,\theta_0)\right| \xrightarrow[n\to+\infty]{\mathbb{P}} 0. \quad (5.5.2)$$

We have

$$\widehat{\rho}_n^2 \sup_{t\in A}|\widehat{g}_n(t) - g(t,\theta_0)| \le \sup_{t\in A}\left|\widehat{\rho}_n^2\widehat{g}_n(t) - \rho_0^2 g(t,\theta_0)\right| + \left(\sup_{t\in A} g(t,\theta_0)\right)\left|\widehat{\rho}_n^2 - \rho_0^2\right|.$$

By ($\mathcal{H}1$)-($\mathcal{H}2$), it follows from Corollary 4.3.3 that $\widehat{\rho}_n$ converges in probability to ρ_0. Further, by ($\mathcal{H}3$) and (5.2.3), $g(.,\theta_0)$ is bounded on $[r_{min}, r_{max}]$. Therefore, we have by (5.5.2) the convergence

$$\sup_{t\in A}|\widehat{g}_n(t) - g(t,\theta_0)| \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

5.5.2 Proof of Lemma 5.5.2

By (5.2.3), we have for all $t \in [r_{min}, r_{max}]$

$$|j_g(t)| = 2\left|w(t)\left(1 - R_{\theta_0}(t)^2\right)^{2c-2} R_{\theta_0}(t) R_{\theta_0}^{(1)}(t)\right|.$$

By ($\mathcal{H}3$), $R_{\theta_0}(.)$ and $R_{\theta_0}^{(1)}(.)$ are continuous on $[r_{min}, r_{max}]$. Further, by ($\mathcal{H}_g1$), $r_{min} > 0$ and as noticed at the beginning of the proof of Theorem 5.2.3, for all $t > 0$, $|R_{\theta_0}(t)| < 1$. Thus by ($\mathcal{H}3$), the function $t \mapsto \left(1 - R_{\theta_0}(t)^2\right)^{2c-2}$ is well defined and continuous on $[r_{min}, r_{max}]$. Finally, by ($\mathcal{H}_g5$), w is continuous on $[r_{min}, r_{max}]$ so the lemma is proved.

5.5.3 Proof of Lemma 5.5.3

Similarly to the proof of Lemma 5.4.4, we have by (5.2.4),

$$\int_{r_{min}}^{r_{max}} H_n^g(t) s^T j_g(t) dt = \sum_{(x,y)\in\mathbf{X}^2} f_n(x,y) - \sum_{x\in\mathbf{X}} h_n(x)$$

where for all $n \in \mathbb{N}$,

$$f_n(x,y) := \mathbf{1}_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} \frac{k\left(\frac{t-|x-y|}{b_n}\right)\mathbf{1}_{\{|x-y|>0, y\in D_n\}}}{\sigma_d t^{d-1} b_n |D_n \cap D_n^{x-y}|} s^T j_g(t) dt$$

and

$$h_n(x) = \frac{2\rho_0}{|D_n|}\mathbf{1}_{\{x \in D_n\}} \int_{r_{min}}^{r_{max}} g(t,\theta_0) s^T j_g(t) dt.$$

For all $n \in \mathbb{N}$ and $x \in \mathbb{R}^d$, $f_n(x,x) = 0$. Thus, by the last equation,

$$\begin{aligned}
&\mathrm{Var}\left(\int_{r_{min}}^{r_{max}} H_n^g(t) s^T j_g(t) dt\right) \\
&= \mathrm{Var}\left(\sum_{(x,y)\in \mathbf{X}^2}^{\neq} f_n(x,y)\right) + \mathrm{Var}\left(\sum_{x \in \mathbf{X}} h_n(x)\right) - 2\mathrm{Cov}\left(\sum_{(x,y)\in \mathbf{X}^2}^{\neq} f_n(x,y), \sum_{x\in\mathbf{X}} h_n(x)\right).
\end{aligned}$$

Then, the proof follows from the assumptions ($\mathcal{H}_g 1$)-($\mathcal{H}_g 5$), ($\mathcal{H}1$)-($\mathcal{H}3$) and Lemmas 4.6.2-4.6.3, by a fastidious but straightforward calculus.

