HAL
open science

Contrast estimation for parametric stationary determinantal point processes

Christophe Ange Napoléon Biscio, Frédéric Lavancier

To cite this version:

Christophe Ange Napoléon Biscio, Frédéric Lavancier. Contrast estimation for parametric stationary determinantal point processes. Scandinavian Journal of Statistics, 2016. hal-01215582

HAL Id: hal-01215582

https://hal.science/hal-01215582

Submitted on 14 Oct 2015

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

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

Contrast estimation for parametric stationary determinantal point processes

Christophe Ange Napoléon Biscio[1] and Frédéric Lavancier[1,2]

[1] Laboratoire de Mathématiques Jean Leray, University of Nantes, France
[2] Inria, Centre Rennes Bretagne Atlantique, France

Abstract

We study minimum contrast estimation for parametric stationary determinantal point processes. These processes form a useful class of models for repulsive (or regular, or inhibitive) point patterns and are already applied in numerous statistical applications. Our main focus is on minimum contrast methods based on the Ripley's K-function or on the pair correlation function. Strong consistency and asymptotic normality of theses procedures are proved under general conditions that only concern the existence of the process and its regularity with respect to the parameters. A key ingredient of the proofs is the recently established Brillinger mixing property of stationary determinantal point processes. This work may be viewed as a complement to the study of Y. Guan and M. Sherman who establish the same kind of asymptotic properties for a large class of Cox processes, which in turn are models for clustering (or aggregation).

Keywords: Ripley's K function, pair correlation function, Brillinger mixing, central limit theorem.

1 Introduction

Determinantal point processes (DPPs) are models for repulsive (or regular, or inhibitive) point processes data. They have been introduced by O. Macchi in [21] to model the position of fermions, which are particles that repel each others. Their probabilistic aspects have been studied thoroughly, in particular in [28], [27] and [13]. Recently, DPPs have been studied and applied from a statistical perspective. A description of their main statistical aspects is conducted in [19] and they actually turn out to be a well-adapted statistical model in domains as statistical learning [16], telecommunications [8, 22], biology and ecology (see the examples in [19] and [18]).

A DPP is defined through a kernel C, basically a covariance function. Assuming a parametric form for C, several estimation procedures are considered in [19], specifically the maximum likelihood method and minimum contrast procedures based on the Ripley's K function or the pair correlation g. These methods are implemented in the `spatstat` library [1, 2] of R [25]. From the simulation study conducted in [19]

and [18], see also Section 2.2, the maximum likelihood procedure seems to be the best method in terms of quadratic loss. However, the expression of the likelihood relies in theory on a spectral representation of C, which is rarely known in practice, and some Fourier approximations are introduced in [19]. The likelihood also involves the determinant of a $n \times n$ matrix, where n is the number of observed points, which is prohibitively time consuming to compute when n is large. In contrast, the estimation procedures based on K or g do not require the knowledge of any spectral representation of C and are faster to compute in presence of large datasets, which explain their importance in practice.

From a theoretical point of view, neither the likelihood method nor the minimum contrast methods for DPPs have been studied thoroughly, even in assuming that a spectral method for C is known. In this work, we focus on parametric stationary DPPs and we prove the strong consistency and the asymptotic normality of the minimum contrast estimators based on K and g. These questions are in connection with the general investigation of Y. Guan and M. Sherman [10], who study the asymptotic properties of the latter estimators for stationary point processes. However the setting in [10] has a clear view to Cox processes and the assumptions involve both α-mixing and Brillinger mixing conditions, which are indeed satisfied for a large class of Cox processes. Unfortunately these results do no apply straightforwardly to DPPs. We consider instead more general versions of the asymptotic theorems in [10] and we prove that they apply nicely to DPPs. Our main ingredient then becomes the Brillinger mixing property of stationary DPPs, recently proved in [4], and we do not need any α-mixing condition. Our asymptotic results finally gather a very large class of stationary DPPs, where the main assumptions are quite standard and only concern the regularity of the kernel C with respect to the parameters. As an extension to the results in [10], it is worth mentioning the study of [32] dealing with constrast estimation for some inhomogeneous spatial point processes, still under a crucial α-mixing condition. We do not address this generalization for DPPs in the present work.

The remainder of this paper is organized as follows. In Section 2, we recall the definition of stationary DPPs, some of their basic properties and we discuss parametric estimation of DPPs. Our main results are presented in Section 3, namely the asymptotic properties of the minimum contrast estimators of a DPP based on the K or the g function. Section 4 gathers the proofs of our main results. In the appendix, we finally present our general asymptotic result for minimum contrast estimators and some auxiliary materials.

2 Stationary DPPs and parametric estimation

2.1 Stationary DPPs

We refer to [6, 7] for a general presentation on point processes. Let $\mathbf{X}$ be a simple point process on $\mathbb{R}^d$. For a bounded set $D \subset \mathbb{R}^d$, denote by $\mathbf{X}(D)$ the number of points of $\mathbf{X}$ in D and let E be the expectation over the distribution of $\mathbf{X}$. If there exists a function $\rho^{(k)} : (\mathbb{R}^d)^k \to \mathbb{R}^+$, for $k \geq 1$, such that for any family of mutually

disjoint subsets $D_1, \dots, D_k$ in $\mathbb{R}^d$

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

then this function is called the joint intensity of order k of $\mathbf{X}$. If $\mathbf{X}$ is stationary, $\rho^{(k)}(x_1, \dots, x_k) = \rho^{(k)}(0, x_2 - x_1, \dots, x_k - x_1)$ and in particular $\rho^{(1)} = \rho$ is a constant. From its definition, the joint intensity of order k is unique up to a Lebesgue nullset. Henceforth, for ease of presentation, we ignore nullsets. In particular we will say that a function is continuous whenever there exists a continuous version of it.

Determinantal point processes (DPPs) are defined through their joint intensities. We refer to the survey by Hough et al. [13] for a general presentation including the non-stationary case and the extension to complex-valued kernels. We focus in this work on stationary DPPs and so we restrict the definition to this subclass. We also consider for simplicity real-valued kernels.

Definition 2.1. *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 exists and satisfies the relation*

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

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

Conditions on C ensuring the existence of $DPP(C)$ are recalled in the next proposition. 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 we consider its extension to $L^2(\mathbb{R}^d)$ by Plancherel's theorem, see [30].

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 $0 \leq \mathcal{F}(C) \leq 1$.

Proposition 2.2 ([28, 19])**.** *Assume C satisfies $\mathcal{K}(\rho)$. Then $DPP(C)$ exists and is unique if and only if $0 \leq \mathcal{F}(C) \leq 1$.*

In short, $DPP(C)$ exists whenever C is a continuous covariance function in $L^2(\mathbb{R}^d)$ with $\mathcal{F}(C) \leq 1$. This makes easy the construction of parametric families of DPPs, simply considering parametric families of covariance functions where the condition $\mathcal{F}(C) \leq 1$ appears as a constraint on the parameters. Some examples are given in [19], [5] and in the next section.

By definition, all moments of a DPP are known, in particular the pair correlation (pcf) and the Ripley's K-function can explicitly be expressed in terms of the kernel. For C satisfying $\mathcal{K}(\rho)$, let $R(x) = C(x)/C(0)$ be the correlation function associated to C. The pcf, defined in the stationary case for all $x \in \mathbb{R}^d$ by $g(x) = \rho^{(2)}(0, x)/\rho^2$, writes

$$g(x) = 1 - R^2(x).$$

The Ripley's K-function is in turn given for all $t \geq 0$ by

$$K(t) = \int_{B(0,t)} g(x)dx = \int_{B(0,t)} (1 - R^2(x))dx \tag{2.1}$$

where $B(0,t)$ is the Euclidean ball centred at 0 with radius t. For later purposes, we denote by $c_{[k]}^{red}$ the density of the reduced factorial cumulant moment measures of order k of $\mathbf{X}$. We refer to [4] for the definition and further details, where the following particular cases are derived. Assuming that the kernel C of $\mathbf{X}$ satisfies $\mathcal{K}(\rho)$, we have for all $(u,v,w) \in \mathbb{R}^{3d}$

$$c_{[2]}^{red}(u) = -C^2(u), \tag{2.2}$$

$$c_{[3]}^{red}(u,v) = \ 2\ C(u)C(v)C(v-u), \tag{2.3}$$

$$\begin{aligned} c_{[4]}^{red}(u,v,w) = -2\Big[&C(u)C(v)C(u-w)C(v-w) + C(u)C(w)C(u-v)C(v-w) \\ &+ C(v)C(w)C(u-v)C(u-w)\Big]. \end{aligned} \tag{2.4}$$

2.2 Parametric estimation of DPPs

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 the point process $\mathbf{X} \sim DPP(C_{\rho_0,\theta_0})$ on a bounded domain $D_n \subset \mathbb{R}^d$.

The standard estimator of the intensity ρ_0 is

$$\widehat{\rho}_n = \frac{1}{|D_n|} \sum_{x \in \mathbf{X}} \mathbf{1}_{\{x \in D_n\}} \tag{2.5}$$

where $|D_n|$ denotes the Lebesgue volume of D_n. Since a stationary DPP is ergodic, see [28], this estimator is strongly consistent by the ergodic theorem, and it is asymptotically normal, cf [29] and [4]. In the following, we focus our attention on the estimation of θ_0. As explained in [18], 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 Fourier approximations are introduced in [18]. Another option is to consider minimum contrast estimators (MCE) as described below.

