Electronic Journal of Statistics
Vol. 12 (2018) 427–460
ISSN: 1935-7524
https://doi.org/10.1214/18-EJS1397

Large and moderate deviations for kernel–type estimators of the mean density of Boolean models

Federico Camerlenghi[*]

Dipartimento di Economia, Metodi Quantitativi e Strategie d'Impresa
Università degli Studi di Milano–Bicocca
Piazza dell'Ateneo Nuovo 1, 20126 Milano, Italy
e-mail: federico.camerlenghi@unimib.it

and

Elena Villa[†]

Dipartimento di Matematica
Università degli Studi di Milano
Via Saldini 50, 20133 Milano, Italy
e-mail: elena.villa@unimi.it

Abstract: The mean density of a random closed set with integer Hausdorff dimension is a crucial notion in stochastic geometry, in fact it is a fundamental tool in a large variety of applied problems, such as image analysis, medicine, computer vision, etc. Hence the estimation of the mean density is a problem of interest both from a theoretical and computational standpoint. Nowadays different kinds of estimators are available in the literature, in particular here we focus on a kernel–type estimator, which may be considered as a generalization of the traditional kernel density estimator of random variables to the case of random closed sets. The aim of the present paper is to provide asymptotic properties of such an estimator in the context of Boolean models, which are a broad class of random closed sets. More precisely we are able to prove large and moderate deviation principles, which allow us to derive the strong consistency of the estimator of the mean density as well as asymptotic confidence intervals. Finally we underline the connection of our theoretical findings with classical literature concerning density estimation of random variables.

MSC 2010 subject classifications: Primary 62F12, 60F10, 60D05.
Keywords and phrases: Large deviations, moderate deviations, random closed sets, confidence intervals, stochastic geometry, Boolean models.

Received September 2017.

Contents

[*]Also affiliated to BIDSA, Bocconi University, 20136 Milano and to the Collegio Carlo Alberto, Piazza Arbarello 8, 10122 Torino, Italy.
[†]Corresponding author.

1. Introduction

The mean density of lower dimensional random closed sets, such as fiber processes and surfaces of full dimensional random sets, is an important quantity which arises in different scientific fields. As a consequence its evaluation and estimation have undergone a growing interest during the last decades [6, 19]. Recent areas of applications include pattern recognition and image analysis [40, 28], computer vision [42], medicine [1, 8, 15, 16, 17], material science [14]. We remind that, given a probability space $(\Omega, \mathscr{F}, \mathbb{P})$, a *random closed set* Θ in $\mathbb{R}^d$ is a measurable map

$$\Theta : (\Omega, \mathscr{F}) \longrightarrow (\mathbb{F}, \sigma_{\mathbb{F}}),$$

where $\mathbb{F}$ denotes the class of the closed subsets in $\mathbb{R}^d$, and $\sigma_{\mathbb{F}}$ is the σ-algebra generated by the so called *Fell topology*, or *hit-or-miss topology*, that is the topology generated by the set system

$$\{\mathscr{F}_G : G \in \mathscr{G}\} \cup \{\mathscr{F}^C : C \in \mathscr{C}\}$$

where $\mathscr{G}$ and $\mathscr{C}$ are the system of the open and compact subsets of $\mathbb{R}^d$, respectively (e.g., see [36]). We say that a random closed set $\Theta : (\Omega, \mathscr{F}) \to (\mathbb{F}, \sigma_{\mathbb{F}})$ satisfies a certain property (e.g., Θ has Hausdorff dimension n) if Θ satisfies that property $\mathbb{P}$-a.s.; throughout the paper we shall deal with countably $\mathscr{H}^n$-rectifiable random closed sets, having denoted by $\mathscr{H}^n$ the n-dimensional Hausdorff measure.

A random closed set Θ_n of locally finite n-dimensional Hausdorff measure $\mathscr{H}^n$ induces a random measure $\mu_{\Theta_n}(A) := \mathscr{H}^n(\Theta_n \cap A)$, $A \in \mathscr{B}_{\mathbb{R}^d}$, and the corresponding expected measure is defined as

$$\mathbb{E}[\mu_{\Theta_n}](A) := \mathbb{E}[\mathscr{H}^n(\Theta_n \cap A)], \quad A \in \mathscr{B}_{\mathbb{R}^d}, \tag{1}$$

where $\mathscr{B}_{\mathbb{R}^d}$ is the Borel σ-algebra of $\mathbb{R}^d$. (The important issue of the measurability of the random variable $\mu_{\Theta_n}(A)$ has been addressed in [5, 45].) Whenever the measure $\mathbb{E}[\mu_{\Theta_n}]$ is absolutely continuous with respect to the d-dimensional Hausdorff measure $\mathscr{H}^d$, its density (i.e. its Radon-Nikodym derivative) with respect to $\mathscr{H}^d$ is called *mean density* of Θ_n, and, according to notation in previous works (e.g., see [18, 20]), denoted by λ_{Θ_n}.

It is worth mentioning that, while the estimation of the mean density in stationary settings has been widely studied in the literature (see, e.g., [6, 23]), only recently the non-stationary case has been addressed, and, to the best of our knowledge, a general density estimation theory for random sets is still missing. The aim of the present paper is the investigation of this area. As a matter of fact, the problem of the local and global approximation of λ_{Θ_n} for non stationary random sets has been tackled by the authors in [2, 18, 19, 20, 44]. More specifically, given an i.i.d. random sample $\Theta_n^{(1)}, \dots, \Theta_n^{(N)}$ of size N for the random closed set Θ_n, the authors have provided two different kinds of estimators for the mean density of Θ_n: the so-called "Minkowski content"–based estimator, introduced in [43] through the notion of the *Minkowski content* of a set (see, e.g., [3]), and the so-called kernel-type estimator, introduced in [10] and denoted here $\widehat{\lambda}_{\Theta_n}^{\kappa,N}$ (for its precise definition see Eq. (6) bellow). We refer to [10] for a discussion on similarities and differences among them; we mention here that, even if the evaluation of $\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)$ is a non–trivial issue for very general random sets, it has been shown in [11] that it approaches the true value of $\lambda_{\Theta_n}(x)$ much faster than the "Minkowski content"–based estimator.

We point out that the importance of the estimator $\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)$ arises in the general theory of random sets, because it may be regarded as a generalization of the classical kernel density estimator of random variables to the case of random sets (see also Section 6); this is the reason why we shall refer to $\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)$ as "kernel-type" estimator (or briefly *kernel density estimator*), and why its investigation plays a pivotal role in the whole theory of random sets, providing a unifying approach to density estimation. While the asymptotic properties of the "Minkowski content"–based estimator, as well as asymptotic confidence intervals and central limit theorems, have been studied in [13], no analogous results are still available for the kernel–type estimator of the mean density. Hence the main aim of the present paper is the investigation of large and moderate deviation principles of $\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)$ for a large class of random closed sets, known as Boolean models, leaving to subsequent works extensions to more general classes. The analysis we will carry out is much in the spirit of [31, 35], who proved similar results for kernel estimators of random variables. Even if Boolean models do not cover all the variety of random sets, as stated in [4], they are usually considered basic random sets models in stochastic geometry. So the present paper may be seen as the first step in extending large and moderate deviation principles for kernel density estimators of random variables to the case of kernel-type estimators of the mean density of random sets. The theorems we are going to prove are interesting in their own right, in addition they provide tools to derive asymptotic normality and strong consistency of kernel-type estimators, which are useful to determine asymptotic confidence intervals, as well.

The paper is organized as follows. In Section 2, we depict the general framework of Boolean models that we want to handle in this paper; besides we briefly recall all the results on stochastic geometry and large deviation theory that are necessary to the aim of the present paper. Large and moderate deviation principles for the kernel–type estimator of the mean density are presented in Section 3, namely in Theorem 2 and Theorem 3, respectively. These theorems are the basic building blocks to derive statistical properties of such an estimator. Indeed we are able to prove its strong consistency and to derive asymptotic confidence intervals (see Section 4). Some noteworthy examples of Boolean models are discussed as well in Section 5. Finally Section 6 contains a discussion on relevant connections with the literature and paves the way for future developments of the present work. For the reader's convenience, the proofs of the main theorems, and some related technical lemmas, are deferred to Appendix A.

2. Preliminaries and notations

This section gathers some basics on stochastic geometry and large deviations, which are necessary to understand our main results. Clearly the treatment is not exhaustive here, thus throughout the paper we provide some interesting references for those readers who want to deepen the results we just recall.

2.1. Point processes, intensity measure and Boolean models

Roughly speaking a point process, denoted here by $\widetilde{\Phi}$, is a locally finite collection $\{\xi_i\}_{i\in\mathbb{N}}$ of random points; more formally $\widetilde{\Phi}$ is a random counting measure, that is a measurable map from a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ into the space of locally finite counting measures on $\mathbb{R}^d$. Throughout the paper we will deal with simple point processes, that is $\widetilde{\Phi}(\{x\}) \leq 1$ $\forall x \in \mathbb{R}^d$, $\mathbb{P}$-a.s.

The measure $\widetilde{\Lambda}(A) := \mathbb{E}[\widetilde{\Phi}(A)]$ on $\mathscr{B}_{\mathbb{R}^d}$ is called *intensity measure* of $\widetilde{\Phi}$; whenever it is absolutely continuous with respect to $\mathscr{H}^d$, its density is called *intensity* of $\widetilde{\Phi}$.

Marked point processes may be regarded as a generalization of point processes. They are collections of random points ξ_i in $\mathbb{R}^d$, each one associated with a mark K_i, which usually belongs to a complete and separable metric space (c.s.m.s.) $\mathbf{K}$. Hence the resulting collection of random points $\Phi = \{(\xi_i, K_i)\}_{i\in\mathbb{N}}$ is a point process on $\mathbb{R}^d \times \mathbf{K}$, with the property that the unmarked process $\{\widetilde{\Phi}(B) : B \in \mathscr{B}_{\mathbb{R}^d}\} := \{\Phi(B \times \mathbf{K}) : B \in \mathscr{B}_{\mathbb{R}^d}\}$ is a point process in $\mathbb{R}^d$. $\mathbf{K}$ is called *mark space*, while the random element K_i of $\mathbf{K}$ is the *mark associated to the point* ξ_i. Φ is said to be *stationary* if the distribution of $\{\xi_i + x, K_i\}_i$ is independent of $x \in \mathbb{R}^d$. The intensity measure of Φ, say Λ, is a σ-finite measure on $\mathscr{B}_{\mathbb{R}^d\times\mathbf{K}}$ defined as $\Lambda(B \times L) := \mathbb{E}[\Phi(B \times L)]$. A common assumption (e.g., see [33]) is that there exists a measurable function $f : \mathbb{R}^d \times \mathbf{K} \to \mathbb{R}_+$ and a probability measure Q on $\mathbf{K}$ such that $\Lambda(\mathrm{d}(x, K)) = f(x, K)\mathrm{d}xQ(\mathrm{d}K)$. We also recall that point processes can be considered on quite general metric spaces. In particular, a point process in $\mathscr{C}^d$, the class of compact subsets of $\mathbb{R}^d$, is called *particle process* (see [4] and references therein). It is well known that, by a *center map*, a particle process can be transformed into a marked point process Φ on $\mathbb{R}^d$ with marks in $\mathscr{C}^d$, by representing any compact set C as a pair (x, Z), where x may be interpreted as the "location" of C and $Z := C - x$ the "shape" (or "form") of C. In this case the marked point process $\Phi = \{(X_i, Z_i)\}$ is also called *germ-grain model*. Every random closed set Θ in $\mathbb{R}^d$ can be represented as a germ-grain model by means of a suitable marked point process $\Phi = \{X_i, Z_i\}$. In a large variety of applications the random sets Z_i are uniquely determined by a suitable random parameter $S \in \mathbf{K}$. Typical examples include: union of random balls, where $\mathbf{K} = \mathbb{R}_+$ and S is the radius of a ball centered at the origin; segment processes in $\mathbb{R}^2$ in which $\mathbf{K} = \mathbb{R}_+ \times [0, 2\pi]$ and $S = (L, \alpha)$ where L and α are the random length and orientation of the segment attached to the origin, respectively.