5.5.4 Proof of Lemma 5.5.4

By ($\mathcal{H}_g 2$), for any $t \in [r_{min}, r_{max}]$ and $(x,y) \in \mathbb{R}^{2d}$,

$$\begin{aligned}
\left|k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{|y-x|>0,\, y\in D_n\}} &\le \left|k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{0<|y-x|<t+Tb_n\}} \\
&\le \left|k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{0<|y-x|<t+T\}}
\end{aligned}$$

as soon as $b_n < 1$ which, by ($\mathcal{H}_g 3$), we assume in the following without loss of generality. Thus, for any $t \in [r_{min}, r_{max}]$ and $(x,y) \in \mathbb{R}^{2d}$,

$$\left|k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{|y-x|>0,\, y\in D_n\}} \le \left|k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{0<|y-x|<r_{max}+T\}}. \tag{5.5.3}$$

Further, by Lemma 5.5.2, j_g is bounded on $[r_{min}, r_{max}]$ by a constant M so by (5.5.3) and Lemma 4.5.3, we have

$$\begin{aligned}
&\left| 1_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} \frac{k\left(\frac{t-|x-y|}{b_n}\right)\mathbf{1}_{\{|x-y|>0, y\in D_n\}}}{\sigma_d t^{d-1} b_n |D_n \cap D_n^{x-y}|} s^T j_g(t) dt \right| \\
&\qquad \le \mathbf{1}_{\{x\in D_n\}} \frac{|s|M}{|D_n^{\ominus r_{max}+T}|} \frac{\mathbf{1}_{\{0<|x-y|\le r_{max}+T\}}}{\sigma_d r_{min}^{d-1} b_n} \int_{r_{min}}^{r_{max}} \left|k\left(\frac{t-|x-y|}{b_n}\right)\right| dt.
\end{aligned}$$

Finally, the result follows by Lemma 4.5.3, the last inequality, ($\mathcal{H}2$) and ($\mathcal{H}_g 2$).

5.5.5 Proof of Lemma 5.5.5

The arguments of this proof are similar the the ones of the proof of Lemma 5.4.5. Notice that

$$\begin{aligned}
\rho_0^2 \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} [\widehat{g}_n(t) - g(t,\theta_0)]\, j_g(t) dt &= \sqrt{|D_n|}\Bigg(\left[\rho_0^2 - \widehat{\rho}_n^2\right] \int_{r_{min}}^{r_{max}} \widehat{g}_n(t) j_g(t) dt \\
+ \int_{r_{min}}^{r_{max}} \left[\widehat{\rho}_n^2 \widehat{g}_n(t) - \mathbb{E}\left[\widehat{\rho}_n^2 \widehat{g}_n(t)\right]\right] j_g(t) dt &+ \int_{r_{min}}^{r_{max}} \left[\mathbb{E}\left[\widehat{\rho}_n^2 \widehat{g}_n(t)\right] - \rho_0^2 g(t,\theta_0)\right] j_g(t) dt \Bigg)
\end{aligned} \tag{5.5.4}$$

and

$$\sqrt{|D_n|}\left[\rho_0-\widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}}\widehat{g}_n(t)j_g(t)dt =$$

$$\sqrt{|D_n|}\left[\rho_0-\widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}}\left[\widehat{g}_n(t)-g(t,\theta_0)\right]j_g(t)dt+\sqrt{|D_n|}\left[\rho_0-\widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}}g(t,\theta_0)j_g(t)dt. \tag{5.5.5}$$