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, that we study in detail in the following. Consider $\widehat{J}_n$ an estimator of J from the observation of $\mathbf{X}$ on D_n. Further, let $c \in \mathbb{R}$, $c \neq 0$, be a parameter such that $\widehat{J}_n(t)^c$ and $J(t,\theta)^c$ are well defined for all $t \in \mathbb{R}$ and $\theta \in \Theta_{\rho_0}$. Finally, define for $0 \leq r_{min} < r_{max}$, the discrepancy measure

$$U_n(\theta) = \int_{r_{min}}^{r_{max}} w(t) \left\{ \widehat{J}_n(t)^c - J(t,\theta)^c \right\}^2 dt \tag{2.6}$$

where w is a smooth weight function. The MCE of θ_0 is

$$\widehat{\theta}_n = \arg\min_{\theta \in \Theta_{\hat{\rho}_n}} U_n(\theta). \tag{2.7}$$

For example, let us consider the parametric family of DPPs with Gaussian kernels

$$C(x) = \rho e^{-\left|\frac{x}{\alpha}\right|^2}, \quad x \in \mathbb{R}^d, \tag{2.8}$$

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 $K(\rho)$. Some realizations are shown in Figure 1. For comparison, we have estimated the parameter α of this model with the MCE (2.7) when J corresponds to K or g, and with the maximum likelihood method (using the Fourier approximation of the spectral representation of C introduced in [19]). The estimators of K and g, in place of $\widehat{J}_n$ in (2.7), are standard and recalled in Sections 3.2-3.3, see also [23, Chapter 4]. For the tuning parameters, we followed the standard choice $w(t) = 1$, $r_{\min} = 0.01$, $r_{\max}$ as one quarter of the side length of the window and $c = 0.5$ as recommended in [9] for repulsive point processes. This simulation study has been carried out with the functions implemented in the `spatstat` library. Table 1 reports the mean squared errors of the three mentioned methods over 500 realisations of $DPP(C)$ with $\rho = 100$ and $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$, observed on $[0,1]^2$, $[0,2]^2$ and $[0,3]^2$.

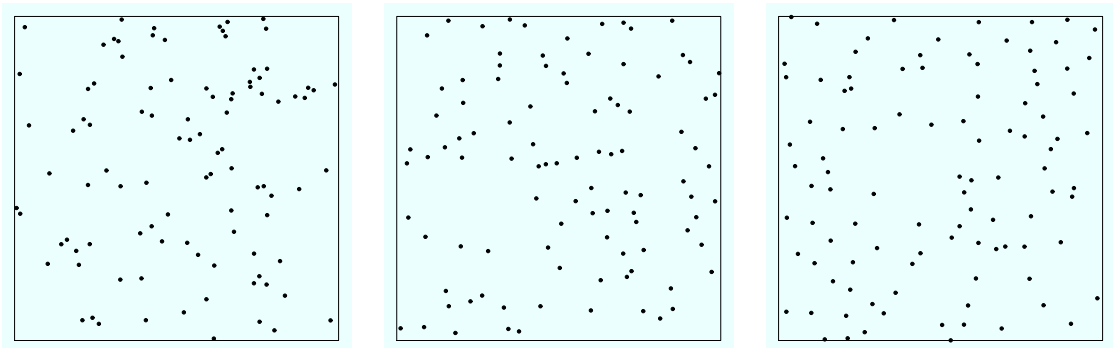

Figure 1: Realizations on $[0,1]^2$ of DPPs with kernel (2.8) where $\rho = 100$ and from left to right $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$.

For all methods considered in Table 1, the estimators seem consistent and the precision, in the sense of the mean squared errors, increases with the size of the observation window. From these results, the maximum likelihood method seems to be the best method in terms of quadradic loss, which agrees with the observations made in [19]. However, MCEs, especially the one based on g, seem to perform reasonably well. Moreover, their computation is faster than the maximum likelihood method and do not rely on an approximated spectral representation of C. For instance, with a regular laptop, the estimation of α for 500 realizations on $[0,3]^2$ took about 30 minutes for the MCEs based on K and g against more than 7 hours

	$[0,1]^2$			$[0,2]^2$			$[0,3]^2$		
	K	g	ML	K	g	ML	K	g	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: Mean squared errors of the MCE (2.7) when $J = K$, $J = g$, and the maximum likelihood method estimator (ML) as approximated in [19]. These values are estimated from 500 realizations of DPPs on $[0,1]^2$, $[0,2]^2$ and $[0,3]^3$ with kernel (2.8), $\rho = 100$ and $\alpha = 0.01,\ 0.03,\ \frac{1}{10\sqrt{\pi}}$. All entries are multiplied by 10^4 to make the table more compact.

by the maximum likelihood method. Finally, it seems that each estimator has an asymptotic Gaussian behaviour, as illustrated in Figure 2 where we have represented the histograms obtained from the estimations of $\alpha = 0.03$ over 500 realizations on $[0,1]^2$ as in Table 1. The remainder of this paper is dedicated to proving the asymptotic normality of the MCE (2.7) when $J = K$ or $J = g$ and $\mathbf{X}$ is a stationary DPP. The asymptotic properties of the maximum likelihood estimator remain an open problem. Note finally that a solution to improve the efficiency of the MCEs, still avoiding the computation of the likelihood, is to construct an optimal linear combination of the MCE based on K and the MCE based on g, see [20] for a general presentation of the procedure and [17] for an example in spatial statistics.

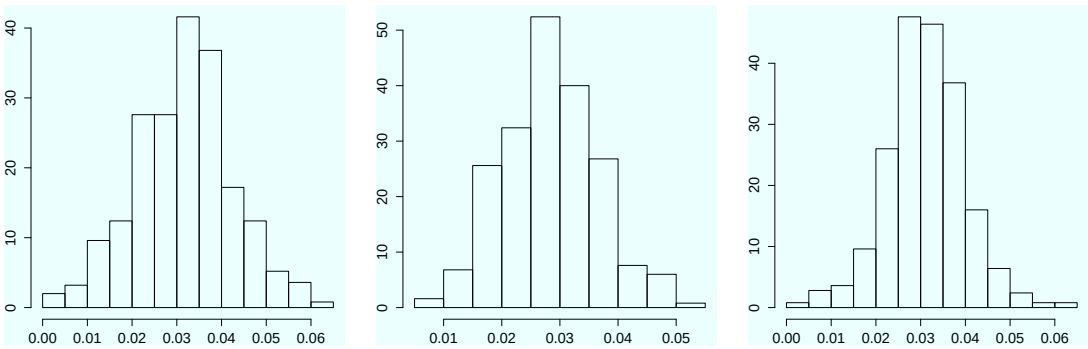

Figure 2: Histograms of the estimations of $\alpha = 0.03$ from 500 realizations of DPPs with kernel (2.8) on $[0,1]^2$. From left to right : MCE (2.7) based on K, MCE (2.7) based on g and maximum likelihood estimator.

3 Asymptotic properties of minimum contrast estimators based on K and g

3.1 Setting

In the next sections we study the asymptotic properties of (2.7) when $J = K$ and $J = g$, respectively. The asymptotic is to be understood in the following way. We assume to observe one realization of $\mathbf{X}$ on D_n and we let D_n to expand to $\mathbb{R}^d$ as detailed below. We denote by ∂D_n the boundary of D_n.

Definition 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*

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

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

Henceforth, we consider the estimator (2.7) under the setting of Section 2.2 where $\{D_n\}_{n\in\mathbb{N}}$ is a sequence of regular subsets of $\mathbb{R}^d$. Moreover, for any $\rho > 0$ and $\theta \in \Theta_\rho$, we assume that the correlation function associated to $C_{\rho,\theta}$, denoted by R_θ, does not depend on ρ but only on θ, i.e. $R_\theta = C_{\rho,\theta}/\rho$. Note that this is the case for all parametric families considered in [19] and [5], including the Whittle-Matèrn, the generalized Cauchy and the generalized Bessel families.

For $r > 0$, we denote by $\Theta_{\rho_0}^{\oplus r} := \Theta_{\rho_0} + \overline{B}(0, r)$ the r-dilation of Θ_{ρ_0}, where $\overline{B}(0, r)$ denotes the closed ball centred at 0 with radius r. Further, for all $x \in \mathbb{R}^d$, denote $R_\theta^{(1)}(x)$ and $R_\theta^{(2)}(x)$, the gradient, respectively the Hessian matrix, of $R_\theta(x)$ with respect to θ. We make the following assumptions. Specific additional hypotheses in the case $J = K$ and $J = g$ are described in the respective sections.

($\mathcal{H}1$) 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.

($\mathcal{H}2$) 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 for all $\theta \in \Theta_{\rho_0}^{\oplus\epsilon}$, $C_{\rho_0,\theta} \in L^2(\mathbb{R}^d)$ and $\mathcal{F}(C_{\rho_0,\theta}) \geq 0$.