In order to be consistent with the notation used in previous works (e.g., [44, 10]), we shall consider random sets Θ_n described by marked point processes Φ in $\mathbb{R}^d$ with marks in a suitable mark space $\mathbf{K}$ so that $Z = Z(S)$ is a random set containing the origin:

$$\Theta_n(\omega) = \bigcup_{(\xi,s)\in\Phi(\omega)} \xi + Z(s), \qquad \omega \in \Omega. \tag{2}$$

Whenever Φ is a marked Poisson point process, Θ_n is said to be a *Boolean model*. Since we are going to consider here Boolean models, we also recall that a marked Poisson point process in $\mathbb{R}^d$ with marks in $\mathbf{K}$ may be seen as a Poisson point process on $\mathbb{R}^d \times \mathbf{K}$ with intensity measure Λ if $\Lambda(\cdot \times \mathbf{K})$ is continuous and locally bounded.

For an exhaustive treatment of point processes we refer to [24, 25], and to [34] for an elegant presentation of Poisson processes. Further, we mention [36, 37, 38, 39] for a unified theory on germ-grain models.

2.2. Basics on large and moderate deviations

The theory of large deviations is concerned with the asymptotic estimation of probabilities of rare events, by giving an asymptotic computation of small probabilities in exponential scale. Assume that $(\mathbb{X}, \mathscr{X})$ is a Polish space equipped with its Borel σ–algebra. The large deviation principle characterizes the asymptotic behavior of a family of probability measures $\{\mu_N\}_{N\geq 1}$ on $(\mathbb{X}, \mathscr{X})$ as N goes to infinity in terms of a *rate function.* A rate function is a map $J^* : \mathbb{X} \to [0, +\infty)$ lower semicontinuous, i.e. the level sets $\{x : J^*(x) \leq \alpha\}$ are closed for every $\alpha \geq 0$; J^* is said to be a *good rate function* if the level sets are compact. The set $\{x : J^*(x) < +\infty\}$ amounts to be the domain of J^*. Let v_N be a velocity, namely a function such that $v_N \to +\infty$ as $N \to \infty$.

A family of probability measure $\{\mu_N\}_{N\geq 1}$ is said to satisfy a *Large Deviation Principle* (LDP) with rate function J^* and velocity v_N if and only if for any

$A \in \mathscr{X}$

$$
\begin{aligned}
-\inf_{x\in \mathring{A}} J^*(x) &\le \liminf_{N\to+\infty} \frac{1}{v_N}\log(\mu_N(A)) \\
&\le \limsup_{N\to+\infty} \frac{1}{v_N}\log(\mu_N(A)) \le -\inf_{x\in\overline{A}} J^*(x),
\end{aligned} \tag{3}
$$

where $\mathring{A}$ and $\overline{A}$ are the interior and the closure af A, respectively, and with the convention that the infimum over the empty set equals $+\infty$. We say that a sequence of random variables satisfies the LDP when the sequence of measures induced by these variables satisfies the LDP.

The Gärtner-Ellis Theorem [26, Theorem 2.3.6] is the main tool to prove large deviations results. For our purposes, we consider the case $\mathbb{X} = \mathbb{R}^m$, with $m \ge 1$, and $\mathscr{X} = \mathscr{B}_{\mathbb{R}^m}$. In what follows $\boldsymbol{a}\cdot\boldsymbol{b} := \sum_{j=1}^m a_j b_j$ denotes the scalar product between two generic vectors $\boldsymbol{a} = (a_1,\dots,a_m)$ and $\boldsymbol{b} = (b_1,\dots,b_m)$ of $\mathbb{R}^m$. We also remind that a convex function $f : \mathbb{R}^m \to (-\infty,\infty]$ is said to be *essentially smooth* (see e.g. Definition 2.3.5 in [26]) if the interior $\mathring{\mathscr{D}}_f$ of $\mathscr{D}_f := \{\boldsymbol{\gamma} \in \mathbb{R}^m : f(\boldsymbol{\gamma}) < \infty\}$ is non-empty, f is differentiable throughout $\mathring{\mathscr{D}}_f$, and f is steep, i.e. $\lim_{h\to\infty} \|\nabla f(\boldsymbol{\gamma}_h)\| = \infty$ whenever $\{\boldsymbol{\gamma}_h : h \ge 1\}$ is a sequence in $\mathring{\mathscr{D}}_f$ converging to some boundary point of $\mathring{\mathscr{D}}_f$.

Theorem 1 (Gärtner-Ellis Theorem)**.** *Let $\{\boldsymbol{Z}_N\}_{N\ge1}$ be a sequence of $\mathbb{R}^m$–valued random variables and define the function $J : \mathbb{R}^m \to [-\infty,+\infty]$ by*

$$
J(\boldsymbol{\gamma}) := \lim_{N\to+\infty} \frac{1}{v_N}\log \mathbb{E}[e^{v_N \boldsymbol{\gamma}\cdot\boldsymbol{Z}_N}],
$$

whenever the limit exists. Assume that the origin $\mathbf{0} = (0,\dots,0) \in \mathbb{R}^m$ belongs to the interior $\mathscr{D}_J := \{\boldsymbol{\gamma} \in \mathbb{R}^m : J(\boldsymbol{\gamma}) < \infty\}$. Then, if J is essentially smooth and lower semi-continuous, then $\{\boldsymbol{Z}_N : N \ge 1\}$ satisfies the LDP with speed v_N and good rate function J^* defined by $J^*(\boldsymbol{y}) := \sup_{\boldsymbol{\gamma}\in\mathbb{R}^m}\{\boldsymbol{\gamma}\cdot\boldsymbol{y} - J(\boldsymbol{\gamma})\}$.

Formally a *Moderate Deviation Principle* (MDP) is nothing else but a LDP. We speak of moderate deviation when, for a suitable class of sequences of positive numbers $\{a_N\}$ such that

$$
\lim_{N\to\infty} a_N = 0 \quad\text{and}\quad \lim_{N\to\infty} w_N a_N = \infty, \tag{4}
$$

where $w_N \to \infty$ as $N \to \infty$, a LDP holds for suitable centered random variables with speed $v_N = 1/a_N$ and the same quadratic rate which does not depend on the choice of $\{a_N\}$. Moderate deviations may be employed to obtain the weak convergence to a centered Normal distribution whose variance is determined by a suitable application of the Gärtner-Ellis Theorem (e.g., see also [9]). This will be clarified in Section 4 where we shall apply LDP and MDP to show that, for every $x \in \mathbb{R}^d$, the kernel estimator $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ of $\lambda_{\Theta_n}(x)$ is strongly consistent and asymptotically Normal, respectively.

2.3. Notations and assumptions

To fix the notation, b_n denotes the volume of the unit ball in $\mathbb{R}^n$, and $B_r(x)$ is the closed ball centered at $x \in \mathbb{R}^d$ with radius $r > 0$. For any $A \subset \mathbb{R}^d$ and $r > 0$, its Minkowski enlargement at size $r > 0$ is denoted by $A_{\oplus r} := \{x \in \mathbb{R}^d : \operatorname{dist}(x, A) \leq r\}$, where $\operatorname{dist}(x, A)$ stands for the euclidean distance of the point x to the set A. The diameter of A will be denoted by $\operatorname{diam}(A)$. It is worth to recall that a compact set $A \in \mathscr{B}_{\mathbb{R}^d}$ is said to be *countably $\mathscr{H}^n$-rectifiable* if there exist countably many n-dimensional Lipschitz graphs $\Gamma_i \subset \mathbb{R}^d$ such that $A \setminus \cup_i \Gamma_i$ is $\mathscr{H}^n$-negligible. For further definitions and properties on rectifiable sets refer to [3, 29, 30].

In the sequel, we will say that Θ_n satisfies a certain property if such a property is satisfied for $\mathbb{P}$-almost every $\omega \in \Omega$; in particular Θ_n will be a Boolean model driven by a Poisson point process Φ in $\mathbb{R}^d \times \mathbf{K}$ with intensity measure $\Lambda(\mathrm{d}(x, s)) = f(x, s)\mathrm{d}x Q(\mathrm{d}s)$, satisfying the following assumptions:

($\overline{A1}$) for any $s \in \mathbf{K}$, $Z(s)$ is a countably $\mathscr{H}^n$-rectifiable and compact subset of $\mathbb{R}^d$, such that there exists a closed set $\Xi(s) \supseteq Z(s)$ such that $\int_{\mathbf{K}} \mathscr{H}^n(\Xi(s))Q(\mathrm{d}s) < \infty$ and

$$\gamma r^n \leq \mathscr{H}^n(\Xi(s) \cap B_r(x)) \leq \widetilde{\gamma} r^n \quad \forall x \in Z(s),\ r \in (0, 1)$$

for some $\gamma, \widetilde{\gamma} > 0$ independent of s;

(A2) for any $s \in \mathbf{K}$, $\mathscr{H}^n(\operatorname{disc}(f(\cdot, s))) = 0$, where $\operatorname{disc}(f(\cdot, s))$ contains the discontinuity points of $f(\cdot, s)$, and $f(\cdot, s)$ is locally bounded such that for any compact $K \subset \mathbb{R}^d$

$$\sup_{x \in K_{\oplus \operatorname{diam}(Z(s))}} f(x, s) \leq \widetilde{\xi}_K(s)$$

for some $\widetilde{\xi}_K(s)$ with $\int_{\mathbf{K}} \mathscr{H}^n(\Xi(s))\widetilde{\xi}_K(s)Q(\mathrm{d}s) < \infty$.

These assumptions may seem to be a little bit technical at a first glance, but they are natural hypotheses fulfilled by a wide class of germ-grain models, and their meaning has been extensively discussed in [10, 44]; indeed, for the reader's convenience, we use here the same notation ($\overline{A1}$) and (A2) introduced in [10] and in [44], respectively. We also recall that the assumption ($\overline{A1}$) guarantees (see Remark 4 and Proposition 5 in [44]) that the measure $\mathbb{E}[\mu_{\Theta_n}]$ defined in (1) is locally bounded and absolutely continuous with density

$$\lambda_{\Theta_n}(x) = \int_{\mathbf{K}} \int_{x - Z(s)} f(y, s) \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s). \tag{5}$$

In order to define the kernel density estimator of the mean density, we remind that a multivariate *kernel* is a probability density function $\kappa : \mathbb{R}^d \to \mathbb{R}$ which is radially symmetric.

Summing up, throughout the paper, unless otherwise specified, we suppose the validity of:

Assumptions.

- Θ_n is a Boolean model with integer Hausdorff dimension $n < d$ as in (2), satisfying ($\overline{A1}$) and (A2).
- $\left\{\Theta_n^{(i)}\right\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random closed sets as Θ_n.
- κ is a continuous kernel with compact support $\mathrm{supp}(\kappa) \subset B_R(0)$, and such that $\kappa(x) \le M$, for all $x \in \mathbb{R}^d$ and for some $M > 0$.

The *kernel-type estimator* $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ of the mean density $\lambda_{\Theta_n}(x)$ at a point $x \in \mathbb{R}^d$ is defined as follows [10]:

$$\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) := \frac{1}{N}\sum_{i=1}^{N}\kappa_{r_N} * \mathscr{H}^n_{|_{\Theta_n^{(i)}}}(x) = \frac{1}{Nr_N^d}\sum_{i=1}^{N}\int_{\Theta_n^{(i)}}\kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(dy), \qquad (6)$$

where $*$ stands for the usual convolution product, while $\kappa_{r_N} := \kappa(x/r_N)/r_N^d$ is the scaled kernel.

It can be shown (see [10, Corollary 7]) that if the *bandwidth* r_N is such that

$$\lim_{N\to\infty} r_N = 0 \quad \text{and} \quad \lim_{N\to\infty} N r_N^{d-n} = +\infty,$$

then $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ is weakly consistent and asymptotically unbiased for $\mathscr{H}^d$–a.e. $x \in \mathbb{R}^d$.

The notion of *approximate tangent space* shall appear in the expression for the rate function both in the LDP and in the MDP stated in Theorem 2 and Theorem 3, respectively. Such a notion is borrowed from geometric measure theory and it is recalled below, for the reader's convenience. Denoted by $\mathbf{G}_n$ the set of unoriented n-dimensional subspaces of $\mathbb{R}^d$, and by $C_c(\mathbb{R}^d;\mathbb{R})$ the space of all the real valued continuous functions with compact support in $\mathbb{R}^d$, we remind that a $\mathscr{H}^n$-rectifiable compact set $A \subset \mathbb{R}^d$ admits approximate tangent space $\pi_x A \in \mathbf{G}_n$ at $x \in A$ if

$$\lim_{r\to 0}\int_{(A-x)/r}\phi(y)\mathscr{H}^n(\mathrm{d}y) = \int_{\pi_x A}\phi(y)\mathscr{H}^n(\mathrm{d}y) \quad \forall \phi \in C_c(\mathbb{R}^d;\mathbb{R}). \qquad (7)$$

By Theorem 2.83 and Proposition 1.62 in [3], $\pi_x A$ exists for $\mathscr{H}^n$-a.e. $x \in A$; moreover, (7) holds for any bounded Borel measurable function $\phi : \mathbb{R}^d \to \mathbb{R}$ with compact support such that $\mathscr{H}^n_{|\pi_x A}(\mathrm{disc}(\phi)) = 0$. For the sake of simplicity, we have assumed that κ is continuous: this allows us to directly apply Eq. (7) in the sequel. We refer to [10, Remark 9] for a more detailed discussion on the non-continuous case.