Denote

$$\begin{aligned}
T_n &= 2\rho_0\sqrt{|D_n|}\left[\rho_0-\widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}}\widehat{g}_n(t)j_g(t)dt\\
U_n &= 2\rho_0\sqrt{|D_n|}\left[\rho_0-\widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}}\left[\widehat{g}_n(t)-g(t,\theta_0)\right]j_g(t)dt\\
V_n &= \sqrt{|D_n|}\int_{r_{min}}^{r_{max}}\left[\mathbb{E}\left[\widehat{\rho}_n^2\widehat{g}_n(t)\right]-\rho_0^2 g(t,\theta_0)\right]j_g(t)dt\\
W_n &= \sqrt{|D_n|}\int_{r_{min}}^{r_{max}}\left[\widehat{\rho}_n^2\widehat{g}_n(t)-2\rho_0\widehat{\rho}_n g(t,\theta_0)-\left(\mathbb{E}\left[\widehat{\rho}_n^2\widehat{g}_n(t)\right]-2\rho_0^2 g(t,\theta_0)\right)\right]j_g(t)dt.
\end{aligned}$$

Therefore, by using (5.4.9) in the proof of Lemma 5.4.5, (5.5.4) and (5.5.5),

$$\rho_0^2\sqrt{|D_n|}\int_{r_{min}}^{r_{max}}\left[\widehat{g}_n(t)-g(t,\theta_0)\right]j_g(t)dt = U_n+V_n+W_n+o\left(T_n\right). \tag{5.5.6}$$

We prove that $U_n+V_n+o(T_n)$ tends in probability to 0 and we conclude by proving that W_n tends in distribution to a Gaussian variable. We may use again (5.4.12) so by Slutsky's theorem and Lemmas 5.5.1-5.5.2 we have $U_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$. Further, since g is continuous on $[r_{min}, r_{max}]$ so bounded, we have by Lemma 5.5.1 that $\widehat{g}_n$ is uniformly bounded in probability on $[r_{min}, r_{max}]$, see [64, Prohorov's theorem]. Thus, by (5.4.12) and Lemma 5.5.2, $o(T_n) \xrightarrow[n\to+\infty]{\mathbb{P}} 0$. Further, under $(\mathcal{H}1)$-$(\mathcal{H}2)$ and $(\mathcal{H}_g1)$-$(\mathcal{H}_g4)$, we may apply Lemma 4.5.2 so by Lemma 5.5.2, $V_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

We prove the normal convergence of W_n by Cramer-Wold device. To shorten, denote for all $n\in\mathbb{N}$ and $s\in\mathbb{R}^p$,

$$X_n^s := \int_{r_{min}}^{r_{max}} H_n^g(t)s^T j_g(t)dt.$$

By Lemma 5.5.2, j_g is bounded on $[r_{min}, r_{max}]$ by a constant M. Then, since for all $t\in[r_{min}, r_{max}]$ and $n\in\mathbb{N}$,

$$H_n^g(t)=\widehat{\rho}_n^2\widehat{g}_n(t)-\rho_0^2 g(t,\theta_0)+(\rho_0-\widehat{\rho}_n)\rho_0 g(t,\theta_0)-\rho_0\widehat{\rho}_n g(t,\theta_0),$$

we have

$$\begin{aligned}
\mathbb{E}\left(\int_{r_{min}}^{r_{max}}\left|H_n^g(t)s^T j_g(t)\right|dt\right) &\le |s|M\mathbb{E}\int_{r_{min}}^{r_{max}}\left(\left|\widehat{\rho}_n^2\widehat{g}_n(t)-\rho_0^2 g(t,\theta_0)\right|\right)dt\\
&\quad + |s|M\left[\mathbb{E}(|\rho_0-\widehat{\rho}_n|)+\mathbb{E}(\widehat{\rho}_n)\right]\int_{r_{min}}^{r_{max}}|\rho_0 g(t,\theta_0)|dt.
\end{aligned} \tag{5.5.7}$$

By $(\mathcal{H}3)$, $g(.,\theta_0)$ is bounded on $[r_{min}, r_{max}]$. Denote $||g||_\infty$ its maximum so by Cauchy Schwartz inequality, Jensen inequality and (5.5.7), we have

$$\int_{r_{min}}^{r_{max}} \mathbb{E}\left|H_n^g(t)s^T j_g(t)\right| dt \leq |s|M(r_{max}-r_{min})^{\frac{1}{2}}\left(\mathbb{E}\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2\widehat{g}_n(t)-\rho_0^2 g(t,\theta_0)\right)^2 dt\right)^{\frac{1}{2}} + |s|M(r_{max}-r_{min})\rho_0||g||_\infty\left(\mathbb{E}(|\rho_0-\widehat{\rho}_n|)+\mathbb{E}(\widehat{\rho}_n)\right). \quad (5.5.8)$$