($\mathcal{H}3$) 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 (2.7) is done over the random set $\Theta_{\hat{\rho}_n}$ in place of Θ_{ρ_0}. The two other assumptions deal with the regularity of the kernel with respect to the parameters.

3.2 MCE based on K

Since for any $\rho > 0$ and $\theta \in \Theta_\rho$ $R_\theta = C_{\rho,\theta}/\rho$ is assumed to not depend on ρ, the K-function (2.1) of $DPP(C_{\rho,\theta})$ does not depend on ρ. Consequently we denote it by $K(.,\theta)$. For all $t \geq 0$ and $n \in \mathbb{N}$, we consider the estimator of the K-function, see [23, Chapter 4],

$$\widehat{K}_n(t) := \frac{1}{\widehat{\rho}_n^2} \sum_{(x,y)\in\mathbf{X}^2}^{\neq} \mathbf{1}_{\{x\in D_n\}}\mathbf{1}_{\{y\in D_n^{\ominus t}\}} \frac{\mathbf{1}_{\{|x-y|\leq t\}}}{|D_n^{\ominus t}|} \tag{3.1}$$

where $\widehat{\rho}_n$ is as in (2.5) and for $t \geq 0$, $D_n^{\ominus t} := \{x \in D_n,\ B(x,t) \in D_n\}$.

For all $t \in [r_{min}, r_{max}]$, denote by $K^{(1)}(t,\theta)$ and $K^{(2)}(t,\theta)$ the gradient and the Hessian matrix of $K(t,\theta)$ 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 technical 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 states the strong consistency and the asymptotic normality of the MCE based on K for stationary DPPs. It is proved in Section 4.1.

Theorem 3.2. *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 (2.6) with $J = K$ and $\widehat{J}_n = \widehat{K}_n$. Assume that ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K1$)-($\mathcal{H}_K4$) hold. Then, the minimum contrast estimator $\widehat{\theta}_n$ defined by (2.7) 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

$$B_{\theta_0} := \int_{r_{min}}^{r_{max}} w(t)K(t,\theta_0)^{2c-2}K^{(1)}(t,\theta_0)K^{(1)}(t,\theta_0)^T dt \tag{3.2}$$

and

$$\Sigma_{\rho_0,\theta_0} = \int_{r_{min}}^{r_{max}} \int_{r_{min}}^{r_{max}} h_{\rho_0,\theta_0}(t_1,t_2) j_K(t_1) j_K(t_2) dt_1 dt_2$$

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

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

where $c^{red}_{[2]}, c^{red}_{[3]}$ *and* $c^{red}_{[4]}$ *are given with respect to* C_{ρ_0,θ_0} *in* (2.2)-(2.4).

Let us 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 [4].

3.3 MCE based on g

We assume in this section that all DPPs of the parametric family are isotropic, which is the usual practice when dealing with the pair correlation function. 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 that the pcf of $DPP(C_{\rho,\theta})$ writes

$$g(x,\theta) = 1 - \widetilde{R}_\theta(|x|)^2 =: \tilde{g}(|x|,\theta) \tag{3.3}$$

and does not depend on ρ. In the following, to alleviate the notation, we omit the symbol tilde and for all $\theta \in \Theta_\rho$, we consider that the domain of definition of $R_\theta(.)$ and $g(.,\theta)$ is $\mathbb{R}^+$. Moreover, by symmetry we extend this domain to $\mathbb{R}$. Denote, for all $d \ge 2$, the surface area of the d-dimensional unit ball,

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

For $n \in \mathbb{N}$ and $t > 0$, we consider the kernel estimator of g, see [23, Section 4.3.5],

$$\widehat{g_n}(t) := \frac{1}{\sigma_d t^{d-1} \widehat{\rho}_n^2} \sum_{(x,y)\in\mathbf{X}^2}^{\neq} \mathbf{1}_{\{x\in D_n,\ y\in D_n\}} \frac{1}{b_n |D_n \cap D_n^{x-y}|} k\left(\frac{t-|x-y|}{b_n}\right) \tag{3.4}$$

where for any $z \in \mathbb{R}^d$ $D_n^z := \{u, u+z \in D_n\}$, $\widehat{\rho}_n$ is as in (2.5) 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 $g^{(1)}(t,\theta)$ and $g^{(2)}(t,\theta)$ the gradient and the Hessian matrix of g with respect to θ. We consider the assumptions:

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

($\mathcal{H}_g2$) w is a positive and continuous function on $[r_{min}, r_{max}]$.

($\mathcal{H}_g3$) The kernel k is positive, symmetric and bounded with compact support included in $[-T, T]$ for a given $T > 0$. Further, $\int_{\mathbb{R}} k(x)dx = 1$.

($\mathcal{H}_g4$) $\{b_n\}_{n\in\mathbb{N}}$ is a positive sequence, $b_n \to 0$, $b_n|D_n| \to +\infty$ and $b_n^4|D_n| \to 0$.

($\mathcal{H}_g5$) There exists $\epsilon > 0$ such that for all $\theta \in \Theta_{\rho_0}^{\oplus\epsilon}$, $R_\theta(.)$ is of class $\mathcal{C}^2$ on $\mathbb{R} \setminus \{0\}$.

($\mathcal{H}_g6$) 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}, w, b_n and k. ($\mathcal{H}_g5$) is not restrictive and is satisfied by all parametric families considered in [19] and [5]. ($\mathcal{H}_g6$) is an identifiability assumption and as in the previous section, the main technical assumption is in fact ($\mathcal{H}_g7$). The proof of the following theorem is postponed to Section 4.2. Put

$$j_g(t) := w(t) g(t,\theta_0)^{2c-2} g^{(1)}(t,\theta_0), \quad t \in [r_{min}, r_{max}].$$

Theorem 3.3. *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 (2.6) 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 (2.7) 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

$$
\begin{aligned}
\Sigma_{\rho_0,\theta_0} = {} & 2\int_{\mathbb{R}^d} \mathbf{1}_{\{r_{min}\le|x|\le r_{max}\}} \frac{j_g(|x|)j_g(|x|)}{\sigma_d^2|x|^{2(d-1)}} \left(c_{[2]}^{red}(x)+\rho_0^2\right) dx \\
& + 4\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{r_{min}\le|x|,|y-x|\le r_{max}\}} \frac{j_g(|x|)j_g(|y-x|)}{\sigma_d^2|x|^{d-1}|y-x|^{d-1}} \left(c_{[3]}^{red}(x,y)+\rho_0 c_{[2]}^{red}(y)\right) dxdy \\
& + 4\rho_0\int_{\mathbb{R}^{2d}} \mathbf{1}_{\{r_{min}\le|x|,|y|\le r_{max}\}} \frac{j_g(|x|)j_g(|y|)}{\sigma_d^2|x|^{d-1}|y|^{d-1}} \left(2c_{[2]}^{red}(x)+\rho_0^2\right) dxdy \\
& + \int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y|)}{\sigma_d^2|x|^{d-1}|z-y|^{d-1}} c_{[4]}^{red}(x,y,z)\ dxdydz \\
& + 4\rho_0\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y|)}{\sigma_d^2|x|^{d-1}|z-y|^{d-1}} c_{[3]}^{red}(y,z)\ dxdydz \\
& + 2\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y+x|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y+x|)}{\sigma_d^2|x|^{d-1}|z-y+x|^{d-1}} c_{[2]}^{red}(y)c_{[2]}^{red}(z)\ dxdydz \\
& + 4\rho_0^2\int_{\mathbb{R}^{3d}} \mathbf{1}_{\{r_{min}\le|x|,|z-y|\le r_{max}\}} \frac{j_g(|x|)j_g(|z-y|)}{\sigma_d^2|x|^{d-1}|z-y|^{d-1}} c_{[2]}^{red}(y)\ dxdydz \\
& - 4\rho_0\left(\int_{r_{min}}^{r_{max}} g(t,\theta_0)j_g(t)dt\right) \\
& \qquad \int_{\mathbb{R}^{2d}} \mathbf{1}_{\{r_{min}\le|x|\le r_{max}\}} \frac{j_g(|x|)}{\sigma_d|x|^{d-1}} \left(c_{[3]}^{red}(x,y)+2\rho_0 c_{[2]}^{red}(y)\right) dxdy \\
& - 8\rho_0\left(\int_{r_{min}}^{r_{max}} g(t,\theta_0)j_g(t)dt\right)\int_{\mathbb{R}^d} \mathbf{1}_{\{r_{min}\le|x|\le r_{max}\}} \frac{j_g(|x|)}{\sigma_d|x|^{d-1}} \left(c_{[2]}^{red}(x)+\rho_0^2\right) dx \\
& + 4\rho_0^2\left(\int_{r_{min}}^{r_{max}} g(t,\theta_0)j_g(t)dt\right)^2 \left(\rho_0 - \int_{\mathbb{R}^d} C_{\rho_0,\theta_0}(x)^2 dx\right)
\end{aligned}
$$

where $c_{[2]}^{red}, c_{[3]}^{red}$ *and* $c_{[4]}^{red}$ *are given in* (2.2)-(2.4).