3. Large and moderate deviations for the kernel-type estimator

In this section we state large and moderate deviation principles for the kernel density estimator defined in (6), by deferring their proof to the Appendix. Such results will be useful to derive statistical properties and confidence intervals for the involved estimator, as we will see in Section 4.

Theorem 2 (LDP)**.** *Let Θ_n and κ be as in the Assumptions. Then the sequence of kernel estimators $\left\{\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)\right\}_{N\geq 1}$ satisfies a LDP with speed $v_N = Nr_N^{d-n}$ and good rate function*

$$J^*_x(y) = \sup_{t\in\mathbb{R}}\Big\{ty - \int_{\boldsymbol{K}}\int_{x-Z(s)}\int_{\mathbb{R}^d}\kappa(w)f(y,s) \times \frac{\exp\left\{t\int_{\pi_y(x-Z(s))}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)\right\}-1}{\int_{\pi_y(x-Z(s))}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)}\mathrm{d}w\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s)\Big\}, \tag{8}$$

where $\pi_y(x-Z(s)) \in \boldsymbol{G}_n$ is the approximate tangent space to $x-Z(s)$ at $y \in x-Z(s)$.

Theorem 3 (MDP)**.** *Let Θ_n and κ be as in the Assumptions, and let $\{b_N\}_{N\geq 1}$ be a sequence of positive real numbers such that*

$$\lim_{N\to+\infty}\frac{b_N}{\sqrt{Nr_N^{d-n}}} = +\infty \quad \textit{and} \quad \lim_{N\to+\infty}\frac{b_N}{Nr_N^{d-n}} = 0. \tag{9}$$

Then the sequence of estimators $\left\{Nr_N^{d-n}/b_N(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)])\right\}_{N\geq 1}$ satisfies a LDP with speed function $v_N := b_N^2/Nr_N^{d-n}$ and good rate function

$$J^*_x(y) := \frac{y^2}{2C_{Var}(x)},$$

where $C_{Var}(x)$ is the quantity so defined

$$C_{Var}(x) := \int_{\mathbf{K}}\int_{\mathbb{R}^d}\int_{x-Z(s)}\int_{\pi_y(x-Z(s))}\kappa(z) \times \kappa(z+w)f(y,s)\mathscr{H}^n(\mathrm{d}w)\mathscr{H}^n(\mathrm{d}y)\mathrm{d}zQ(\mathrm{d}s). \tag{10}$$

4. Statistical properties and confidence intervals

In the previous section we stated large and moderate deviation principles for the kernel estimator of the mean densities of random closed sets; these results allow to derive useful statistical properties for such an estimator. Indeed, proceeding along the same lines of [12, Remark 2], we can show how an estimate of the rate of convergence of $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ to $\lambda_{\Theta_n}(x)$ follows as a byproduct of Theorem 2 and that an immediate application of the Borel-Cantelli Lemma leads to a strong consistency result:

Proposition 4 (Convergence rate)**.** *Let Θ_n and κ be as in the Assumptions, and let $C_\delta := \{y \in \mathbb{R} : |y - \lambda_{\Theta_n}(x)| \geq \delta\}$, with $\delta > 0$. Denoted by $\Gamma^*_\delta := \inf_{y\in C_\delta} J^*_x(y)$, where $J^*_x(y)$ has been defined in Theorem 2, we have that for any $0 < \eta < \Gamma^*_\delta$ there exists N_0 such that*

$$\mathbb{P}\Big(\Big|\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \lambda_{\Theta_n}(x)\Big| \geq \delta\Big) \leq \exp\Big(-Nr_N^{d-n}(\Gamma^*_\delta - \eta)\Big) \qquad \forall N \geq N_0.$$

Proof. It is known that when we can apply the Gärtner-Ellis Theorem (see Theorem 1), the rate function $J^*(\boldsymbol{y})$ uniquely vanishes at $\boldsymbol{y} = \boldsymbol{y}_0$, where $\boldsymbol{y}_0 := \nabla J(\mathbf{0})$. Denoted for any $\delta > 0$

$$C_\delta := \{\boldsymbol{y} \in \mathbb{R}^m : \|\boldsymbol{y} - \boldsymbol{y}_0\| \geq \delta\},$$

we have that $\inf_{\boldsymbol{y}\in C_\delta} J^*(\boldsymbol{y}) > 0$, since J^* is non-negative and uniquely vanishes at $\boldsymbol{y}_0$. Therefore, as a consequence of the large deviation upper bound in (3) for the closed set C_δ, we have

$$\limsup_{N\to\infty} \frac{1}{v_N} \log\Big(\mathbb{P}(\boldsymbol{Z}_N \in C_\delta)\Big) \leq -\inf_{\boldsymbol{y}\in C_\delta} J^*(\boldsymbol{y}). \tag{11}$$

By virtue of Theorem 2, the previous bound holds true for $\boldsymbol{Z}_N = \widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$, and $v_n = Nr_N^{d-n}$. Besides, using equations (5) and (29), it can be easily seen that in our setup $y_0 := J'_x(0) = \lambda_{\Theta_n}(x)$. Hence, in view of these remarks and (11), one concludes that for all η such that $0 < \eta < \Gamma^*_\delta$, there exists N_0 such that

$$P\Big(|\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \lambda_{\Theta_n}(x)| \geq \delta\Big) \leq \exp\Big(-Nr_N^{d-n}(\Gamma^*_\delta - \eta)\Big)$$

for all $N \geq N_0$. □

Corollary 5 (Strong consistency)**.** *Let Θ_n and κ be as in the Assumptions, with $r_N \to 0$ such that $Nr_N^{d-n}/N^\alpha \to C$ for some $C, \alpha > 0$ as $N \to \infty$.*

Then the kernel estimator $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ of $\lambda_{\Theta_n}(x)$ is strongly consistent for every $x \in \mathbb{R}^d$, i.e.

$$\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) \overset{a.s.}{\to} \lambda_{\Theta_n}(x), \quad \text{as } N \to \infty.$$

Proof. Let $H := (\Gamma^*_\delta - \eta)$, with Γ^*_δ defined as in Proposition 4 and $\eta \in (0, \Gamma^*_\delta)$. Then H is a positive quantity independent of N, and observe that $\sum_{N\geq 1} \exp\Big(-Nr_N^{d-n}H\Big) < \infty$, since $Nr_N^{d-n} \sim N^\alpha$ for some $\alpha > 0$. Thus the result follows by Proposition 4 and a standard application of the Borel–Cantelli lemma. □

At the end of Section 2.2, we mentioned that the term *moderate deviation* is used when for a sequence $\{a_N\}$ of positive numbers satisfying the conditions in (4), a LDP holds for suitable centered random variables with speed $v_N = 1/a_N$. If we choose $w_N = Nr_N^{d-n}$, we may observe that by Theorem 3 we are in the case $a_N = Nr_N^{d-n}/b_N^2$, with b_N satisfying the conditions in (9).

Moreover we also mention that the case $a_N = 1/w_N$ (so here $b_N = Nr_N^{d-n}$) and $a_N = 1$ (so here $b_N = \sqrt{Nr_N^{d-n}}$) should correspond to the convergence to zero and to the weak convergence to a centered normal distribution, respectively, of the associated centered random variables (here $\Big\{\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)]\Big\}_{N\geq 1}$ and $\Big\{\sqrt{Nr_N^{d-n}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)])\Big\}_{N\geq 1}$, respectively). This is in accordance with the corollary above and with the proposition below.

Proposition 6 (Asymptotic Normality)**.** *Let Θ_n and κ be as in the Assumptions. Then the sequence $\left\{\sqrt{Nr_N^{d-n}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)])\right\}_{N\geq 1}$ converges weakly, as $N \to +\infty$, to the normal distribution* $\mathrm{N}(0, C_{Var}(x))$*, where $C_{Var}(x)$ is the quantity defined in* (10)*.*

Proof. One can proceed as in the proof of Theorem 3 with $b_N = \sqrt{Nr_N^{d-n}}$, noticing that the proof is still valid, even if the first condition in (9) is violated. As a consequence one is able to show that

$$\lim_{N\to+\infty} \mathbb{E}\Big[\exp\Big\{t\sqrt{Nr_N^{d-n}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)])\Big\}\Big] = e^{t^2 C_{Var}(x)/2}$$

which is tantamount to saying that $\left\{\sqrt{Nr_N^{d-n}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)])\right\}_{N\geq 1}$ converges weakly to the normal distribution $\mathrm{N}(0, C_{Var}(x))$, as $N \to +\infty$. □

We conclude the investigation of the statistical properties related to $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ providing asymptotic confidence intervals for $\lambda_{\Theta_n}(x)$, relying on Proposition 6. In order to do this we have to choose a specific bandwidth r_N, which is assumed to be the optimal bandwidth determined in [10]. Here we recall some useful results in this direction.

We remind that the best choice for r_N should be the one which minimizes the *mean square error* (MSE), given by

$$MSE(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)) := \mathbb{E}[(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x) - \lambda_{\Theta_n}(x))^2] = Bias^2(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)) + Var(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)).$$

The minimization of the MSE is a quite challenging problem, which cannot be solved even in the simplest case of kernel density estimators of random variables. Hence one should look for an r_N which minimizes the *asymptotic mean square error* (AMSE). For Θ_n and κ as in the Assumptions, the following asymptotic approximation of the variance may be deduced by the proof of Theorem 8 in [10]:

$$Var(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)) = \frac{C_{Var}(x)}{Nr_N^{d-n}} + o\Big(\frac{1}{Nr_N^{d-n}}\Big), \quad \text{for } \mathscr{H}^d\text{–a.e. } x \in \mathbb{R}^d, \tag{12}$$

where $C_{Var}(x)$ is the quantity defined in (10). For what concernes the asymptotic approximation of the bias, further differentiability assumptions on f are required. To fix the notation (the same used in [10] for the reader's convenience), in the sequel $\alpha := (\alpha_1, ..., \alpha_d)$ will denote a multi-index of $\mathbb{N}_0^d$; we will further define

$$\begin{aligned} |\alpha| &:= \alpha_1 + \cdots + \alpha_d \\ \alpha! &:= \alpha_1! \cdots \alpha_d! \\ y^\alpha &:= y_1^{\alpha_1} \cdots y_d^{\alpha_d} \\ D_y^\alpha f(y,s) &:= \frac{\partial^{|\alpha|} f(y,s)}{\partial y_1^{\alpha_1} \cdots \partial y_d^{\alpha_d}}; \end{aligned}$$

besides, for all $s \in \mathbf{K}$, we will put

$$\mathscr{D}^{(\alpha)}(s) := \operatorname{disc}(D_y^\alpha f(y,s)), \qquad \mathscr{D}(s) := \operatorname{disc}(f(\cdot,s)).$$

For now on we assume that $f(\cdot,s)$ is at least twice differentiable, and that the following assumption is fulfilled for any $|\alpha| = 2$:

($A2$bis) for any $s \in \mathbf{K}$, $\mathscr{H}^n(\mathscr{D}^{(\alpha)}(s)) = 0$ and $D_y^\alpha f(y,s)$ is locally bounded such that for any compact $C \subset \mathbb{R}^d$

$$\sup_{y \in C_{\oplus \operatorname{diam} Z(s)}} |D_y^\alpha f(y,s)| \leq \widetilde{\xi}_C^{(\alpha)}(s)$$

for some $\widetilde{\xi}_C^{(\alpha)}(s)$ with

$$\int_{\mathbf{K}} \mathscr{H}^n(\Xi(s))\widetilde{\xi}_C^{(\alpha)}(s)Q(\mathrm{d}s) < \infty.$$

An asymptotic approximation of the bias has been proved in [10, Theorem 8]:

$$Bias(\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)) = C_{Bias}(x)r_N^2 + o(r_N^2), \quad \text{for } \mathscr{H}^d\text{-a.e. } x \in \mathbb{R}^d, \tag{13}$$

where

$$C_{Bias}(x) := \sum_{|\alpha|=2} \frac{1}{\alpha!} \int_{\mathbb{R}^d} \kappa(z) z^\alpha \mathrm{d}z \int_{\mathbf{K}} \int_{x-Z(s)} D_y^\alpha f(y,s) \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s).$$

From (12) and (13) one gets the following asymptotic expansion of the MSE for $\mathscr{H}^d$–a.e. $x \in \mathbb{R}^d$

$$AMSE(\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)) = C_{Bias}^2(x) r_N^4 + \frac{1}{N r_N^{d-n}} C_{Var}(x),$$

from which, for N sufficiently large, the optimal bandwidth $r_N^{o,AMSE}(x)$ amounts to be (see [10, Eq. (17)])

$$r_N^{o,AMSE}(x) := \arg\min_{r_N} AMSE(\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)) = \sqrt[4+d-n]{\frac{(d-n)C_{Var}(x)}{4NC_{Bias}^2(x)}}, \tag{14}$$

$\mathscr{H}^d$–a.e. $x \in \mathbb{R}^d$, provided that $C_{Bias}(x) \neq 0$. (For a discussion on the case $C_{Bias}(x) = 0$ we refer to [10].)