By the same arguments as in the proof of Lemma 5.5.1, we have

$$\mathbb{E}\left[\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2\widehat{g}_n(t)-\rho_0^2 g(t,\theta_0)\right)^2 dt\right] = \frac{2\rho_0^2}{b_n|D_n|}\int_{r_{min}}^{r_{max}}\frac{g(t,\theta_0)}{\sigma_d t^{d-1}}dt\int_{\mathbb{R}} k(x)^2 dx + O\left(\frac{1}{|D_n|}\right) + O(b_n^4).$$

Thus by $(\mathcal{H}_g3)$, $\mathbb{E}\left(\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2\widehat{g}_n(t)-\rho_0^2 g(t,\theta_0)\right)^2 dt\right)$ tends to 0. Moreover, as noticed in [24], $\widehat{\rho}_n$ converge in L^1 to ρ_0 so $\mathbb{E}(|\rho_0-\widehat{\rho}_n|)+\mathbb{E}(\widehat{\rho}_n)$ converges to ρ_0. Hence, by (5.5.7)-(5.5.8) $\mathbb{E}\int_{r_{min}}^{r_{max}}\left|H_n^g(t)s^T j_g(t)\right| dt$ is bounded. Then, by Fubini theorem,

$$\mathbb{E}(X_n^s) = \int_{r_{min}}^{r_{max}} \mathbb{E}\left(H_n^g(t)\right) s^T j_g(t) dt$$

which implies that

$$s^T W_n = \sqrt{|D_n|}\left(X_n^s - \mathbb{E}(X_n^s)\right).$$

By (5.2.4), we have

$$X_n^s = \sum_{(x,y)\in\mathbf{X}^2} f_{D_n}(x,y), \quad (5.5.9)$$

where

$$f_{D_n}(x,y) := \mathbf{1}_{\{x\in D_n\}}\int_{r_{min}}^{r_{max}}\left(\frac{k\left(\frac{t-|x-y|}{b_n}\right)\mathbf{1}_{\{|x-y|>0, y\in D_n\}}}{\sigma_d t^{d-1} b_n |D_n\cap D_n^{x-y}|} - \frac{2\rho_0 g(t,\theta_0)}{|D_n|}\mathbf{1}_{\{x-y=0\}}\right) s^T j_g(t) dt.$$

By Lemma 5.5.4,

$$|f_{D_n}(x,y)| \leq \frac{|s|M\mathbf{1}_{\{x\in D_n\}}}{|D_n^{\ominus r_{max}+T}|}\left(\frac{1}{\sigma_d r_{min}^{d-1}}\mathbf{1}_{\{0<|x-y|\leq r_{max}+T\}} + 2\rho_0||g||_\infty\mathbf{1}_{\{x-y=0\}}\right). \quad (5.5.10)$$

The right-hand term in (5.5.10) is bounded and compactly supported. Therefore, by Lemma 5.5.3 and Theorem 4.3.2, we have the convergence

$$\sqrt{|D_n|}\left(X_n^s - \mathbb{E}(X_n^s)\right) \xrightarrow[n\to+\infty]{distr.} N(0, s^T\Sigma_{\rho_0,\theta_0}s).$$

Finally, we deduce from the last convergence that

$$W_n \xrightarrow[n\to+\infty]{distr.} N(0, \Sigma_{\rho_0,\theta_0}).$$

Bibliography

[1] Abramowitz, M., and Stegun, I. A. *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, vol. 55 of *National Bureau of Standards Applied Mathematics Series.* U.S. Government Printing Office, Washington, D.C., 1964. 44, 50, 57, 58, 59, 60, 73, 74

[2] Álvarez-Nodarse, R., and Moreno-Balcázar, J. J. Asymptotic properties of generalized Laguerre orthogonal polynomials. *Indagationes Mathematicae. New Series 15*, 2 (2004), 151–165. 59