4 Proofs

4.1 Proof of Theorem 3.2

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 $n \ge N$, $\Theta_{\hat{\rho}_n} \subset \Theta_{\rho_0}^{\oplus\epsilon}$ 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 apply below the general Theorems 5.1-5.2 of the appendix to prove that the estimator $\widetilde{\theta}_n$ defined in (5.2) with $\Theta = \Theta_{\rho_0}^{\oplus\epsilon}$, $J = K$ and $\widehat{J}_n = \widehat{K}_n$ is consistent and asymptotically normal. 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 $n \ge N_r$, $\widetilde{\theta}_n \in B(\theta_0, r)$. From Lemma 5.4 in the appendix and $(\mathcal{H}1)$, we deduce 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 that $\widetilde{\theta}_n$ in (5.2) and $\widehat{\theta}_n$ in (2.7) coincide.

Let us now prove the strong consistency and asymptotic normality of $\widetilde{\theta}_n$ in (5.2) when $\Theta = \Theta_{\rho_0}^{\oplus\epsilon}$, $J = K$ and $\widehat{J}_n = \widehat{K}_n$. To that end, we verify all the assumptions of Theorems 5.1-5.2. The general setting in Section 3.1, Assumptions $(\mathcal{H}1)$ and $(\mathcal{H}_K1)$

imply directly ($\mathcal{A}1$)-($\mathcal{A}2$). For all $\theta \in \Theta$, we have

$$K(t,\theta) = \sigma_d t^d - \int_{x\in B(0,t)} R_\theta(x)^2 dx \tag{4.1}$$

where $\mathcal{F}(R_\theta) \geq 0$ by ($\mathcal{H}2$). Further, by [26, 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 (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 ($\mathcal{A}8$) 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 4.1, $\widehat{K}_n(t)$ 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,

$$K^{(1)}(t,\theta) = -2 \int_{x\in B(0,t)} R_\theta(x) R_\theta^{(1)}(x) dx. \tag{4.2}$$

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. Assumptions ($\mathcal{A}6$) and ($\mathcal{A}9$) are directly implied by ($\mathcal{H}_K3$) and ($\mathcal{H}_K4$), respectively. The assumption ($\mathcal{A}5$)$'$ is proved by Lemma 4.1 below, while Lemmas 4.2-4.3 are preliminary results for Lemma 4.4 which proves the remaining assumption ($\mathcal{TCL}$).

Lemma 4.1. *Let K be the Ripley's K-function of a DPP with kernel C verifying $\mathcal{K}(\rho_0)$ and $\widehat{K}_n$ the estimator given by* (3.1)*. Then, for all $r_{max} > r_{min} \geq 0$,*

$$\sup_{t\in[r_{min},r_{max}]} \left| \widehat{K}_n(t) - K(t) \right| \xrightarrow[n\to+\infty]{a.s.} 0,$$

Proof. Since a stationary DPP is ergodic by [28, Theorem 7], we have

$$\widehat{\rho}_n \xrightarrow[n\to+\infty]{a.s.} \rho_0 \tag{4.3}$$

and

$$\sup_{t\in[r_{min},r_{max}]}\left|\widehat{\rho}_n^2\widehat{K}_n(t)-\rho_0^2K(t)\right|\xrightarrow[n\to+\infty]{a.s.}0, \tag{4.4}$$

see for instance [12, Section 4.2.2]. Further, as K is an increasing function, we have

$$\begin{aligned}\widehat{\rho}_n^2&\sup_{t\in[r_{min},r_{max}]}\left|\widehat{K}_n(t)-K(t)\right|\\ &\leq\sup_{t\in[r_{min},r_{max}]}\left|\widehat{\rho}_n^2\widehat{K}_n(t)-\rho_0^2K(t)\right|+K(r_{max})\sup_{t\in[r_{min},r_{max}]}\left|\widehat{\rho}_n^2-\rho_0^2\right|.\end{aligned}$$

Hence, by (4.3)-(4.4) and the last equation, we have the convergence

$$\sup_{t\in[r_{min},r_{max}]}\left|\widehat{K}_n(t)-K(t)\right|\xrightarrow[n\to+\infty]{a.s.}0.$$

□

Lemma 4.2. *If ($\mathcal{H}_K1$)-($\mathcal{H}_K2$) and ($\mathcal{H}3$) hold, then for all $r_{max}>r_{min}\geq 0$,*

$$\int_{r_{min}}^{r_{max}}|j_K(t)|\,dt<+\infty.$$

Proof. By (4.2), we have

$$\int_{r_{min}}^{r_{max}}|j_K(t)|\,dt=2\int_{r_{min}}^{r_{max}}\left|w(t)K(t,\theta_0)^{2c-2}\int_{x\in B(0,t)}R_{\theta_0}(x)R_{\theta_0}^{(1)}(x)dx\right|dt. \tag{4.5}$$

By ($\mathcal{H}3$), the function defined for all $t\geq 0$ by

$$t\mapsto\int_{x\in B(0,t)}R_{\theta_0}(x)R_{\theta_0}^{(1)}(x)dx$$

is continuous so bounded on $[r_{min},r_{max}]$. As already noticed after (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 ($\mathcal{H}_K1$). Finally, if $r_{min}=0$, the integrability at 0 of the function $t\mapsto|j_K(t)|$ follows from ($\mathcal{H}_K2$). □

To shorten, define for all $n\in\mathbb{N}$ and $t\in[r_{min},r_{max}]$,

$$H_n(t):=\widehat{\rho}_n^2\widehat{K}_n(t)-2\rho_0K(t,\theta_0)\widehat{\rho}_n.$$

Lemma 4.3. *If ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K1$) hold, for all $s\in\mathbb{R}^d$, we have*

$$\lim_{n\to+\infty}|D_n|\mathrm{Var}\Big(\int_{r_{min}}^{r_{max}}H_n(t)s^Tj_K(t)dt\Big)=\int_{[r_{min},r_{max}]^2}h(t_1,t_2)s^Tj_K(t_1)s^Tj_K(t_2)dt_1dt_2$$

where h_{ρ_0,θ_0} is defined as in Theorem 3.2.

Proof. From (3.1), 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 $n \in \mathbb{N}$,

$$f_n(x,y) := \mathbf{1}_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} \frac{1}{|D_n^{\ominus t}|} \mathbf{1}_{\{y\in D_n^{\ominus t}\}} \mathbf{1}_{\{0<|x-y|\le t\}} s^T j_K(t) dt$$

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,

$$\begin{aligned}
&\mathrm{Var}\left(\int_{r_{min}}^{r_{max}} H_n(t) s^T j_K(t) dt\right) \\
&= \mathrm{Var}\left(\sum_{(x,y)\in\mathbf{X}^2}^{\neq} f_n(x,y)\right) + \mathrm{Var}\left(\sum_{x\in\mathbf{X}} h_n(x)\right) - 2\,\mathrm{Cov}\left(\sum_{(x,y)\in\mathbf{X}^2}^{\neq} f_n(x,y), \sum_{x\in\mathbf{X}} h_n(x)\right).
\end{aligned}$$

These terms are developed in Lemmas 7.1-7.3 of [4], whereby we deduce the limit by a long but straightforward calculus. □

Lemma 4.4. *If ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_K1$)-($\mathcal{H}_K2$) hold, then*

$$\sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t) dt \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(0, \Sigma_{\rho_0,\theta_0})$$

where Σ_{ρ_0,θ_0} is defined as in Theorem 3.2.

Proof. For all $n \in \mathbb{N}$, we have

$$\begin{aligned}
\rho_0^2 \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t) dt &= \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\rho_0^2 - \widehat{\rho}_n^2\right] \widehat{K}_n(t) j_K(t) dt \\
&\quad + \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\widehat{\rho}_n^2 \widehat{K}_n(t) - \rho_0^2 K(t,\theta_0)\right] j_K(t) dt. \quad (4.6)
\end{aligned}$$

Since $\mathbf{X}$ is ergodic by [28, Theorem 7], $\widehat{\rho}_n$ converges almost surely to ρ_0. Then, by Taylor expansion of the function $x \to x^2$ at ρ_0, we have almost surely

$$\left[\rho_0^2 - \widehat{\rho}_n^2\right] = 2\rho_0 \left[\rho_0 - \widehat{\rho}_n\right] + o\left(\rho_0 - \widehat{\rho}_n\right). \quad (4.7)$$

Moreover,

$$\begin{aligned}
&2\rho_0 \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\rho_0 - \widehat{\rho}_n\right] \widehat{K}_n(t) j_K(t) dt \\
&\qquad = 2\rho_0 \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\rho_0 - \widehat{\rho}_n\right] \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t) dt \\
&\qquad\qquad + 2\rho_0 \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\rho_0 - \widehat{\rho}_n\right] K(t,\theta_0) j_K(t) dt. \quad (4.8)
\end{aligned}$$