Proposition 7. *Let Θ_n and κ be as in the Assumptions, and such that* ($A2$bis) *is fulfilled. If r_N is the asymptotic optimal bandwidth $r_N^{o,AMSE}$ in* (14)*, then*

$$\sqrt{\frac{N r_N^{d-n}}{C_{Var}(x)}} (\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x) - \lambda_{\Theta_n}(x)) \xrightarrow{d} Z \quad \text{as } N \to \infty,$$

where $Z \sim \mathrm{N}(\sqrt{(d-n)/2}, 1)$, for $\mathscr{H}^d$–a.e. $x \in \mathbb{R}^d$.

Proof. First of all note that

$$\begin{aligned}\sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\lambda_{\Theta_n}(x)) &= \sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)]) \\ &\quad+\sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}(\mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)]-\lambda_{\Theta_n}(x))\end{aligned} \tag{15}$$

and the first term in (15) converges weakly to the standard normal distribution as $N\to+\infty$, by Proposition 6. Let us notice now that the non–random term

$$\begin{aligned}\sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}(\mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)]-\lambda_{\Theta_n}(x)) &= \sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}Bias(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)) \\ &\overset{(13)}{=} \frac{C_{Bias}(x)}{\sqrt{C_{Var}(x)}}\sqrt{Nr_N^{d-n+4}}+o\Big(\sqrt{Nr_N^{d-n+4}}\Big) \overset{(14)}{=} \frac{\sqrt{d-n}}{2}+o(1),\end{aligned}$$

as $N\to+\infty$, for $\mathscr{H}^d$–a.e. $x\in\mathbb{R}^d$, if $r_N=r_N^{o,AMSE}$, and so the assertion. □

Corollary 8 (Asymptotic confidence interval)**.** *Under the assumptions of Proposition 7, if $r_N\equiv r_N^{o,AMSE}$, then an asymptotic confidence interval for $\lambda_{\Theta_n}(x)$ of level α is*

$$\Big[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\sqrt{\frac{C_{Var}(x)}{Nr_N^{d-n}}}\Big(J_\alpha+\frac{\sqrt{d-n}}{2}\Big),\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)+\sqrt{\frac{C_{Var}(x)}{Nr_N^{d-n}}}\Big(J_\alpha-\frac{\sqrt{d-n}}{2}\Big)\Big]$$

for $\mathscr{H}^d$–a.e. $x\in\mathbb{R}^d$, where J_α is such that $\mathbb{P}(-J_\alpha\le Z\le J_\alpha)=1-\alpha$ with $Z\sim\mathrm{N}(0,1)$.

Proof. Thanks to Proposition 7 we can state that

$$\sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\lambda_{\Theta_n}(x))-\frac{\sqrt{d-n}}{2}\sim AN(0,1),$$

hence

$$\begin{aligned}1-\alpha &\simeq \mathbb{P}\Big(-J_\alpha\le\sqrt{\frac{Nr_N^{d-n}}{C_{Var}(x)}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\lambda_{\Theta_n}(x))-\frac{\sqrt{d-n}}{2}\le J_\alpha\Big) \\ &= \mathbb{P}\Big(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\sqrt{\frac{C_{Var}(x)}{Nr_N^{d-n}}}\Big(J_\alpha+\frac{\sqrt{d-n}}{2}\Big)\le\lambda_{\Theta_n}(x) \\ &\qquad\le\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)+\sqrt{\frac{C_{Var}(x)}{Nr_N^{d-n}}}\Big(J_\alpha-\frac{\sqrt{d-n}}{2}\Big)\Big).\end{aligned}$$ □

The asymptotic confidence intervals we have derived in the present section are based on the assumption $C_{Bias}(x)\neq 0$, this does not happen for stationary Boolean models. However in such a situation the kernel–estimator is unbiased (see [10]) and Proposition 6 gives immediately the following:

Proposition 9 (Stationary case)**.** *Let Θ_n and κ be as in the Assumptions; furthermore suppose that Θ_n is a stationary Boolean model with constant mean density $\lambda_{\Theta_n} = f\mathbb{E}_Q[\mathscr{H}^n(Z)]$, where f is the intensity of the underlying Poisson point process. Then the sequence $\left\{\sqrt{Nr_N^{d-n}}(\widehat{\lambda}^{\kappa,N}_{\Theta_n} - \lambda_{\Theta_n})\right\}_{N\geq 1}$ converges weakly, as $N \to +\infty$, to the normal distribution $\mathrm{N}(0, C_{Var})$, being*

$$C_{Var} := f\int_{\mathbf{K}}\int_{\mathbb{R}^d}\int_{\{-Z(s)\}}\int_{\pi_y(-Z(s))}\kappa(z)\kappa(z+w)\mathscr{H}^n(\mathrm{d}w)\mathscr{H}^n(\mathrm{d}y)\mathrm{d}zQ(\mathrm{d}s).$$

The previous Proposition is the basic building block to determine asymptotic confidence intervals for stationary Boolean models as well. Indeed, proceeding along the same lines as in the proof of Corollary 8, an asymptotic confidence interval for λ_{Θ_n} of level α is

$$\left[\widehat{\lambda}^{\kappa,N}_{\Theta_n} - \sqrt{\frac{C_{Var}}{Nr_N^{d-n}}}J_\alpha, \widehat{\lambda}^{\kappa,N}_{\Theta_n} + \sqrt{\frac{C_{Var}(x)}{Nr_N^{d-n}}}J_\alpha\right]$$

where J_α is such that $\mathbb{P}(-J_\alpha \leq Z \leq J_\alpha) = 1-\alpha$ with $Z \sim \mathrm{N}(0,1)$. Note that here r_N can be any bandwidth.

5. Noteworthy examples

Here we discuss some relevant examples of Boolean models, in particular a Boolean segment process, the Poisson point process and the Matérn cluster process.

5.1. A Boolean segment process

As simple example of applicability of the previous results we discuss the Boolean segment process already introduced in [10].

Let $n = 1$ and assume that Θ_1 is an inhomogeneous Boolean model of segments in $\mathbb{R}^2$ with random length L and uniform orientation, so that the mark space is $\mathbf{K} = \mathbb{R}^+ \times [0, 2\pi]$. For all $s = (\ell, \alpha) \in \mathbf{K}$, let $Z(s) := \{(u, v) \in \mathbb{R}^2 : u = \tau cos\alpha, v = \tau sin\alpha, \tau \in [0, \ell]\}$ be the segment with length $\ell \in \mathbb{R}^+$, and orientation $\alpha \in [0, 2\pi]$. Denoted by $P_L(\mathrm{d}\ell)$ the probability law of the random length L, we assume that $\mathbb{E}[L^3] < +\infty$. Finally the segment process Θ_1 is driven by the marked Poisson process Φ in $\mathbb{R}^2 \times \mathbf{K}$ having intensity measure $\Lambda(\mathrm{d}(y, s)) = f(y)\mathrm{d}yQ(\mathrm{d}s)$, where $f(y) = f(y_1, y_2) = y_1^2 + y_2^2$ and $Q(\mathrm{d}s) = \frac{1}{2\pi}\mathrm{d}\alpha P_L(\mathrm{d}\ell)$. We are going to consider the kernel $k(z) = \mathbb{1}_{B_1(0)}(z)/\pi$, which is not continuous, anyway the theory developed here apply for this kernel thanks to [10, Remark 9]. More precisely, $\widehat{\lambda}^{\kappa,N}_{\Theta_1}(x)$ is given here by

$$\widehat{\lambda}^{\kappa,N}_{\Theta_1}(x) = \frac{1}{N\pi r_N^2}\sum_{i=1}^{N}\mathcal{H}^1(\Theta_1^{(i)} \cap B_{r_N}(x)),$$

whereas in [10] it is shown that

$$C_{Var}(x) = \frac{16}{3\pi^2}\lambda_{\Theta_1}(x), \quad \text{for any } x \in \mathbb{R}^2,$$

and the asymptotic optimal bandwidth is

$$r_N^{o,AMSE} = \sqrt[5]{\frac{16(\mathbb{E}[L](x_1^2+x_2^2)+\mathbb{E}[L^3]/3)}{3\pi^2 N(\mathbb{E}[L])^2}}.$$

Hence Proposition 7 and Corollary 8 now apply with the previous specifications of $\widehat{\lambda}^{\kappa,N}_{\Theta_1}(x)$, r_N and $C_{Var}(x)$.

5.2. Poisson point processes

Let Ψ be a Poisson point process in $\mathbb{R}^d$ with a continuous intensity λ_Ψ. We recall that Ψ may be seen as a particular Boolean model with Hausdorff dimension $n = 0$ and mean density λ_Ψ, by choosing $\mathbf{K} = \mathbb{R}^d$ as mark space, $Z(s) = s \in \mathbb{R}^d$ as trivial typical grain, and $\Lambda(\mathrm{d}(y,s)) := \lambda_\Psi(y)\mathrm{d}y\delta_0(s)\mathrm{d}s$. As expected, observe that

$$(5) = \int_{\mathbf{K}}\int_{x-Z(s)} \lambda_\Psi(y)\mathscr{H}^0(\mathrm{d}y)\delta_0(s)\mathrm{d}s = \int_{\mathbf{K}} \lambda_\Psi(x-s)\delta_0(s)\mathrm{d}s = \lambda_\Psi(x).$$

Let $\left\{\Psi^i\right\}_{i\in\mathbb{N}}$ be a sequence of i.i.d. point processes as Ψ, and κ as in the Assumptions. Then, by noticing that Assumptions $(\overline{A1})$ and (A2) are trivially fulfilled, all the previous results for $\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)$ specialize now for the sequence of kernel estimators $\left\{\widehat{\lambda}^{\kappa,N}_{\Psi}(x)\right\}_{N\geq 1}$ of $\lambda_\Psi(x)$ defined by

$$\widehat{\lambda}^{\kappa,N}_{\Psi}(x) := \frac{1}{Nr_N^d}\sum_{i=1}^{N}\sum_{x_j\in\Psi^i}\kappa\Big(\frac{x-x_j}{r_N}\Big).$$

In particular, by observing that $\pi_y(x - Z(s)) = \{0\}$, we get

$$\int_{\pi_y(x-Z(s))} \kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta) = \int_{\{0\}} \kappa(\theta+w)\mathscr{H}^0(\mathrm{d}\theta) = \kappa(w),$$

and

$$\begin{aligned} C_{Var}(x) &\overset{(10)}{=} \int_{\mathbb{R}}\int_{\mathbb{R}^d}\int_{x-s}\int_{\{0\}} \kappa(z)\kappa(z+w)\lambda_\Psi(y)\mathscr{H}^0(\mathrm{d}w)\mathscr{H}^0(\mathrm{d}y)\mathrm{d}z\delta_0(s)\mathrm{d}s \\ &= \lambda_\Psi(x)\int_{\mathbb{R}^d}\kappa^2(z)\mathrm{d}z. \end{aligned}$$

Hence we can specialize large and moderate deviation principles for $\widehat{\lambda}^{\kappa,N}_{\Psi}(x)$ by a direct application of Theorem 2 and Theorem 3, respectively:

(LDP) the sequence of kernel estimators $\left\{\widehat{\lambda}_{\Psi}^{\kappa,N}(x)\right\}_{N\geq 1}$ satisfies a LDP with speed $v_N = Nr_N^d$ and good rate function

$$J_x^*(y) = \sup_{t\in\mathbb{R}}\left\{ty - \lambda_\Psi(x)\int_{\mathbb{R}^d}(e^{t\kappa(w)} - 1)\mathrm{d}w\right\}; \tag{16}$$

(MDP) assume that $\left\{b_N\right\}_{N\geq 1}$ is a sequence of positive real numbers satisfying

$$\lim_{N\to+\infty}\frac{b_N}{\sqrt{Nr_N^d}} = +\infty \quad \text{and} \quad \lim_{N\to+\infty}\frac{b_N}{Nr_N^d} = 0;$$

then the sequence $\left\{(Nr_N^d/b_N)(\widehat{\lambda}_{\Psi}^{\kappa,N}(x) - \mathbb{E}[\widehat{\lambda}_{\Psi}^{\kappa,N}(x)])\right\}_{N\geq 1}$ satisfies a LDP with speed $v_N = b_N^2/Nr_N^d$ and good rate function

$$J_x^*(y) = \frac{y^2}{2||\kappa||_2^2\lambda_\Psi(x)}. \tag{17}$$

Finally, as a direct consequence of Proposition 6 it follows that the sequence $\left\{\sqrt{Nr_N^d}(\widehat{\lambda}_{\Psi}^{\kappa,N}(x) - \mathbb{E}[\widehat{\lambda}_{\Psi}^{\kappa,N}(x)])\right\}_{N\geq 1}$ converges weakly, as $N \to +\infty$, to the normal distribution $\mathrm{N}(0, ||\kappa||_2^2\lambda_\Psi(x))$.

5.3. Matérn cluster processes

Clustering is a fundamental operation on point processes, well-known in stochastic geometry, and it allows to construct new point processes (see [23] for a more exhaustive treatment). The clustering operation consists in replacing each point x of a given point process Φ_p, called *parent point process*, by a cluster N_x of points, called *daughter points*. Each cluster N_x is itself a point process, and it is assumed to have only a finite mean number of points. The resulting point process given by the union of all the clusters N_x is said to be a *cluster point process*. Let us assume that the parent point process Φ_p is a homogeneous Poisson point process with intensity λ_p, and the clusters N_x are of the form $N_{x_i} = N_i + x_i$ for each $x_i \in \Phi_p$, where the sequence $\{N_i\}_i$ is independent of Φ_p, and independent and identically distributed as N_0 (the *representative cluster*, centered at 0). Assuming that the number of points of N_0 is distributed according to a Poisson random variable with parameter n_c, and that these points are independently and uniformly distributed in the ball $B_R(0)$, where R is a further parameter of the model, then the resulting cluster point process

$$\Phi = \bigcup_{x_i\in\Phi_p} x_i + N_i$$

is called *Matérn cluster process*. It follows that Φ has constant intensity $\lambda_\Phi = \lambda_p n_c$, and may be regarded as a Boolean model Θ_0 with dimension $n = 0$, underlying Poisson point process Φ_p, and typical grain $Z_0 := N_0$ given by a Poisson point process restricted to $B_R(0)$ whose intensity equals $\lambda_{N_0}(x) = \frac{n_c}{b_d R^d}\mathbb{1}_{B_R(0)}(x)$. The resulting Boolean model $\Theta_0 \equiv \Phi$ is driven by a marked Poisson point process in $\mathbb{R}^d \times \mathscr{S}$ having intensity measure $\Lambda(\mathrm{d}(\xi,\eta)) = \lambda_p \mathrm{d}\xi Q(\mathrm{d}\eta)$, where the mark space coincides with $\mathbf{K} := \mathscr{S}$ the space of all sequences of points in $\mathbb{R}^d$ and Q is the probability distribution of N_0.

Note that all the assumptions ($\overline{A1}$) and (A2) are trivially fulfilled; as a consequence, all the previous results on $\widehat{\lambda}^{\kappa,N}_{\Theta_0}(x) \equiv \widehat{\lambda}^{\kappa,N}_{\Phi}(x)$ hold in such a context. A LDP follows from Theorem 2, more specifically one can observe that the general expression for J_x^* appearing in the statement of that theorem simplifies in the context of Matérn cluster processes, indeed

$$\int_{\mathbf{K}}\int_{x-Z(s)}\int_{\mathbb{R}^d} \frac{\exp\left\{t\int_{\pi_y(x-Z(s))}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)\right\}-1}{\int_{\pi_y(x-Z(s))}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)} \times \kappa(w) f(y,s)\mathrm{d}w\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s)$$

now equals

$$\begin{aligned}
&\int_{\mathscr{S}}\int_{x-\eta}\int_{\mathbb{R}^d} \frac{\exp\left\{t\int_{\{0\}}\kappa(\theta+w)\mathscr{H}^0(\mathrm{d}\theta)\right\}-1}{\int_{\{0\}}\kappa(\theta+w)\mathscr{H}^0(\mathrm{d}\theta)}\kappa(w)\mathrm{d}w\lambda_p\mathscr{H}^0(\mathrm{d}y)Q(\mathrm{d}\eta) \\
&\quad = \int_{\mathscr{S}}\int_{x-\eta}\Big(\int_{R^d}(e^{t\kappa(w)}-1)\mathrm{d}w\Big)\lambda_p\mathscr{H}^0(\mathrm{d}y)Q(\mathrm{d}\eta) \\
&\quad = \Big(\int_{\mathbb{R}^d}(e^{t\kappa(w)}-1)\mathrm{d}w\Big)\lambda_p\underbrace{\int_{\mathscr{S}}\mathscr{H}^0(x-\eta)Q(\mathrm{d}\eta)}_{=n_c} \\
&\quad = \lambda_\Phi\int_{\mathbb{R}^d}(e^{t\kappa(w)}-1)\mathrm{d}w.
\end{aligned}$$

Hence one can see that the same large deviation principle (LDP) stated in Section 5.2 for a Poisson point process Ψ holds even for the Matérn cluster process Φ, replacing the intensity λ_Ψ with λ_Φ in the expression for the rate function J_x^* (16). In a similar vein one can prove the validity of the MDP stated in Section 5.2 for the Matérn cluster process Φ, where again the intensity λ_Ψ is replaced with λ_Φ in (17).

Finally it is easy to see that $C_{Var(x)} \overset{(10)}{=} ||\kappa||_2^2\lambda_\Phi$, from which we may claim that the sequence $\left\{\sqrt{Nr_N^d}(\widehat{\lambda}^{\kappa,N}_{\Phi}(x) - \mathbb{E}\widehat{\lambda}^{\kappa,N}_{\Phi}(x))\right\}_{N\geq 1}$ converges weakly, as $N \to +\infty$, to the normal distribution $\mathrm{N}(0, ||\kappa||_2^2\lambda_\Phi)$.

6. Discussion and concluding remarks

We have proved large and moderate deviation principles for kernel–type estimators of the mean density of Boolean models. Thanks to these results, we have been able to derive the consistency of the estimator, and asymptotic confidence intervals as well.

Theorems 2 and 3 are connected with classical results concerning the kernel–type estimator of the density function of an absolutely continuous random variable due to [31, 35]. Here we want to pinpoint the connection with the classical literature in view of future developments. More specifically, let X be a random variable taking values in $\mathbb{R}^d$ with probability density function f_X, and let $X_1, \dots, X_N$ be a random sample for X. The kernel density estimator $\widehat{f}_X^N(x)$ of $f(x)$ at a point $x \in \mathbb{R}^d$ is traditionally [27, 32, 41] defined as

$$\widehat{f}_X^N(x) := \frac{1}{N r_N^d} \sum_{i=1}^{N} \kappa\Big(\frac{x - X_i}{r_N}\Big), \quad x \in \mathbb{R}^d.$$

The scaling parameter r_N, known as the bandwidth, determines the smoothness of the estimator, and it has to be chosen such that

$$r_N \to 0 \quad N r_N^d \to \infty$$

to obtain an asymptotically unbiased and weakly consistent estimator $\widehat{f}_X^N(x)$. The kernel density estimator $\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)$ of $\lambda_{\Theta_n}(x)$ defined in Eq.(6) may be seen as the natural extension of $\widehat{f}_X^N(x)$ to the case of very general random geometric objects in $\mathbb{R}^d$ of Hausdorff dimension $n > 0$, i.e. not necessarily Boolean models. See also [10, Section 3.3.1].

Large and moderate deviation principles for kernel density estimators of f_X have been investigated in [31, 35] with different techniques, in particular, in [31] the author establishes pointwise, as well as uniform, moderate and large deviations principles for the sequence $\Big\{\widehat{f}_X^N(x) - \mathbb{E}[\widehat{f}_X^N(x)]\Big\}_{N\geq 1}$, even for more general kernel functions κ. We recall here the pointwise results for large and moderate deviations given in [31, Proposition 3.1] and [31, Proposition 2.1], respectively, specializing them with our notation and assumptions on κ:

(LDP) the sequence $\Big\{\widehat{f}_X^N(x) - \mathbb{E}[\widehat{f}_X^N(x)]\Big\}_{N\geq 1}$ satisfies a LDP with speed $v_N = N r_N^d$ and rate function

$$J_x^*(y) = \sup_{t\in\mathbb{R}} \Big\{ty - \Big(f_X(x)\int_{\mathbb{R}^d}(e^{t\kappa(w)} - 1)\mathrm{d}w - t f_X(x)\Big)\Big\}; \tag{18}$$

(MDP) assume that $\Big\{b_N\Big\}_{N\geq 1}$ is a sequence of positive real numbers satisfying

$$\lim_{N\to+\infty} \frac{b_N}{\sqrt{N r_N^d}} = +\infty \quad \text{and} \quad \lim_{N\to+\infty} \frac{b_N}{N r_N^d} = 0;$$

then the sequence $\Big\{(N r_N^d / b_N)(\widehat{f}_X^N(x) - \mathbb{E}[\widehat{f}_X^N(x)])\Big\}_{N\geq 1}$ satisfies a LDP with speed $v_N = b_N^2 / N r_N^d$ and rate function

$$J_x^*(y) = \frac{y^2}{2||\kappa||_2^2 f_X(x)}. \tag{19}$$

If Theorems 2 and 3 were true for a general germ–grain model (not only for Boolean models), then the results concerning random variables just recalled here would follow as a particular case. Indeed a random variable $X \equiv \Theta_0$ can be seen as a trivial germ–grain process driven by the marked point process $\Phi = \{(X, s)\}$ in $\mathbb{R}^d$ with mark space $\mathbf{K} = \mathbb{R}^d$, consisting of one point (X) only, with grain $Z(s) := s$, and intensity measure $\Lambda(\mathrm{d}(y, s)) = f(y)\mathrm{d}y\delta_0(s)\mathrm{d}s$. With these choices equation (5) implies that $\lambda_{\Theta_0}(x) = f(x)$, i.e. the mean density of X amounts to be its probability density function, and the expressions in (18) and (19) follow (formally) by replacing $\Lambda(\mathrm{d}(y, s)) = f(y)\mathrm{d}y\delta_0(s)\mathrm{d}s$ and $\Theta_n = X$ in Theorem 2 and Theorem 3, respectively, in analogous way as we did in Section 5.2. Note that the further term $tf_X(x)$ appearing in (18) is due to having considered now the sequence $\left\{\widehat{f}_X^N(x) - \mathbb{E}[\widehat{f}_X^N(x)]\right\}_{N\geq 1}$ instead of $\left\{\widehat{f}_X^N(x)\right\}_{N\geq 1}$.

Hence we may ask whether the theorems obtained for Boolean models extend to more general random closed sets, e.g. germ–grain models. In such a case, as just observed here, the results of [31, 35] would follow as a particular case of a more general theory. Otherwise, if the extension is not possible, the independence property of the underlying Poisson point processes would be peculiar in obtaining such expressions. This problem remains open and requires different kinds of techniques with respect to the ones employed here, which are mainly based on the availability of the Laplace functional of a Poisson point process.

Finally it is worth to underline that the theoretical results proved in this paper and in [12] may be useful in many applications, for example to determine confidence intervals for the estimators. A future work in this direction, we are working on, will be focused on simulation studies of the kernel–type estimator in comparison with other estimators, such as the "Minkowski content"–based estimator mentioned in the Introduction.