[3] Baddeley, A. Spatial point patterns: models and statistics. In *Stochastic geometry, spatial statistics and random fields*, vol. 2068 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2013, pp. 49–114. 9, 23

[4] Baddeley, A., and Turner, R. Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software 12*, 6 (2005), 1–42. URL: `www.jstatsoft.org`, ISSN: 1548-7660. 89

[5] Baraud, Y. Estimation of the density of a determinantal process. *Confluentes Mathematici 5*, 1 (2013), 3–21. 18, 32

[6] Billingsley, P. *Probability and measure.* John Wiley & Sons, New York-Chichester-Brisbane, 1979. Wiley Series in Probability and Mathematical Statistics. 103

[7] Biscio, C. A. N., and Lavancier, F. Quantifying repulsiveness of determinantal point processes. *To appear in Bernoulli* (2015). 61, 62, 64, 66, 67

[8] Borodin, A., and Serfaty, S. Renormalized energy concentration in random matrices. *Communications in Mathematical Physics 320*, 1 (2013), 199–244. 43

[9] Daley, D. J., and Vere-Jones, D. *An Introduction to the Theory of Point Processes, Vol. I, Elementary Theory and Methods*, 2 ed. Springer, 2003. 11, 24, 38, 40, 69, 72, 86, 87

[10] Daley, D. J., and Vere-Jones, D. *An Introduction to the Theory of Point Processes, Vol. II, General Theory and Structure*, 2 ed. Springer, 2008. 11, 24, 40, 69, 71

[11] Deng, N., Zhou, W., and Haenggi, M. The Ginibre point process as a model for wireless networks with repulsion. *to appear in IEEE Transactions on Wireless Communications. doi: 10.1109/TWC.2014.2332335* (2014). 18, 32, 37

[12] Domar, Y. An extremal problem for positive definite functions. *Journal of Mathematical Analysis and Applications 52* (1975), 56–63. 44

[13] Dudley, R. M. *Real Analysis and Probability*, vol. 74. Cambridge University Press, 2002. 71

[14] Ehm, W., Gneiting, T., and Richards, D. Convolution roots of radial positive definite functions with compact support. *Transactions of the American Mathematical Society 356*, 11 (2004), 4655–4685. 45, 52, 55

[15] Evans, L. *Partial Differential Equations*, 2 ed., vol. 19 of *Graduate Studies in Mathematics*. American Mathematical Society, 2010. 56, 57

[16] Fasshauer, G. E., and Zhang, J. G. Iterated approximate moving least squares approximation. In *Advances in Meshfree Techniques*. Springer, 2007, pp. 221–239. 48

[17] Flyer, N. Exact polynomial reproduction for oscillatory radial basis functions on infinite lattices. *Computers & Mathematics with Applications. 51*, 8 (2006), 1199–1208. 47, 58

[18] Fornberg, B., Larsson, E., and Wright, G. A new class of oscillatory radial basis functions. *Computers & Mathematics with Applications 51*, 8 (2006), 1209–1222. 47

[19] Garsia, A., Rodemich, E., and Rumsey, H. On some extremal positive definite functions. *Journal of Mathematics and Mechanics 18*, 9 (1969), 805–834. 44

[20] Giaquinta, M., and Hildebrandt, S. *Calculus of Variations I.* Springer, 1996. 57

[21] Gradshteyn, I. S., and Ryzhik, I. M. *Table of Integrals, Series, and Products*, seventh ed. Elsevier/Academic Press, Amsterdam, 2007. 58, 60

[22] Grafakos, L. *Classical Fourier Analysis.* Graduate Texts in Mathematics. Springer, 2008. 42, 58, 60, 61

[23] Guan, Y., and Sherman, M. On least squares fitting for stationary spatial point processes. *Journal of the Royal Statistical Society. Series B. Statistical Methodology 69*, 1 (2007), 31–49. 89, 94

[24] Heinrich, L. Minimum contrast estimates for parameters of spatial ergodic point processes. In *Transactions of the 11th Prague conference on random processes, information theory and statistical decision functions* (1992), vol. 479, p. 492. 109