Using the notation

$$A_n = 2\rho_0\sqrt{|D_n|}\,[\rho_0 - \widehat{\rho}_n]\int_{r_{min}}^{r_{max}} \widehat{K}_n(t) j_K(t)dt,$$
$$B_n = 2\rho_0\sqrt{|D_n|}\,[\rho_0 - \widehat{\rho}_n]\int_{r_{min}}^{r_{max}} \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t)dt,$$
$$C_n = \sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left([\rho_0 - \widehat{\rho}_n]\,2\rho_0 K(t,\theta_0) + \left[\widehat{\rho}_n^2\widehat{K}_n(t) - \rho_0^2 K(t,\theta_0)\right]\right) j_K(t)dt,$$

we have by (4.6)-(4.8),

$$\rho_0^2\sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\widehat{K}_n(t) - K(t,\theta_0)\right] j_K(t)dt = B_n + C_n + o\left(A_n\right). \tag{4.9}$$

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 (4.9). By Lemma 4.1,

$$\sup_{t\in[r_{min},r_{max}]} \left|\widehat{K}_n(t) - K(t,\theta_0)\right| \xrightarrow[n\to+\infty]{a.s.} 0$$

so

$$\int_{r_{min}}^{r_{max}} \widehat{K}_n(t) j_K(t)dt \xrightarrow[n\to+\infty]{a.s.} \int_{r_{min}}^{r_{max}} K(t,\theta_0) j_K(t)dt. \tag{4.10}$$

Since $K(.,\theta_0)$ is continuous on $[r_{min}, r_{max}]$, $\int_{r_{min}}^{r_{max}} K(t,\theta_0) j_K(t)dt$ is finite by Lemma 4.2. Hence, by Corollary 5.6, (4.10) and Slutsky's theorem, we deduce that $B_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$ and $o(A_n) \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

As to the term C_n, notice that

$$C_n = \sqrt{|D_n|}\left(\int_{r_{min}}^{r_{max}} H_n(t) j_K(t)dt - \left[-\int_{r_{min}}^{r_{max}} \rho_0^2 K(t,\theta_0) j_K(t)dt\right]\right). \tag{4.11}$$

We prove the convergence in distribution of C_n by the Cramer-Wold device, see for instance [3, Theorem 29.4]. For all $t \in [r_{min}, r_{max}]$ and $s \in \mathbb{R}^p$, we have

$$s^T C_n = \sqrt{|D_n|}\left(\int_{r_{min}}^{r_{max}} H_n(t) s^T j_K(t)dt - \left[-\int_{r_{min}}^{r_{max}} \rho_0^2 K(t,\theta_0) s^T j_K(t)dt\right]\right).$$

By (3.1), we have

$$\int_{r_{min}}^{r_{max}} H_n(t) s^T j_K(t)dt = \sum_{(x,y)\in\mathbf{X}^2} f_{D_n}(x,y) \tag{4.12}$$

where

$$f_{D_n}(x,y) := \mathbf{1}_{\{x\in D_n\}}\int_{r_{min}}^{r_{max}} \left(\frac{\mathbf{1}_{\{y\in D_n^{\ominus t}\}}}{|D_n^{\ominus t}|}\mathbf{1}_{\{0<|x-y|\le t\}} - 2\rho_0\frac{K(t,\theta_0)}{|D_n|}\mathbf{1}_{\{x-y=0\}}\right) s^T j_K(t)dt.$$

Notice that for $t \in [r_{min}, r_{max}]$, $s^T j_K(t) \leq |j_K(t)||s|$ and $K(t, \theta_0) \leq K(r_{max}, \theta_0)$ so we have

$$\begin{aligned} &|f_{D_n}(x,y)| \\ &\leq \frac{|s|}{|D_n^{\ominus r_{max}}|} \mathbf{1}_{D_n}(x) \left(\mathbf{1}_{\{0<|x-y|\leq r_{max}\}} + \mathbf{1}_{\{x-y=0\}} 2\rho_0 K(r_{max}, \theta_0) \right) \int_{r_{min}}^{r_{max}} |j_K(t)| \, dt. \end{aligned} \tag{4.13}$$

The right-hand term in (4.13) is compactly supported and is bounded by Lemma 4.2. Moreover,

$$\begin{aligned} &\mathbb{E}\left(\int_{r_{min}}^{r_{max}} \left| H_n(t) s^T j_K(t) \right| dt \right) \\ &\qquad \leq |s| \left[\mathbb{E}\left(\left| \widehat{\rho}_n^2 \widehat{K}_n(t) \right| \right) + 2\rho_0 K(r_{max}, \theta_0) \mathbb{E}\left(|\widehat{\rho}_n| \right) \right] \int_{r_{min}}^{r_{max}} |j_K(t)| dt. \end{aligned}$$

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 [12, 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 4.2. Then, by Fubini's theorem, (4.12) 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 (4.12) and Lemma 4.3,

$$\lim_{n\to+\infty} Var\left(\sqrt{|D_n|} \sum_{(x,y)\in \mathbf{X}^2} f_{D_n}(x,y) \right) = s^T \Sigma_{\rho_0,\theta_0} s.$$

Therefore, by (4.11)-(4.13), the last two equations and Theorem 5.5, we have

$$s^T C_n \xrightarrow[n\to+\infty]{distr.} N(0, s^T \Sigma_{\rho_0,\theta_0} s).$$

which proves that $C_n \xrightarrow[n\to+\infty]{distr.} N(0, \Sigma_{\rho_0,\theta_0})$. □

4.2 Proof of Theorem 3.3

As in the proof of Theorem 3.2, we consider without loss of generality $\epsilon > 0$ 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 (5.2) 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 5.3, 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 5.1 and 5.2. The assumptions $(\mathcal{A}1)$, $(\mathcal{A}2)$, $(\mathcal{A}6)$, $(\mathcal{A}7)$ and $(\mathcal{A}9)$ are directly implied by $(\mathcal{H}1)$-$(\mathcal{H}3)$, $(\mathcal{H}_g1)$, $(\mathcal{H}_g2)$, $(\mathcal{H}_g6)$ and $(\mathcal{H}_g7)$. Moreover, $r_{min} > 0$ by $(\mathcal{H}_g1)$ so $(\mathcal{A}4)$ is directly implied by (3.4), $(\mathcal{H}_g3)$, $(\mathcal{H}_g4)$ and the ergodic theorem, see [24] or [12]. By $(\mathcal{H}2)$, $R_{\theta_0}(.)$ is continuous on $[r_{min}, r_{max}]$ so is g. By [26, 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 (3.3) 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 4.5 and 4.9, respectively while the other lemmas are auxiliary results.

Lemma 4.5. *If $(\mathcal{H}1)$-$(\mathcal{H}3)$, $(\mathcal{H}_g1)$ and $(\mathcal{H}_g3)$-$(\mathcal{H}_g4)$ hold then, for all $r_{max} > r_{min} > 0$, there exists a set A verifying $|[r_{min}, r_{max}] \setminus A| = 0$ such that*

$$\sup_{t\in A} |\widehat{g}_n(t) - g(t, \theta_0)| \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

Proof. From $(\mathcal{H}2)$-$(\mathcal{H}3)$ and $(\mathcal{H}_g3)$-$(\mathcal{H}_g4)$ we can use Proposition 4.5 in [4] that gives

$$\mathbb{E}\left[\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2 \widehat{g}_n(t) - \rho_0^2 g(t, \theta_0)\right)^2 dt\right] = \frac{2\rho_0^2}{b_n|D_n|} \int_{r_{min}}^{r_{max}} \frac{g(t, \theta_0)}{\sigma_d t^{d-1}} dt \int_{\mathbb{R}} k(x)^2 dx + O\left(\frac{1}{|D_n|}\right) + O(b_n^4). \quad (4.14)$$

By $(\mathcal{H}_g1)$, $(\mathcal{H}_g3)$ 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 $(\mathcal{H}_g4)$, the right-hand term in (4.14) tends to 0 as n tends to infinity. Moreover, the term inside the expectation in (4.14) is positive so there exists a set A as in Lemma 4.5 such that

$$\sup_{t\in A} \left|\widehat{\rho}_n^2 \widehat{g}_n(t) - \rho_0^2 g(t, \theta_0)\right| \xrightarrow[n\to+\infty]{\mathbb{P}} 0. \quad (4.15)$$

We have

$$\widehat{\rho}_n^2 \sup_{t\in A} |\widehat{g}_n(t) - g(t, \theta_0)| \leq \sup_{t\in A} \left|\widehat{\rho}_n^2 \widehat{g}_n(t) - \rho_0^2 g(t, \theta_0)\right| + \left(\sup_{t\in A} g(t, \theta_0)\right) \left|\widehat{\rho}_n^2 - \rho_0^2\right|.$$

By $(\mathcal{H}1)$-$(\mathcal{H}2)$, it follows from Corollary 5.6 that $\widehat{\rho}_n$ converges in probability to ρ_0. Further, by $(\mathcal{H}3)$ and (3.3), $g(., \theta_0)$ is bounded on $[r_{min}, r_{max}]$. Therefore, we have by (4.15) the convergence

$$\sup_{t\in A} |\widehat{g}_n(t) - g(t, \theta_0)| \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

□

Lemma 4.6. *If ($\mathcal{H}1$)-($\mathcal{H}3$), ($\mathcal{H}_g1$)-($\mathcal{H}_g2$) hold then $j_g(.)$ is continuous on $[r_{min}, r_{max}]$.*