Appendix A: Proofs of the main theorems

A.1. Proof of Theorem 2

Before proving Theorem 2, we provide two thecnical lemmas. For the sake of simplifying notation we define

$$h(\xi, s) := \frac{1}{r^n} \int_{\xi + Z(s)} \kappa\Big(\frac{x-y}{r}\Big) \mathscr{H}^n(\mathrm{d}y), \tag{20}$$

and we shall write $h_N(\xi, s)$ if $r = r_N$ in the above definition.

Lemma 10. *Let Θ_n and κ be as in the Assumptions. For any $r < 1$ and $t \in \mathbb{R}$*

$$\begin{aligned} \mathbb{E}\Big[\exp\Big\{\frac{t}{r^n}\int_{\Theta_n} \kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}\Big] \\ = \exp\Big\{\int_{\mathbb{R}^d \times \mathbf{K}} (e^{th(\xi,s)} - 1) f(\xi, s)\mathrm{d}\xi Q(\mathrm{d}s)\Big\}. \end{aligned} \tag{21}$$

Proof. First of all we remind that (see [34, pg. 28]) if Ψ is a Poisson point process on $\mathbb{X}$ with intensity measure μ, then, for any measurable function $g : \mathbb{X} \to \mathbb{R}$ such that $\int_{\mathbb{X}} \min\{|g(x)|, 1\}\, \mu(\mathrm{d}x) < \infty$, it holds

$$\mathbb{E}\Big[\exp\Big\{\vartheta \sum_{x\in\Psi} g(x)\Big\}\Big] = \exp\Big\{\int_{\mathbb{X}} (e^{\vartheta g(x)} - 1)\mu(\mathrm{d}x)\Big\} \tag{22}$$

for any complex number ϑ.

By observing that $\min\{h(\xi,s),1\} \le h(\xi,s)$, and that $1/r^n \le 1/r^d$ if $r \le 1$, we have

$$\begin{aligned}
&\int_{\mathbb{R}^d\times\mathbf{K}} \min\Big\{h(\xi,s),1\Big\} f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s) \\
&\quad \le \frac{1}{r^d}\int_{\mathbf{K}}\int_{\mathbb{R}^d}\int_{\xi+Z(s)} \kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y) f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s) \\
&\quad = \frac{1}{r^d}\int_{\mathbf{K}}\int_{\mathbb{R}^d}\int_{Z(s)} \kappa\Big(\frac{x-\tilde{y}-\xi}{r}\Big)\mathscr{H}^n(\mathrm{d}\tilde{y}) f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s) \\
&\quad = \int_{\mathbf{K}}\int_{\mathbb{R}^d}\kappa(w)\int_{Z(s)} f(x-\tilde{y}-rw,s)\mathscr{H}^n(\mathrm{d}\tilde{y})\mathrm{d}w Q(\mathrm{d}s)
\end{aligned}$$

We remind that κ is a kernel with $\mathrm{supp}(\kappa) \in B_R(0)$, $Z(s) \subseteq \Xi(s)$, and we notice that if $\tilde{y} \in Z(s)$ and $w \in B_R(0)$ then $x - rw \in B_{Rr}(x) \subseteq B_R(x)$. Therefore $x - \tilde{y} - rw \in B_R(x) - Z(s) \subset B_R(x)_{\oplus \mathrm{diam}(Z(s))}$, so that by (A2) we get

$$\begin{aligned}
&\int_{\mathbf{K}}\int_{\mathbb{R}^d}\kappa(w)\int_{Z(s)} f(x-\tilde{y}-rw,s)\mathscr{H}^n(\mathrm{d}\tilde{y})\mathrm{d}w Q(\mathrm{d}s) \\
&\qquad\qquad \le \int_{\mathbf{K}} \tilde{\xi}_{B_R(x)}(s)\mathscr{H}^n(\Xi(s))Q(\mathrm{d}s) \overset{(A2)}{<} \infty.
\end{aligned} \tag{23}$$

Thus we may write

$$\mathbb{E}\Big[\exp\Big\{t\sum_{(\xi,s)\in\Phi} h(\xi,s)\Big\}\Big] \overset{(22)}{=} \exp\Big\{\int_{\mathbb{R}^d\times\mathbf{K}} (e^{th(\xi,s)}-1) f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s)\Big\}.$$

Finally, Lemma 3 in [44] guarantees that the event that different grains of Θ_n overlap in a subset of $\mathbb{R}^d$ of positive $\mathscr{H}^n$-measure has null probability, therefore we may claim that

$$\begin{aligned}
&\mathbb{E}\Big[\exp\Big\{\frac{t}{r^n}\int_{\Theta_n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}\Big] \\
&\qquad = \mathbb{E}\Big[\exp\Big\{\frac{t}{r^n}\sum_{(\xi,s)\in\Phi}\int_{\xi+Z(s)}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}\Big],
\end{aligned}$$

that is the assertion. □

Lemma 11. *Let Θ_n and κ be as in the Assumptions. If $r < \min\left\{1, 1/(2R)\right\}$, the following bound holds for any $s \in \mathbf{K}$, $t \in \mathbb{R}$, $w \in \mathbb{R}^d$, and $\mathscr{H}^n$-a.e. $y \in x - Z(s)$:*

$$\left|\frac{\left(\exp\left\{t\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})\right\}-1\right)}{\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})}\kappa(w)f(y-wr,s)\right| \\ \leq \Psi(t)\kappa(w)\tilde{\xi}_{B_R(x)}(s), \tag{24}$$

where

$$\Psi(t) = \begin{cases} |t| & \text{if } t \leq 0 \\ \dfrac{e^{tM\tilde{\gamma}(2R)^n}-1}{M\tilde{\gamma}(2R)^n} & \text{if } t > 0 \end{cases}$$

with

$$\int_{\mathbf{K}}\int_{x-Z(s)}\int_{\mathbb{R}^d}\Psi(t)\kappa(w)\tilde{\xi}_{B_R(x)}(s)\mathrm{d}w\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s) < +\infty. \tag{25}$$

Proof. First of all consider the case $t \leq 0$.

$$\begin{aligned}&\left|\frac{\left(\exp\left\{t\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})\right\}-1\right)}{\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})}\kappa(w)f(y-wr,s)\right| \\ &\quad = \frac{\left(1-\exp\left\{t\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})\right\}\right)}{\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})}\kappa(w)f(y-wr,s) \\ &\quad \leq |t|\kappa(w)f(y-wr,s),\end{aligned}$$

being $1 - e^{\alpha} \leq -\alpha$ for any $\alpha \in \mathbb{R}$.

The case $t > 0$ is less trivial, and we employ the Taylor series expansion of the exponential

$$\begin{aligned}&\left|\frac{\left(\exp\left\{t\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})\right\}-1\right)}{\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})}\kappa(w)f(y-wr,s)\right| \\ &\quad = \frac{\left(\exp\left\{t\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})\right\}-1\right)}{\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})}\kappa(w)f(y-wr,s) \\ &\quad = \kappa(w)f(y-wr,s)\sum_{k\geq 1}\frac{t^k}{k!}\Big(\int_{[(x-Z(s))-y]/r}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)\Big)^{k-1}.\end{aligned} \tag{26}$$

Now we focus on the integral in (26):

$$\begin{aligned}\int_{[(x-Z(s))-y]/r} & \kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta) \\ &\le \int_{[(x-Z(s))-y]/r} M\mathbb{1}_{B_R(0)}(\theta+w)\mathscr{H}^n(\mathrm{d}\theta) \\ &= \frac{M}{r^n}\int_{Z(s)}\mathbb{1}_{B_R(0)}\Big(w+\frac{x-\tilde{\theta}-y}{r}\Big)\mathscr{H}^n(\mathrm{d}\tilde{\theta}) \\ &\le \frac{M}{r^n}\int_{Z(s)}\mathbb{1}_{B_{rR}(wr+(x-y))}(\tilde{\theta})\mathscr{H}^n(\mathrm{d}\tilde{\theta}) \\ &\le \frac{M}{r^n}\mathscr{H}^n(\Xi(s)\cap B_{rR}(wr+(x-y))).\end{aligned} \tag{27}$$

By replacing this in (26) and by remembering that $\operatorname{supp}(\kappa)\subset B_R(0)$, we obtain

$$\begin{aligned}\Big|&\frac{\Big(\exp\Big\{t\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})\Big\}-1\Big)}{\int_{[(x-Z(s))-y]/r}\kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y})}\kappa(w)f(y-wr,s)\Big| \\ &\le \kappa(w)f(y-wr,s) \\ &\qquad\times\sum_{k\ge1}\frac{t^kM^{k-1}}{r^{(k-1)n}k!}(\mathscr{H}^n(\Xi(s)\cap B_{rR}(wr+(x-y))))^{k-1} \\ &= \mathbb{1}_{B_R(0)}(w)\kappa(w)f(y-wr,s) \\ &\qquad\times\sum_{k\ge1}\frac{t^kM^{k-1}}{r^{(k-1)n}k!}(\mathscr{H}^n(\Xi(s)\cap B_{rR}(wr+(x-y))))^{k-1} \\ &\le \mathbb{1}_{B_R(0)}(w)\kappa(w)f(y-wr,s) \\ &\qquad\times\sum_{k\ge1}\frac{t^kM^{k-1}}{r^{(k-1)n}k!}(\mathscr{H}^n(\Xi(s)\cap B_{2rR}(x-y)))^{k-1}\end{aligned} \tag{28}$$

where we have used the fact that $w\in B_R(0)$. By assumption, $x-y\in Z(s)$ and $2Rr\le1$; thus ($\overline{A1}$) implies

$$\begin{aligned}\sum_{k\ge1}\frac{t^kM^{k-1}}{r^{(k-1)n}k!}(\mathscr{H}^n(\Xi(s)\cap B_{2rR}(x-y)))^{k-1} &\le \sum_{k\ge1}\frac{t^kM^{k-1}}{r^{(k-1)n}k!}(\tilde{\gamma}(2rR)^n)^{k-1} \\ &= \frac{1}{M\tilde{\gamma}(2R)^n}\sum_{k\ge1}\frac{(tM\tilde{\gamma}(2R)^n)^k}{k!}=\Psi(t)\end{aligned}$$

Moreover, being $y\in x-Z(s)$ and $r\le1$, we observe that, for any $w\in B_R(0)$,

$$y-wr\in x-Z(s)+B_R(0)\subseteq B_R(x)_{\oplus\mathrm{diam}(Z(s))},$$

therefore $\mathbb{1}_{B_R(0)}(w)f(y-wr,s)\overset{(A2)}{\le}\tilde{\xi}_{B_R(x)}(s)$, and (24) it is now proved by replacing the above inequalities in (28).