[25] Heinrich, L. Asymptotic methods in statistics of random point processes. In *Stochastic Geometry, Spatial Statistics and Random Fields*, vol. 2068 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2013, pp. 115–150. 69, 77, 90, 101, 102, 104, 105

[26] Heinrich, L., and Klein, S. Central limit theorem for the integrated squared error of the empirical second-order product density and goodness-of-fit tests for stationary point processes. *Statistics & Risk Modeling with Applications in Finance and Insurance 28*, 4 (2011), 359 – 397. 20, 21, 34, 75, 77, 78

[27] Heinrich, L., and Klein, S. Central limit theorems for empirical product densities of stationary point processes. *Statistical Inference for Stochastic Processes. An International Journal Devoted to Time Series Analysis and the Statistics of Continuous Time Processes and Dynamical Systems 17*, 2 (2014), 121–138. 20, 21, 34, 75, 77, 83

[28] Heinrich, L., and Pawlas, Z. Absolute regularity and Brillinger-mixing of stationary point processes. *Lithuanian Mathematical Journal 53*, 3 (2013), 293–310. 69

[29] Heinrich, L., and Prokešová, M. On estimating the asymptotic variance of stationary point processes. *Methodology and Computing in Applied Probability 12*, 3 (2010), 451–471. 20, 34, 75, 86

[30] Heinrich, L., and Schmidt, V. Normal convergence of multidimensional shot noise and rates of this convergence. *Advances in Applied Probability 17*, 4 (1985), 709–730. 69

[31] Hiriart-Urruty, J.-B., and Lemaréchal, C. *Convex Analysis and Minimization Algorithms: Part 1: Fundamentals*, vol. 305. Springer, 1993. 57

[32] Hough, J. B., Krishnapur, M., Peres, Y., and Virág, B. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, vol. 51 of *University Lecture Series*. American Mathematical Society, Providence, RI, 2009. 15, 28, 29, 37, 40, 41, 46

[33] Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley and Sons, Chichester, 2008. 9, 13, 23, 26, 38, 39

[34] Janson, S. Normal convergence by higher semi-invariants with applications to sums of dependent random variables and random graphs. *The Annals of Probability 16*, 1 (1988), 305–312. 80, 83

[35] Jolivet, E. Central limit theorem and convergence of empirical processes for stationary point processes. In *Point processes and queuing problems (Colloqium, Keszthely, 1978)*, vol. 24 of *Colloq. Math. Soc. János Bolyai*. North-Holland, Amsterdam-New York, 1981, pp. 117–161. 76, 79, 80, 81

[36] Jolivet, E. Upper bound of the speed of convergence of moment density estimators for stationary point processes. *Metrika. International Journal for Theoretical and Applied Statistics 31*, 6 (1984), 349–360. 20, 34, 75

[37] Jolivet, E. Moment estimation for stationary point processes in $\mathbf{R}^d$. In *Spatial statistics and imaging (Brunswick, ME, 1988)*, vol. 20 of *IMS Lecture Notes Monogr. Ser.* Inst. Math. Statist., Hayward, CA, 1991, pp. 138–149. 20, 34, 75

[38] Kölbig, K. S., and Scherb, H. On a Hankel transform integral containing an exponential function and two Laguerre polynomials. *Journal of Computational and Applied Mathematics 71*, 2 (1996), 357–363. 59

[39] Kolountzakis, M., and Révész, S. On a problem of Turán about positive definite functions. *Proceedings of the American Mathematical Society 131*, 11 (2003), 3423–3430. 44

[40] Krickeberg, K. Processus ponctuels en statistique. In *Tenth Saint Flour Probability Summer School—1980 (Saint Flour, 1980)*, vol. 929 of *Lecture Notes in Math.* Springer, Berlin-New York, 1982, pp. 205–313. 20, 34, 69, 76

[41] Kulesza, A., and Taskar, B. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning 5* (2012), 123–286. 18, 32, 37

[42] Kuna, T., Lebowitz, J. L., and Speer, E. R. Realizability of point processes. *Journal of Statistical Physics 129*, 3 (2007), 417–439. 39