Proof. By (3.3), we have for all $t \in [r_{min}, r_{max}]$

$$|j_g(t)| = 2\left|w(t)\left(1 - R_{\theta_0}(t)^2\right)^{2c-2} R_{\theta_0}(t)R^{(1)}_{\theta_0}(t)\right|.$$

By ($\mathcal{H}3$), $R_{\theta_0}(.)$ and $R^{(1)}_{\theta_0}(.)$ 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 3.3, for all $t > 0$, $|R_{\theta_0}(t)| < 1$. Thus by ($\mathcal{H}3$), the function $t \mapsto (1 - R_{\theta_0}(t)^2)^{2c-2}$ is well defined and continuous on $[r_{min}, r_{max}]$. Finally, by ($\mathcal{H}_g2$), w is continuous on $[r_{min}, r_{max}]$ so the lemma is proved. □

To abbreviate, we define for all $n \in \mathbb{N}$ and $t \in [r_{min}, r_{max}]$,

$$H^g_n(t) := \widehat{\rho}^2_n \widehat{g}_n(t) - 2\rho_0\widehat{\rho}_n g(t, \theta_0).$$

Lemma 4.7. *If ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_g1$)-($\mathcal{H}_g5$) hold, we have for all $s \in \mathbb{R}^d$,*

$$\lim_{n\to+\infty} |D_n|\mathrm{Var}\left(\int_{r_{min}}^{r_{max}} H^g_n(t)s^T j_g(t)dt\right) = s^T\Sigma_{\rho_0,\theta_0}s$$

with Σ_{ρ_0,θ_0} defined as in Theorem 3.3.

Proof. Similarly to the proof of Lemma 4.3, we have by (3.4),

$$\int_{r_{min}}^{r_{max}} H^g_n(t)s^T j_g(t)dt = \sum_{(x,y)\in\mathbf{X}^2} f_n(x, y) - \sum_{x\in\mathbf{X}} h_n(x)$$

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

$$h_n(x) = \frac{2\rho_0}{|D_n|}\mathbf{1}_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} g(t, \theta_0)s^T j_g(t)dt.$$

The result follows similarly as in the proof of Lemma 4.3 using Lemmas 7.1-7.3 in [4]. □

Lemma 4.8. *Assume that ($\mathcal{H}1$)-($\mathcal{H}3$) and ($\mathcal{H}_g1$)-($\mathcal{H}_g4$) hold. For a given $s \in \mathbb{R}^d$ and all $n \in \mathbb{N}$, let f_{D_n} be defined for any $(x, y) \in \mathbb{R}^{2d}$ by*

$$f_{D_n}(x, y) := \mathbf{1}_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} \left(\frac{k\left(\frac{t-|x-y|}{b_n}\right)\mathbf{1}_{\{|x-y|>0,y\in D_n\}}}{\sigma_d t^{d-1}b_n|D_n \cap D_n^{x-y}|} - \frac{2\rho_0 g(t, \theta_0)}{|D_n|}\mathbf{1}_{\{x-y=0\}}\right) s^T j_g(t)dt.$$

Then, there exists $M > 0$ such that for all $(x, y) \in \mathbb{R}^{2d}$,

$$|f_{D_n}(x, y)| \le \frac{|s|M\mathbf{1}_{\{x\in D_n\}}}{|D_n^{\ominus r_{max}+T}|}\left(\frac{1}{\sigma_d r_{min}^{d-1}}\mathbf{1}_{\{0<|x-y|\le r_{max}+T\}} + 2\rho_0||g||_\infty\mathbf{1}_{\{x-y=0\}}\right).$$

Proof. By $(\mathcal{H}_g 3)$, for any $t \in [r_{min}, r_{max}]$ and $(x,y) \in \mathbb{R}^{2d}$,

$$\left| k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{|y-x|>0,\, y\in D_n\}} \leq \left| k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{0<|y-x|<t+Tb_n\}}$$
$$\leq \left| k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{0<|y-x|<t+T\}}$$

whenever $b_n < 1$ which, by $(\mathcal{H}_g 4)$, we assume in the following without loss of generality. Thus, for any $t \in [r_{min}, r_{max}]$ and $(x,y) \in \mathbb{R}^{2d}$,

$$\left| k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{|y-x|>0,\, y\in D_n\}} \leq \left| k\left(\frac{t-|x-y|}{b_n}\right)\right| \mathbf{1}_{\{0<|y-x|<r_{max}+T\}}. \tag{4.16}$$

Further, by Lemma 4.6, j_g is bounded on $[r_{min}, r_{max}]$ by a constant M so by (4.16) and Lemma 6.3 in [4], we have

$$\left| 1_{\{x\in D_n\}} \int_{r_{min}}^{r_{max}} \frac{k\left(\frac{t-|x-y|}{b_n}\right) \mathbf{1}_{\{|x-y|>0, y\in D_n\}}}{\sigma_d t^{d-1} b_n |D_n \cap D_n^{x-y}|} s^T j_g(t) dt \right|$$
$$\leq \mathbf{1}_{\{x\in D_n\}} \frac{|s|M}{|D_n^{\ominus r_{max}+T}|} \frac{\mathbf{1}_{\{0<|x-y|\leq r_{max}+T\}}}{\sigma_d r_{min}^{d-1} b_n} \int_{r_{min}}^{r_{max}} \left| k\left(\frac{t-|x-y|}{b_n}\right)\right| dt.$$

Finally, the result follows by the last inequality, $(\mathcal{H}2)$ and $(\mathcal{H}_g 3)$. □

Lemma 4.9. *If $(\mathcal{H}1)$-$(\mathcal{H}3)$ and $(\mathcal{H}_g 1)$-$(\mathcal{H}_g 5)$ hold, then*

$$\sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\widehat{g}_n(t) - g(t,\theta_0)\right] j_g(t) dt \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(0, \Sigma_{\rho_0,\theta_0})$$

with Σ_{ρ_0,θ_0} defined as in Theorem 3.3.

Proof. The arguments of this proof are similar the the ones of the proof of Lemma 4.4. Notice that

$$\rho_0^2 \sqrt{|D_n|} \int_{r_{min}}^{r_{max}} \left[\widehat{g}_n(t) - g(t,\theta_0)\right] j_g(t) dt = \sqrt{|D_n|} \Big(\left[\rho_0^2 - \widehat{\rho}_n^2\right] \int_{r_{min}}^{r_{max}} \widehat{g}_n(t) j_g(t) dt$$
$$+ \int_{r_{min}}^{r_{max}} \left[\widehat{\rho}_n^2 \widehat{g}_n(t) - \mathbb{E}\left[\widehat{\rho}_n^2 \widehat{g}_n(t)\right]\right] j_g(t) dt + \int_{r_{min}}^{r_{max}} \left[\mathbb{E}\left[\widehat{\rho}_n^2 \widehat{g}_n(t)\right] - \rho_0^2 g(t,\theta_0)\right] j_g(t) dt \Big) \tag{4.17}$$

and

$$\sqrt{|D_n|} \left[\rho_0 - \widehat{\rho}_n\right] \int_{r_{min}}^{r_{max}} \widehat{g}_n(t) j_g(t) dt =$$
$$\sqrt{|D_n|} \left[\rho_0 - \widehat{\rho}_n\right] \int_{r_{min}}^{r_{max}} \left[\widehat{g}_n(t) - g(t,\theta_0)\right] j_g(t) dt + \sqrt{|D_n|} \left[\rho_0 - \widehat{\rho}_n\right] \int_{r_{min}}^{r_{max}} g(t,\theta_0) j_g(t) dt. \tag{4.18}$$

Denote

$$T_n = 2\rho_0\sqrt{|D_n|}\left[\rho_0 - \widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}} \widehat{g}_n(t) j_g(t)dt$$
$$U_n = 2\rho_0\sqrt{|D_n|}\left[\rho_0 - \widehat{\rho}_n\right]\int_{r_{min}}^{r_{max}} \left[\widehat{g}_n(t) - g(t,\theta_0)\right] j_g(t)dt$$
$$V_n = \sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\mathbb{E}\left[\widehat{\rho}_n^2\widehat{g}_n(t)\right] - \rho_0^2 g(t,\theta_0)\right] j_g(t)dt$$
$$W_n = \sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\widehat{\rho}_n^2\widehat{g}_n(t) - 2\rho_0\widehat{\rho}_n g(t,\theta_0) - \left(\mathbb{E}\left[\widehat{\rho}_n^2\widehat{g}_n(t)\right] - 2\rho_0^2 g(t,\theta_0)\right)\right] j_g(t)dt.$$

Using (4.7) in the proof of Lemma 4.4, (4.17) and (4.18), we get

$$\rho_0^2\sqrt{|D_n|}\int_{r_{min}}^{r_{max}} \left[\widehat{g}_n(t) - g(t,\theta_0)\right] j_g(t)dt = U_n + V_n + W_n + o\left(T_n\right). \tag{4.19}$$

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. From Corollary 5.6, Lemmas 4.5-4.6 and Slutsky's theorem, 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 4.5 that $\widehat{g}_n$ is uniformly bounded in probability on $[r_{min}, r_{max}]$, see [31, Prohorov's theorem]. Thus, by Corollary 5.6 and Lemma 4.6, $o(T_n) \xrightarrow[n\to+\infty]{\mathbb{P}} 0$. Further, under $(\mathcal{H}1)$-$(\mathcal{H}2)$ and $(\mathcal{H}_g1)$-$(\mathcal{H}_g5)$, we deduce from Lemma 6.2 in [4] that $\sup_{t\in[r_{min},r_{max}]}\left(\mathbb{E}\left[\widehat{\rho}_n^2\widehat{g}_n(t)\right] - \rho_0^2 g(t,\theta_0)\right) < \kappa b_n^2$ with $\kappa > 0$, which combined with $(\mathcal{H}_g4)$ and Lemma 4.6 proves that $V_n \xrightarrow[n\to+\infty]{\mathbb{P}} 0$.