Finally, the integrability condition (25) easy follows:

$$\begin{aligned}\int_{\mathbf{K}}\int_{x-Z(s)}\int_{\mathbb{R}^d}&\Psi(t)\kappa(w)\tilde{\xi}_{B_R(x)}(s)\mathrm{d}w\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s)\\ &\le \Psi(t)Mb_dR^d\int_{\mathbf{K}}\tilde{\xi}_{B_R(x)}(s)\Big(\int_{x-Z(s)}\mathscr{H}^n(\mathrm{d}y)\Big)Q(\mathrm{d}s)\\ &\le \Psi(t)Mb_dR^d\int_{\mathbf{K}}\mathscr{H}^n(\Xi(s))\tilde{\xi}_{B_R(x)}(s)Q(\mathrm{d}s)\overset{(A2)}{<}\infty. \qquad\square\end{aligned}$$

Proof of Theorem 2. The proof relies on the Gärtner-Ellis Theorem. First of all we will show that

$$\begin{aligned}J(t) &:= \lim_{N\to\infty}\frac{1}{v_N}\log\mathbb{E}[e^{tv_N\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)}] \qquad (29)\\ &= \int_{\mathbf{K}}\int_{x-Z(s)}\int_{\mathbb{R}^d}\frac{\exp\Big\{t\int_{\pi_y(x-Z(s))}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)\Big\}-1}{\int_{\pi_y(x-Z(s))}\kappa(\theta+w)\mathscr{H}^n(\mathrm{d}\theta)}\\ &\qquad\qquad\times\kappa(w)\mathrm{d}wf(y,s)\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s)\end{aligned}$$

then we observe that J is a smooth function defined on $\mathbb{R}$, hence satisfying the assumptions of the Gärtner-Ellis Theorem. Since $\Big\{\Theta_n^{(i)}\Big\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random sets, then for N sufficiently big so that $r_N<1$

$$\begin{aligned}&\frac{1}{v_N}\log\mathbb{E}[e^{tv_N\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)}]\\ &\quad=\frac{1}{v_N}\log\mathbb{E}\Big[\exp\Big\{tNr_N^{d-n}\frac{1}{Nr_N^d}\sum_{i=1}^N\int_{\Theta_n^{(i)}}\kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}\Big]\\ &\quad=N\frac{1}{v_N}\log\mathbb{E}\Big[\exp\Big\{\frac{t}{r_N^n}\int_{\Theta_n}\kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}\Big]\\ &\quad=\frac{1}{r_N^{d-n}}\log\mathbb{E}\Big[\exp\Big\{\frac{t}{r_N^n}\sum_{(\xi,s)\in\Phi}\int_{\xi+Z(s)}\kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}\Big]\\ &\quad\overset{(21)}{=}\frac{1}{r_N^{d-n}}\int_{\mathbb{R}^d\times\mathbf{K}}\Big(\exp\Big\{\frac{t}{r_N^n}\int_{\xi+Z(s)}\kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big\}-1\Big)f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s)\\ &\quad=\frac{1}{r_N^{d-n}}\int_{\mathbf{K}}\int_{\mathbb{R}^d}\Big(\exp\Big\{t\int_{[(x-Z(s))-\xi]/r_N}\kappa(\tilde{y})\mathscr{H}^n(\mathrm{d}\tilde{y})\Big\}-1\Big)f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s).\end{aligned}$$

It is worth to multiply and divide the above integrand by $\int_{[(x-Z(s))-\xi]/r_N}\kappa(y)\times\mathscr{H}^n(\mathrm{d}y)$; then, by suitable changes of variable, the following chain of equality holds:

$$
\begin{aligned}
&\frac{1}{v_N} \log \mathbb{E}[e^{tv_N \widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)}] \\
&\quad = \frac{1}{r_N^{d-n}} \int_{\mathbf{K}} \int_{\mathbb{R}^d} \frac{\exp\left\{t \int_{[(x-Z(s))-\xi]/r_N} \kappa(\tilde{y}) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{[(x-Z(s))-\xi]/r_N} \kappa(\tilde{y}) \mathscr{H}^n(\mathrm{d}\tilde{y})} \\
&\qquad\qquad \times \int_{\xi+Z(s)} \kappa\Big(\frac{x-y}{r_N}\Big) r_N^{-n} \mathscr{H}^n(\mathrm{d}y) f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s) \\
&\quad = \frac{1}{r_N^{d}} \int_{\mathbf{K}} \int_{Z(s)} \int_{\mathbb{R}^d} \frac{\exp\left\{t \int_{[(x-Z(s))-\xi]/r_N} \kappa(\tilde{y}) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{[(x-Z(s))-\xi]/r_N} \kappa(\tilde{y}) \mathscr{H}^n(\mathrm{d}\tilde{y})} \\
&\qquad\qquad \times \kappa\Big(\frac{x-y-\xi}{r_N}\Big) f(\xi,s)\mathrm{d}\xi \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s) \\
&\quad = \int_{\mathbf{K}} \int_{Z(s)} \int_{\mathbb{R}^d} \frac{\exp\left\{t \int_{[(y-Z(s))/r_N+w]} \kappa(\tilde{y}) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{[(y-Z(s))/r_N+w]} \kappa(\tilde{y}) \mathscr{H}^n(\mathrm{d}\tilde{y})} \\
&\qquad\qquad \times \kappa(w) f(x-y-wr_N,s)\mathrm{d}w \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s) \\
&\quad = \int_{\mathbf{K}} \int_{Z(s)} \int_{\mathbb{R}^d} \frac{\exp\left\{t \int_{(y-Z(s))/r_N} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{(y-Z(s))/r_N} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})} \\
&\qquad\qquad \times \kappa(w) f(x-y-wr_N,s)\mathrm{d}w \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s) \\
&\quad = \int_{\mathbf{K}} \int_{x-Z(s)} \int_{\mathbb{R}^d} \frac{\exp\left\{t \int_{[(x-Z(s))-y]/r_N} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{[(x-Z(s))-y]/r_N} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})} \\
&\qquad\qquad \times \kappa(w) f(y-wr_N,s)\mathrm{d}w \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s).
\end{aligned}
$$

Denoted by $\mathscr{D}_f(s)$ the set of discontinuity points of $f(\cdot,s)$ for any $s \in \mathbf{K}$, assumption (A2) implies $\mathscr{H}^n(\mathscr{D}_f(s)) = 0$, therefore we can see that

$$
\begin{aligned}
&\lim_{N\to+\infty} \frac{\exp\left\{t \int_{[(x-Z(s))-y]/r_N} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{[(x-Z(s))-y]/r_N} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})} \kappa(w) f(y-wr_N,s) \\
&\qquad \overset{(7)}{=} \frac{\exp\left\{t \int_{\pi_y(x-Z(s))} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})\right\} - 1}{\int_{\pi_y(x-Z(s))} \kappa(\tilde{y}+w) \mathscr{H}^n(\mathrm{d}\tilde{y})} \kappa(w) f(y,s)
\end{aligned}
$$

for any $s \in \mathbf{K}$, $w \in \mathbb{R}^d$, and $\mathscr{H}^n$-a.e. $y \in x - Z(s)$.

Thus the (29) follows by a simple application of the dominated convergence theorem, whose validity is guaranteed by Lemma 11.

To conclude the proof we observe that J satisfies the assumptions of Theorem 1. More precisely, as a byproduct of the application of the dominated convergence theorem, $J(t) < +\infty$ for any $t \in \mathbb{R}$. Finally we show that J is differentiable on $\mathbb{R}$ with

$$J'(t_0) = \int_{\mathbf{K}} \int_{x-Z(s)} \int_{\mathbb{R}^d} \exp\left\{ t_0 \int_{\pi_y(x-Z(s))} \kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y}) \right\} \times k(w) f(y,s) \mathrm{d}w \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s)$$

for any $t_0 \in \mathbb{R}$. In order to prove this, fix $t_0 \in \mathbb{R}$ and $\delta > 0$ sufficiently small; following [7, Theorem 16.8], we need to show that the integrand

$$\exp\left\{ t \int_{\pi_y(x-Z(s))} \kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y}) \right\} k(w) f(y,s) \mathbb{1}_{(x-Z(s))}(y) \tag{30}$$

is bounded from above for any $t \in (t_0 - \delta, t_0 + \delta)$ by an integrable function. To this end, the definition of approximate tangent space and similar arguments as in (27) give

$$\int_{\pi_y(x-Z(s))} \kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y}) = \lim_{r\to 0} \int_{\frac{x-Z(s)-y}{r}} \kappa(\tilde{y}+w)\mathscr{H}^n(\mathrm{d}\tilde{y}) \le M\tilde{\gamma}(2R)^n,$$

with $w \in B_R(0) \supset \operatorname{supp}(k)$. Therefore

$$(30) \le \max\left\{ e^{(t_0+\delta)M\tilde{\gamma}(2R)^n}, e^{(t_0-\delta)M\tilde{\gamma}(2R)^n} \right\} k(w) f(y,s) \mathbb{1}_{(x-Z(s))}(y)$$

when $t \in (t_0 - \delta, t_0 + \delta)$, with

$$\begin{aligned}\int_{\mathbf{K}} \int_{x-Z(s)} \int_{\mathbb{R}^d} &\max\left\{ e^{(t_0+\delta)M\tilde{\gamma}(2R)^n}, e^{(t_0-\delta)M\tilde{\gamma}(2R)^n} \right\} \\ &\times k(w) f(y,s) \mathrm{d}w \mathscr{H}^n(\mathrm{d}y) Q(\mathrm{d}s) \\ &\qquad \le \lambda_{\Theta_n}(x) \max\left\{ e^{(t_0+\delta)M\tilde{\gamma}(2R)^n}, e^{(t_0-\delta)M\tilde{\gamma}(2R)^n} \right\} < +\infty,\end{aligned}$$

hence the thesis follows. □

A.2. Proof of Theorem 3

Before proving Theorem 3 we need some useful lemmas.

Lemma 12. *Let Θ_n and κ be as in the Assumptions, and let us define*

$$\tau_q(u) := \int_{\mathbf{K}} \int_{\mathbb{R}^d} e^{-uh(\xi,s)} h^q(\xi,s) f(\xi,s) \mathrm{d}\xi Q(\mathrm{d}s), \quad q \ge 1 \tag{31}$$

for any measurable function $h : \mathbb{R}^d \times \mathbf{K} \to \mathbb{R}^+$ and $u \ge 0$.

If $r < \min\{1, 1/(2R)\}$, the function $h(\xi, s)$ defined in (20) satisfies $\tau_q(u) < +\infty$ for any $u \ge 0$ and $q \ge 1$.

Proof. For any $q \geq 1$, and $u \geq 0$ we have

$$
\begin{aligned}
e^{-uh(\xi,s)}h^q(\xi,s) &\leq h^q(\xi,s) \\
&= \frac{1}{r^{nq}}\Big[\int_{\xi+Z(s)} \kappa\Big(\frac{x-\tilde{y}}{r}\Big)\mathscr{H}^n(\mathrm{d}\tilde{y})\Big]^{q-1}\int_{\xi+Z(s)} \kappa\Big(\frac{x-\tilde{y}}{r}\Big)\mathscr{H}^n(\mathrm{d}\tilde{y}),
\end{aligned}
$$

therefore

$$
\begin{aligned}
\tau_q(u) &\leq \frac{1}{r^{qn}}\int_{\mathbf{K}}\int_{\mathbb{R}^d}\int_{Z(s)}\Big[\int_{\xi+Z(s)}\kappa\Big(\frac{x-\tilde{y}}{r}\Big)\mathscr{H}^n(\mathrm{d}\tilde{y})\Big]^{q-1} \\
&\qquad\qquad \times \kappa\Big(\frac{x-\xi-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)f(\xi,s)\mathrm{d}\xi Q(\mathrm{d}s) \\
&\leq \frac{1}{r^{qn-d}}\int_{\mathbf{K}}\int_{Z(s)}\int_{\mathbb{R}^d}\Big(\int_{Z(s)}\kappa\Big(\frac{y-\tilde{y}}{r}+w\Big)\mathscr{H}^n(\mathrm{d}\tilde{y})\Big)^{q-1} \\
&\qquad\qquad \times \kappa(w)f(x-y-rw,s)\mathrm{d}w\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s).
\end{aligned} \tag{32}
$$

By remembering that $\kappa(x) \leq M\mathbb{1}_{B_R(0)}(x)$ for all $x \in \mathbb{R}^d$, and $Z(s) \subseteq \Xi(s)$,

$$
\begin{aligned}
\int_{Z(s)}\kappa\Big(\frac{y-\tilde{y}}{r}+w\Big)\mathscr{H}^n(\mathrm{d}\tilde{y}) &\leq M\int_{Z(s)}\mathbb{1}_{B_R(0)}\Big(\frac{y-\tilde{y}}{r}+w\Big)\mathscr{H}^n(\mathrm{d}\tilde{y}) \\
&= M\int_{Z(s)}\mathbb{1}_{B_{rR}(y+rw)}(\tilde{y})\mathscr{H}^n(\mathrm{d}\tilde{y}) \leq M\mathscr{H}^n(\Xi(s)\cap B_{rR}(y+rw)).
\end{aligned} \tag{33}
$$

Note that $B_{rR}(y+rw) \subseteq B_{2rR}(y)$ if $w \in B_R(0)$, and that

$$
M\mathscr{H}^n(\Xi(s)\cap B_{2rR}(y)) \overset{(\overline{A1})}{\leq} M\tilde{\gamma}(2rR)^n \qquad \forall y \in Z(s),\ r<1.
$$

Thus, by replacing (33) in (32) we get

$$
\begin{aligned}
\tau_q(u) &\leq \frac{1}{r^{qn-d}}\int_{\mathbf{K}}\int_{Z(s)}\int_{\mathbb{R}^d}(M\tilde{\gamma}(2rR)^n)^{q-1} \\
&\qquad\qquad \times \kappa(w)f(x-y-rw,s)\mathrm{d}w\mathscr{H}^n(\mathrm{d}y)Q(\mathrm{d}s) \\
&\overset{(23)}{\leq} (M\tilde{\gamma})^{q-1}(2R)^{n(q-1)}r^{d-n}\int_{\mathbf{K}}\mathscr{H}^n(\Xi(s))\tilde{\xi}_{B_R(x)}(s)Q(\mathrm{d}s) < \infty
\end{aligned}
$$