[43] Lavancier, F., Møller, J., and Rubak, E. Determinantal point process models and statistical inference : Extended version. *arXiv:1205.4818v5* (2014). 10, 18, 24, 32, 37, 38, 39, 42, 74, 90

[44] Lavancier, F., Møller, J., and Rubak, E. Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 77*, 4 (2015), 853–877. 15, 16, 17, 18, 19, 28, 29, 30, 31, 32, 33, 37, 38, 41, 42, 43, 46, 47, 90

[45] Lieb, E. H., and Loss, M. *Analysis*, second ed., vol. 14 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. 51, 55, 56, 60

[46] Macchi, O. The coincidence approach to stochastic point processes. *Advances in Applied Probability 7* (1975), 83–122. 14, 18, 28, 32, 37, 41, 43, 74

[47] Miyoshi, N., and Shirai, T. A cellular network model with Ginibre configured base stations. *Advances in Applied Probability 46*, 3 (2014), 832–845. 18, 32, 37

[48] Møller, J., Huber, M. L., and Wolpert, R. L. Perfect simulation and moment properties for the Matérn type III process. *Stochastic Processes and their Applications 120*, 11 (2010), 2142–2158. 12, 26

[49] Møller, J., and Waagepetersen, R. P. *Statistical Inference and Simulation for Spatial Point Processes*, vol. 100 of *Monographs on Statistics and Applied Probability.* Chapman & Hall/CRC, Boca Raton, FL, 2004. 10, 12, 13, 24, 25, 27, 39, 77, 93

[50] Nguyen, X.-X., and Zessin, H. Ergodic theorems for spatial processes. *Z. Wahrsch. Verw. Gebiete 48*, 2 (1979), 133–158. 105

[51] Sasvári, Z. *Multivariate Characteristic and Correlation Functions*, vol. 50. Walter de Gruyter, 2013. 41, 52, 99, 105

[52] Scardicchio, A., Zachary, C. E., and Torquato, S. Statistical properties of determinantal point processes in high-dimensional Euclidean spaces. *Physical Review E. Statistical, Nonlinear, and Soft Matter Physics 79*, 4 (2009), 041108, 19. 18, 32

[53] Schoenberg, I. J. Metric spaces and completely monotone functions. *Annals of Mathematics. Second Series 39*, 4 (1938), 811–841. 47, 58

[54] Shirai, T., and Takahashi, Y. Random point fields associated with certain Fredholm determinants. I. Fermion, Poisson and boson point processes. *Journal of Functional Analysis 205*, 2 (2003), 414–463. 73

[55] Shiryaev, A. *Probability.* Graduate Texts in Mathematics. Springer, 1995. 50, 53, 80

[56] Soshnikov, A. Determinantal random point fields. *Russian Mathematical Surveys 55* (2000), 923–975. 37, 77, 101, 102

[57] Soshnikov, A. Gaussian limit for determinantal random point fields. *The Annals of Probability 30*, 1 (2002), 171–187. 20, 34, 77

[58] Stein, E., and Weiss, G. *Introduction to Fourier Analysis on Euclidean Spaces (PMS-32)*, vol. 1. Princeton university press, 1971. 15, 28, 41

[59] Stoyan, D., Kendall, W. S., and Mecke, J. *Stochastic Geometry and Its Applications.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons Ltd., Chichester, 1987. 12, 25, 26, 38

[60] Stoyan, D., and Stoyan, H. On one of Matérn's hard-core point process models. *Mathematische Nachrichten 122* (1985), 205–214. 12, 26, 69

[61] Stoyan, D., and Stoyan, H. *Fractals, Random Shapes and Point Fields.* Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics. John Wiley & Sons, Ltd., Chichester, 1994. Methods of geometrical statistics, Translated from the 1992 German original by N. Bamber and R. B. Johnson. 12, 26

[62] Szegö, G. *Orthogonal Polynomials*, vol. 23. AMS Bookstore, 1992. 59

[63] Teichmann, J., Ballani, F., and van den Boogaart, K. Generalizations of Matérn's hard-core point processes. *Spatial Statistics 3* (2013), 33–53. 12, 26