We prove the convergence in distribution of W_n by the 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 4.6, 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

$$\mathbb{E}\left(\int_{r_{min}}^{r_{max}} \left|H_n^g(t)s^T j_g(t)\right| dt\right) \leq |s|M\mathbb{E}\int_{r_{min}}^{r_{max}} \left(\left|\widehat{\rho}_n^2\widehat{g}_n(t) - \rho_0^2 g(t,\theta_0)\right|\right) dt$$
$$+ |s|M\left[\mathbb{E}(|\rho_0 - \widehat{\rho}_n|) + \mathbb{E}(\widehat{\rho}_n)\right]\int_{r_{min}}^{r_{max}} |\rho_0 g(t,\theta_0)|dt. \tag{4.20}$$

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 (4.20), we have

$$\int_{r_{min}}^{r_{max}} \mathbb{E}\left|H_n^g(t)s^T j_g(t)\right| dt \leq |s|M(r_{max}-r_{min})^{\frac{1}{2}}\left(\mathbb{E}\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2\widehat{g}_n(t) - \rho_0^2 g(t,\theta_0)\right)^2 dt\right)^{\frac{1}{2}}$$
$$+ |s|M(r_{max} - r_{min})\rho_0||g||_\infty\left(\mathbb{E}(|\rho_0 - \widehat{\rho}_n|) + \mathbb{E}(\widehat{\rho}_n)\right). \tag{4.21}$$

By the same arguments as in the proof of Lemma 4.5, 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 $(\mathcal{H}_g 4)$, $\mathbb{E}\left(\int_{r_{min}}^{r_{max}} \left(\widehat{\rho}_n^2 \widehat{g}_n(t) - \rho_0^2 g(t,\theta_0)\right)^2 dt\right)$ tends to 0. Moreover, as noticed in [11], $\widehat{\rho}_n$ 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 (4.20)-(4.21), $\mathbb{E}\int_{r_{min}}^{r_{max}} \left|H_n^g(t) s^T j_g(t)\right| dt$ 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 (3.4), we have

$$X_n^s = \sum_{(x,y)\in\mathbf{X}^2} f_{D_n}(x,y), \tag{4.22}$$

where $f_{D_n}(x,y)$ is given in Lemma 4.8 and satisfies

$$|f_{D_n}(x,y)| \leq \frac{|s| M \mathbf{1}_{\{x\in D_n\}}}{|D_n^{\ominus r_{max}+T}|}\left(\frac{1}{\sigma_d r_{min}^{d-1}}\mathbf{1}_{\{0<|x-y|\leq r_{max}+T\}} + 2\rho_0 ||g||_\infty \mathbf{1}_{\{x-y=0\}}\right). \tag{4.23}$$

The right-hand term in (4.23) is bounded and compactly supported. Therefore, by Lemma 4.7 and Theorem 5.5, we have for all $s \in \mathbb{R}^p$

$$\sqrt{|D_n|}\left(X_n^s - \mathbb{E}(X_n^s)\right) \xrightarrow[n\to+\infty]{distr.} N(0, s^T \Sigma_{\rho_0,\theta_0} s),$$

which implies that $W_n \xrightarrow[n\to+\infty]{distr.} N(0, \Sigma_{\rho_0,\theta_0})$. □

5 Appendix

5.1 A general result for minimum contrast estimation

We present in this section two general theorems concerning the consistency and asymptotic normality of the estimator defined in (2.7). Contrary to the results in Sections 3.2-3.3, these theorems hold for an arbitrary stationary point process and an arbitrary statistic J, generalizing a study by [10]. The results of Sections 3.2-3.3 are in fact consequences in the particular case of a DPP and $J = K$ or $J = g$, which simplifies the general assumptions below.

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 $p \geq 1$. 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 $J^{(1)}(t, \theta)$ and $J^{(2)}(t, \theta)$ the gradient, respectively the Hessian matrix, of $J(t, \theta)$ with respect to θ. Define for all $\theta \in \Theta$,

$$B(\theta) := \int_{r_{min}}^{r_{max}} w(t) J(t, \theta)^{2c-2} J^{(1)}(t, \theta) J^{(1)}(t, \theta)^T dt, \tag{5.1}$$

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 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 $J(t, \theta) > 0$.

($\mathcal{A}4$) There exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\widehat{J}_n(.)$ and $\widehat{J}_n(.)^c$ are almost surely bounded on $[r_{min}, r_{max}]$.

($\mathcal{A}5$) 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 5.1. *Let* $\mathbf{X}$ *be a stationary point process with distribution ruled by a given* θ_0*, assumed to be an interior point of* Θ*. For all* $n \in \mathbb{N}$*, let* U_n *be defined as in* (2.6)*. Assume that*($\mathcal{A}1$)*-*($\mathcal{A}6$)* hold. Then, the minimum contrast estimator* $\widetilde{\theta}_n$ *defined by*

$$\widetilde{\theta}_n = \arg\min_{\theta\in\Theta} U_n(\theta) \tag{5.2}$$

exists almost surely, is consistent for θ_0 *and strongly consistent if* $(\mathcal{A}5)'$ *holds.*

Proof. For a sequence $\{\theta_m\}_{m\in\mathbb{N}}$ belonging to Θ, we have for all $n \in \mathbb{N}$,

$$|U_n(\theta_m) - U_n(\theta)| \leq \int_{r_{min}}^{r_{max}} |w(t)| \left(\left|2\widehat{J}_n(t)^c\right| |J(t,\theta_m)^c - J(t,\theta)^c| + \left|J(t,\theta_m)^{2c} - J(t,\theta)^{2c}\right|\right) dt. \tag{5.3}$$

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) 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 large enough, U_n is almost surely continuous so the almost sure existence of $\widetilde{\theta}_n$ follows by ($\mathcal{A}1$). Define for all $\theta \in \Theta$,

$$U_n^*(\theta) = U_n(\theta) - U_n(\theta_0). \tag{5.4}$$

By (2.6) and (5.4),

$$U_n^*(\theta) = 2\int_{r_{min}}^{r_{max}} w(t)\Big[\widehat{J}_n(t)^c - J(t,\theta_0)^c\Big]\Big[J(t,\theta_0)^c - J(t,\theta)^c\Big]dt + \int_{r_{min}}^{r_{max}} w(t)\Big[J(t,\theta_0)^c - J(t,\theta)^c\Big]^2 dt.$$

Note that from (5.4) $U_n^*(\widetilde{\theta}_n) \leq U_n^*(\theta_0) = 0$, so

$$\int_{r_{min}}^{r_{max}} w(t)\Big[J(t,\theta_0)^c - J(t,\widetilde{\theta}_n)^c\Big]^2 dt \leq 2\int_{r_{min}}^{r_{max}} w(t)\Big|\widehat{J}_n(t)^c - J(t,\theta_0)^c\Big|\Big|J(t,\theta_0)^c - J(t,\widetilde{\theta}_n)^c\Big| dt. \tag{5.5}$$

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 (5.5) tends in probability to 0. Hence, we have

$$\int_{r_{min}}^{r_{max}} w(t)\Big[J(t,\theta_0)^c - J(t,\widetilde{\theta}_n)^c\Big]^2 dt \xrightarrow[n\to+\infty]{\mathbb{P}} 0.$$

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 (5.5) that this last convergence is almost sure if $(\mathcal{A}5)'$ holds. □

Theorem 5.2. *Under the same setting as in Theorem 5.1, if in addition ($\mathcal{A}7$)-($\mathcal{A}9$) and ($\mathcal{TCL}$) hold true, then*

$$\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} \mathcal{N}\left(m, B(\theta_0)^{-1}\Sigma\left(B(\theta_0)^{-1}\right)^T\right)$$

where B is defined as in (5.1) *and Σ comes from ($\mathcal{TCL}$).*

Proof. Denote by 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 $s \in (0,1)$ and $\theta_j^* = \theta_0 + s(\widetilde{\theta}_n - \theta_0)$ such that

$$\partial_j U_n(\widetilde{\theta}_n) - \partial_j U_n(\theta_0) = \left(\partial^2_{ij} U_n(\theta_j^*)\right)_{i=1,\ldots,p}\left(\widetilde{\theta}_n - \theta_0\right).$$