[64] van der Vaart, A. W. *Asymptotic Statistics*, vol. 3 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, 1998. 108

[65] van Lieshout, M. N. M. *Markov Point Processes and Their Applications.* Imperial College Press, London, 2000. 13, 27

[66] Waagepetersen, R., and Guan, Y. Two-step estimation for inhomogeneous spatial point processes. *Journal of the Royal Statistical Society. Series B. Statistical Methodology 71*, 3 (2009), 685–702. 89

[67] Watson, G. N. *A Treatise on the Theory of Bessel Functions.* Cambridge university press, 1995. 57

l'université nantes angers le mans
COMMUNAUTÉ D'UNIVERSITÉS ET ÉTABLISSEMENTS

UNIVERSITÉ DE NANTES

Thèse de Doctorat

Christophe A. N. Biscio

Contribution à la modélisation et à l'estimation paramétrique des processus ponctuels déterminantaux

Contribution to the modelling and the parametric estimation of determinantal point processes

Résumé

Ce manuscrit est dédié à l'étude des processus ponctuels déterminantaux (DPPs) ainsi qu'à leur estimation paramétrique. Ces processus sont connus pour modéliser des phénomènes répulsifs (le cas où les points ont tendance à se repousser entre eux). Dans la première partie, nous étudions la richesse de ce modèle en proposant deux définitions de la répulsion, basées sur les moments d'ordre deux du processus, plus précisément la fonction de corrélation par paires (pcf) g. Cela nous conduit à la détermination du DPP le plus répulsif. Nous présentons ensuite de nouvelles familles paramétriques de DPPs couvrant toute la plage de répulsion offerte par cette classe. Dans la seconde partie, nous démontrons que les DPPs stationnaires sont Brillinger-mélangeants. Cette propriété basée sur les moments du processus permet d'obtenir des résultats asymptotiques sur différentes statistiques. En particulier, nous démontrons un théorème limite central sur une large classe de fonctionnelles d'ordre p et nous adaptons au cadre des DPPs certains résultats déjà connus sur les estimateurs à noyaux de la pcf g d'un processus Brillinger-mélangeant. Finalement, dans la dernière partie, nous étudions l'estimation d'un DPP paramétrique par minimum de contraste, basée sur les fonctions de Ripley K et la pcf g. Nous prouvons la consistance et la normalité asymptotique de ces méthodes et en particulier, nous obtenons une forme explicite de la variance asymptotique. Ces résultats sont en fait des cas particuliers d'un théorème plus général sur les propriétés asymptotiques des estimateurs par minimum de contraste d'un processus ponctuel stationnaire que nous énonçons et prouvons dans ce chapitre.

Mots clés

processus ponctuel déterminantal, fonction de corrélation par paires, mélange de Brillinger, fonction de covariance, estimation par minimum de contraste.

Abstract

This manuscript is devoted to the study of determinantal point processes (DPPs) and their parametric estimation. These processes are known to be well adapted to inhibitive point patterns, where the points tend to repel each others. In the first chapter, we study the flexibility of this model by suggesting two definitions of the repulsiveness, both based on the second order moments of the process, namely the pair correlation function (pcf) g. This leads us to identify the most repulsive DPP. Then, we introduce new parametric families of DPPs that cover a large range of DPPs, from the stationary Poisson process to the most repulsive DPP. In the second chapter, we prove that stationary DPPs are Brillinger mixing. This property allows us to deduce asymptotic results on several statistics. Namely, we prove a central limit theorem for a wide class of functionals of order p of the process and we adapt to DPPs some results already known on kernel estimators of the function g of Brillinger mixing point processes. Finally, in the last chapter we study the minimum contrast estimation based on the Ripley's K-function and the pcf g. We prove the consistency and the asymptotic normality of these methods for stationary DPPs. In particular, we obtain an explicit form of the asymptotic variance. These results are in fact particular cases of a more general theorem dealing with the asymptotic properties of minimum contrast estimation for stationary point processes that we state and prove in this chapter.

Key Words

determinantal point process, pair correlation function, Brillinger mixing, covariance function, minimum contrast estimation.