To shorten, denote by $U_n^{(1)}$ the gradient of U_n and by $U_n^{(2)}(\theta_n^*)$ the matrix with entries $\partial^2_{ij}U_n(\theta_j^*)$. Since U_n is minimal at $\widetilde{\theta}_n$, $U_n^{(1)}(\widetilde{\theta}_n) = 0$ and the last equation becomes

$$\begin{aligned} U_n^{(2)}(\theta_n^*)(\widetilde{\theta}_n - \theta_0) &= -U_n^{(1)}(\theta_0) \\ &= 2c\int_{r_{min}}^{r_{max}} w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_0)^c\right] J(t,\theta_0)^{c-1} J^{(1)}(t,\theta_0)dt. \end{aligned} \tag{5.6}$$

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 $t \in A$,

$$\widehat{J}_n(t)^c - J(t,\theta_0)^c = cJ(t,\theta_0)^{c-1}\left(\widehat{J}_n(t) - J(t,\theta_0)\right) + o\left(\widehat{J}_n(t) - J(t,\theta_0)\right).$$

Therefore, by ($\mathcal{A}5$), (5.6) and the last equation,

$$\sqrt{|D_n|}\, U_n^{(2)}(\theta_n^*)(\widetilde{\theta}_n - \theta_0) = 2c^2 A_n(\theta_0) + o(A_n(\theta_0)) \tag{5.7}$$

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 (5.7),

$$\sqrt{|D_n|}\, U_n^{(2)}(\theta_n^*)(\widetilde{\theta}_n - \theta_0) \xrightarrow[n\to+\infty]{distr.} 2c^2 N(m,\Sigma). \tag{5.8}$$

Moreover, we have that

$$U_n^{(2)}(\theta_n^*) = 2c^2 B(\theta_n^*) - E_n \tag{5.9}$$

where B is as in (5.1) and

$$\begin{aligned} E_n := 2c(c-1)\int_{r_{min}}^{r_{max}} w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_n^*)^c\right] J(t,\theta_n^*)^{c-2} J^{(1)}(t,\theta_n^*) J^{(1)}(t,\theta_n^*)^T dt \\ + 2c\int_{r_{min}}^{r_{max}} w(t)\left[\widehat{J}_n(t)^c - J(t,\theta_n^*)^c\right] J(t,\theta_n^*)^{c-1} J^{(2)}(t,\theta_n^*)dt. \end{aligned}$$

By Theorem 5.1, $\widetilde{\theta}_n \xrightarrow[n\to+\infty]{\mathbb{P}} \theta_0$ 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 (5.9)

$$U_n^{(2)}(\theta_n^*)B(\theta_0)^{-1} \xrightarrow[n\to+\infty]{\mathbb{P}} 2c^2. \tag{5.10}$$

Since the group of invertible matrix is an open set, it follows from the last convergence that for n large enough, $U_n^{(2)}(\theta_n^*)$ is invertible so we can write

$$B(\theta_0)\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0) = B(\theta_0)\left[U_n^{(2)}(\theta_n^*)\right]^{-1} U_n^{(2)}(\theta_n^*)\sqrt{|D_n|}(\widetilde{\theta}_n - \theta_0).$$

By (5.8)-(5.10) and Slutsky's theorem, we get

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

and the conclusion of the theorem follows. □

5.2 Auxiliary results

The two following lemmas are of topological nature and useful for the proofs of Theorems 3.2-3.3.

Lemma 5.3. *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 $\delta \geq 0$, $B(y,\delta) \nsubseteq \Xi^{\oplus\delta}$.*

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

$$(y - p_\Xi(y)).(u - p_\Xi(y)) \leq 0. \tag{5.11}$$

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 (5.11) 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 $B(y,\delta) \nsubseteq \Xi^{\oplus\delta}$. □

Lemma 5.4. *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 $n \geq N$,*

$$B(x,r) \subset \Theta_n.$$

Proof. Since $B(x,r)$ belongs to the interior of Θ, there exists $\delta > 0$ such that

$$B(x, r+\delta) \subset \Theta. \tag{5.12}$$

Assume that the lemma is wrong, then for all $N \in \mathbb{N}$, there exists $n \geq N$ such that $B(x,r) \nsubseteq \Theta_n$. Denote y a point in $B(x,r)$ that does not belong to Θ_n. By Lemma 5.3, $B(y,\delta) \nsubseteq \Theta_n^{\oplus\delta}$. But by (5.12), $B(y,\delta) \subset \Theta$ so $\Theta \nsubseteq \Theta_n^{\oplus\delta}$ which contradicts the convergence of the sequence $\{\Theta_n\}_{n\in\mathbb{N}}$ to Θ. □

The following theorem appears in [14] in a slightly less general framework, see also [15], and is proved in [4] in its present form. It is used in the proofs of our main results, Theorems 3.2 and 3.3.

Theorem 5.5. *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 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 $(x_1,\ldots,x_p) \in \mathbb{R}^{dp}$,*

$$|f_{D_n}(x_1,\ldots,x_p)| \leq \frac{1}{|\widetilde{D}_n|}\mathbf{1}_{\{x_1\in D_n\}}F(x_2-x_1,\ldots,x_p-x_1). \tag{5.13}$$

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

$$\mathrm{Cum}_k\left(\sqrt{|D_n|}N_p\left(f_{D_n}\right)\right) = O\left(|D_n|^{1-\frac{k}{2}}\right). \tag{5.14}$$

Moreover, if there exists $\sigma > 0$ such that

$$\mathrm{Var}\left(\sqrt{|D_n|}N_p\left(f_{D_n}\right)\right) \xrightarrow[n\to+\infty]{} \sigma^2, \tag{5.15}$$

then we have the convergence

$$\sqrt{|D_n|}\left[N_p\left(f_{D_n}\right) - \mathbb{E}\left(N_p\left(f_{D_n}\right)\right)\right] \xrightarrow[n\to+\infty]{distr.} \mathcal{N}(0,\sigma^2) \tag{5.16}$$

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

As a corollary when $p = 1$, we retrieve a theorem from [29] giving the asymptotic normality of the estimator of the intensity of a DPP.

Corollary 5.6. *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 $\sigma^2 = \lim_{n\to+\infty} Var\left(\sqrt{|D_n|}\widehat{\rho}_n\right) = \rho - \int_{\mathbb{R}^d} C(x)^2 dx$.

References

[1] Baddeley, A., Rubak, E., and Turner, R. *Spatial Point Patterns: Methodology and Applications with R.* Chapman and Hall/CRC Press, London, 2015. In press.

[2] Baddeley, A., and Turner, R. spatstat: An R package for analyzing spatial point patterns. *Journal of Statistical Software 12*, 6 (2005), 1–42.

[3] Billingsley, P. *Probability and measure.* John Wiley & Sons, New York-Chichester-Brisbane, 1979. Wiley Series in Probability and Mathematical Statistics.

[4] Biscio, C. A. N., and Lavancier, F. Brillinger mixing of determinantal points processes and statistical applications. *arXiv:1507.06506* (2015).

[5] Biscio, C. A. N., and Lavancier, F. Quantifying repulsiveness of determinantal point processes. *To appear in Bernoulli* (2015).

[6] 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.

[7] 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.

[8] Deng, N., Zhou, W., and Haenggi, M. The Ginibre point process as a model for wireless networks with repulsion. *Wireless Communications, IEEE Transactions on 14*, 1 (2015), 107–121.

[9] Diggle, P. *Statistical Analysis of Spatial Point Patterns*, second ed. Hodder Arnold, London, 2003.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] Kulesza, A., and Taskar, B. Determinantal point processes for machine learning. *Foundations and Trends in Machine Learning 5* (2012), 123–286.

[17] Lavancier, F., and Møller, J. Modelling aggregation on the large scale and regularity on the small scale in spatial point pattern datasets. *Scandinavian Journal of Statistics* (2015). To appear.

[18] Lavancier, F., Møller, J., and Rubak, E. Determinantal point process models and statistical inference : Extended version. *arXiv:1205.4818v5* (2014).

[19] 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.

[20] Lavancier, F., and Rochet, P. A general procedure to combine estimators. *Computational Statistics and Data Analysis* (2015). To appear.

[21] Macchi, O. The coincidence approach to stochastic point processes. *Advances in Applied Probability 7* (1975), 83–122.

[22] Miyoshi, N., and Shirai, T. A cellular network model with ginibre configurated base stations. Technical report, Department of Mathematical and Computing Sciences Tokyo Institute of Technology, series B: Applied Mathematical Science, 2013.

[23] 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.

[24] Nguyen, X.-X., and Zessin, H. Ergodic theorems for spatial processes. *Z. Wahrsch. Verw. Gebiete 48*, 2 (1979), 133–158.

[25] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2015.

[26] Sasvári, Z. *Multivariate Characteristic and Correlation Functions*, vol. 50. Walter de Gruyter, 2013.

[27] 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.

[28] Soshnikov, A. Determinantal random point fields. *Russian Mathematical Surveys 55* (2000), 923–975.

[29] Soshnikov, A. Gaussian limit for determinantal random point fields. *The Annals of Probability 30*, 1 (2002), 171–187.

[30] Stein, E., and Weiss, G. *Introduction to Fourier Analysis on Euclidean Spaces (PMS-32)*, vol. 1. Princeton university press, 1971.

[31] van der Vaart, A. W. *Asymptotic Statistics*, vol. 3 of *Cambridge Series in Statistical and Probabilistic Mathematics.* Cambridge University Press, 1998.

[32] 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.