□

In order to make the proof of the next lemma more readable, we recall here some basics on Stirling numbers. The Stirling numbers of the second kind $S(n,k)$ count the number of partitions of n objects in k groups. They are extensively studied in [21, 22]: refer to them for additional details on the subject. The Stirling number of the second kind is defined by

$$
S(n,k) = \frac{n!}{k!}\sum_{(\star)}\frac{1}{q_1!\dots q_k!} \tag{34}
$$

where the summation is extended over all positive integers which are solution of the equation $q_1 + \ldots + q_n = n$. It is worth noticing that the summation

$$B_n := \sum_{k=0}^{n} S(n,k)$$

is knows as the Bell number, which amounts to be the number of partition of k objects in distinct sets (see [21, pg. 292]). By [21, pg. 97] we have the following representation

$$B_n = e^{-1} \sum_{j=0}^{+\infty} \frac{j^n}{j!}. \tag{35}$$

The Stirling number of the second kind satisfy a useful recurrence relation

$$S(n+1,k) = S(n,k-1) + kS(n,k), \quad k = 1,\ldots,n-1, \quad n = 0,1,\ldots$$

with initial conditions

$$S(0,0) = 1, \quad S(n,0) = 0 \quad \text{for } n > 0, \quad S(n,k) = 0 \quad \text{for } k > n.$$

Lemma 13. *Let Θ_n and κ be as in the Assumptions; then for any $t_0 > 0$*

$$\sum_{k\geq 3} \frac{t_0^k}{k!} \mathbb{E}\Big[\Big(\int_{\Theta_n} \frac{1}{r^n} \kappa\Big(\frac{x-y}{r}\Big) \mathscr{H}^n(\mathrm{d}y)\Big)^k\Big] = O(r^{d-n}) \qquad \text{as } r \to 0. \tag{36}$$

Proof. With the notation introduced in (20), the same argument at the end of the froof of Lemma 10, together with traditional combinatorial arguments, show that for any $k \geq 3$:

$$\mathbb{E}\Big[\Big(\int_{\Theta_n} \frac{1}{r^n} \kappa\Big(\frac{x-y}{r}\Big) \mathscr{H}^n(\mathrm{d}y)\Big)^k\Big] = \mathbb{E}\left[\left(\sum_{(\xi,s)\in\Phi} h(\xi,s)\right)^k\right]$$

$$= \sum_{i=1}^{k} \sum_{(\star)} \binom{k}{q_1 \ldots q_i} \mathbb{E}\left[\sum_{\substack{\xi_{\ell_1} < \ldots < \xi_{\ell_i} \\ (\xi_{\ell_r}, s_{\ell_r}) \in \Phi}} \prod_{r=1}^{i} h^{q_r}(\xi_{\ell_r}, s_{\ell_r})\right]$$

where the sum over $(\star)$ runs over all the vectors $(q_1,\ldots,q_i)$ of positive integers such that $q_1 + \cdots + q_i = k$. Besides we have used the fact that the marked point process is simple.

Since there are $i!$ possible permutations of the points $\xi_{\ell_1},\dots,\xi_{\ell_i}$ we can write

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_{\Theta_n}\frac{1}{r^n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big]\\
&\quad=\sum_{i=1}^{k}\frac{1}{i!}\sum_{(\star)}\binom{k}{q_1\dots q_i}\mathbb{E}\left[\sum_{\substack{\xi_{\ell_r}\neq\xi_{\ell_s}\\(\xi_{\ell_r},s_{\ell_r})\in\Phi}}\prod_{r=1}^{i}h^{q_r}(\xi_{\ell_r},s_{\ell_r})\right]\\
&\quad=\sum_{i=1}^{k}\frac{1}{i!}\sum_{(\star)}\binom{k}{q_1\dots q_i}\int_{(\mathbb{R}^d\times\mathbf{K})^i}\prod_{r=1}^{i}h^{q_r}(\xi_{\ell_r},s_{\ell_r})\nu_{[i]}(\mathrm{d}(x_1,s_1),\dots\mathrm{d}(x_i,s_i)),
\end{aligned}$$

where $\nu_{[i]}$ is the i–th factorial moment measure of Φ (e.g. see [23]). Being Φ a marked Poisson pont processes, then $\nu_{[i]}=\bigotimes_i\Lambda$ (see for example [39, Corollary 3.2.4]); hence we get

$$\mathbb{E}\Big[\Big(\int_{\Theta_n}\frac{1}{r^n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big]=\sum_{i=1}^{k}\frac{1}{i!}\sum_{(\star)}\binom{k}{q_1\dots q_i}\tau_{q_1}(0)\dots\tau_{q_i}(0),$$

where τ is the function defined in (31). By the end of the proof of Lemma 12 we konw that

$$\tau_q(0)\le M^{q-1}\tilde{\gamma}^{q-1}(2R)^{n(q-1)}r^{d-n}\mathbb{E}[\mathscr{H}^n(\Xi)\tilde{\xi}_{B_R(x)}]$$

whenever r is sufficiently small, i.e. $r\le\min\Big\{1,1/(2R)\Big\}$, therefore

$$tau_{q_1}(0)\cdots\tau_{q_i}(0)\le(M\tilde{\gamma}(2R)^n)^{k-i}r^{(d-n)i}\mathbb{E}[\mathscr{H}^n(\Xi)\tilde{\xi}_{B_R(x)}]^i,$$

and so

$$\begin{aligned}
&\mathbb{E}\Big[\Big(\int_{\Theta_n}\frac{1}{r^n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big]\\
&\qquad\overset{(34)}{=}\sum_{i=1}^{k}\Big(\frac{r^{d-n}\mathbb{E}[\mathscr{H}^n(\Xi)\tilde{\xi}_{B_R(x)}]}{M\tilde{\gamma}(2R)^n}\Big)^iS(k,i)(M\tilde{\gamma}(2R)^n)^k.
\end{aligned}$$

Now we define the constant function $C:=C(\Theta_n,\kappa)$, depending on Θ_n and the kernel κ,

$$C:=\max\Big\{1,\frac{\mathbb{E}[\mathscr{H}^n(\Xi)\tilde{\xi}_{B_R(x)}]}{M\tilde{\gamma}(2R)^n}\Big\};$$

as a consequence

$$\begin{aligned}
\mathbb{E}\Big[\Big(\int_{\Theta_n}\frac{1}{r^n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big]&\le\sum_{i=1}^{k}C^kS(k,i)(M\tilde{\gamma}(2R)^n)^kr^{(d-n)i}\\
&=\sum_{i=1}^{k}\tilde{C}^kS(k,i)r^{(d-n)i}
\end{aligned}$$

where we have put

$$\tilde{C} := \max\Big\{M\tilde{\gamma}(2R)^n, \mathbb{E}[\mathscr{H}^n(\Xi)\tilde{\xi}_{B_R(x)}]\Big\}.$$

Recalling the definition of the Bell numbers B_k (see (35)) we have

$$\begin{aligned}\mathbb{E}\Big[\Big(\int_{\Theta_n}\frac{1}{r^n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big] &\leq \tilde{C}^k\sum_{i=1}^k S(k,i)r^{d-n} = \tilde{C}^k r^{d-n}\sum_{i=1}^k S(k,i)\\ &= B_k\tilde{C}^k r^{d-n} = \tilde{C}^k r^{d-n}\frac{1}{e}\sum_{m=0}^{\infty}\frac{m^k}{m!}.\end{aligned}$$

Now we consider the summation in (36) for any $t_0>0$. Finally, the previous bound for the expectation yields

$$\begin{aligned}\sum_{k\geq 3}\frac{t_0^k}{k!}\mathbb{E}\Big[\Big(\int_{\Theta_n}\frac{1}{r^n}\kappa\Big(\frac{x-y}{r}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big] &\leq \sum_{k\geq 3}\frac{t_0^k}{k!}\frac{\tilde{C}^k r^{d-n}}{e}\sum_{m=0}^{\infty}\frac{m^k}{m!}\\ &= \frac{r^{d-n}}{e}\sum_{k=3}^{\infty}\frac{t_0^k\tilde{C}^k}{k!}\sum_{m=0}^{\infty}\frac{m^k}{m!} \leq \frac{r^{d-n}}{e}\sum_{m=0}^{\infty}\frac{1}{m!}\sum_{k=0}^{\infty}\frac{(\tilde{C}mt_0)^k}{k!}\\ &= \frac{r^{d-n}}{e}\sum_{m=0}^{\infty}\frac{e^{t_0\tilde{C}m}}{m!} = r^{d-n}\exp\Big\{e^{t_0\tilde{C}}-1\Big\},\end{aligned}$$

and the r.h.s. of this inequality turns out to be a $O(r^{d-n})$, which implies the assertion. ☐

Finally we recall that the discrete version of the Hölder inequality can be written as follows

$$\sum_{i=1}^n x_i y_i \leq \left(\sum_{i=1}^n x_i^p\right)^{1/p}\cdot\left(\sum_{i=1}^n y_i^q\right)^{1/q}, \tag{37}$$

where $x_i, y_i \geq 0$ for any $i=1,\dots,n$, and $p,q>0$ are such that $1/p+1/q=1$. By specializing (37) with $n=2$, $y_i=1$ for $i=1,2$, $p=k$ and $q=k/(k-1)$, it directly follows that

$$(x_1+x_2)^k \leq \Big[\Big(\sum_{i=1}^2 x_i^k\Big)^{1/k}\cdot\Big(\sum_{i=1}^2 1\Big)^{k/(k-1)}\Big]^k = (x_1^k+x_2^k)\cdot 2^{k-1} \tag{38}$$

for any $x_1,x_2>0$ and for any integer $k>0$.

We are now ready to prove the theorem.

Proof of Theorem 3. Let us define

$$H_N(t) := \mathbb{E}\Big[\exp\Big\{v_N t\frac{Nr_N^{d-n}}{b_N}\big(\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)-\mathbb{E}[\widehat{\lambda}^{\kappa,N}_{\Theta_n}(x)]\big)\Big\}\Big],$$

and

$$J_x(t) := \lim_{N\to+\infty} \frac{1}{v_N} \log H_N(t), \tag{39}$$

after proving that the limit exists finite for any $t \in \mathbb{R}$; then the good rate function will turn out to be

$$J_x^*(y) = \sup_{t\in\mathbb{R}} \Big\{ ty - J_x(t) \Big\},$$

as a direct application of Theorem 1.

First of all observe that, for any $t \in \mathbb{R}$,

$$\begin{aligned} H_N(t) &= \mathbb{E}\Big[\exp\Big\{ \frac{tb_N}{Nr_N^d} \Big(\sum_{i=1}^N \Big(\int_{\Theta_n^{(i)}} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \\ &\qquad - \mathbb{E} \int_{\Theta_n^{(i)}} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \Big)\Big) \Big\} \Big] \\ &= \Big(\mathbb{E}\Big[\exp\Big\{ \frac{tb_N}{Nr_N^d} \Big(\int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \\ &\qquad - \mathbb{E} \int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \Big) \Big\} \Big] \Big)^N \\ &= \Big(1 + \frac{1}{2} \Big(\frac{tb_N}{Nr_N^d} \Big)^2 Var\Big(\int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \Big) + R(N) \Big)^N, \end{aligned} \tag{40}$$

In order to bound the term $R(N)$ appearing in the previous equation, we note that for any real valued random variable X the following inequality holds

$$\mathbb{E}(X - \mathbb{E}X)^k \le \mathbb{E}|X - \mathbb{E}X|^k \overset{(38)}{\le} 2^{k-1}\mathbb{E}(|X|^k + (\mathbb{E}|X|)^k) \le 2^k \mathbb{E}[|X|^k],$$

where the last inequality follows from a standard application of the Hölder inequality, namely $(\mathbb{E}|X|)^k \le \mathbb{E}|X|^k$. Hence, if $X = \int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)$, the remainder term $R(N)$ in (40) may be estimated as follows

$$\begin{aligned} R(N) &:= \sum_{k\ge 3} \frac{1}{k!} \Big(\frac{tb_N}{Nr_N^d} \Big)^k \mathbb{E}\Big[\Big(\int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \\ &\qquad - \mathbb{E} \int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \Big)^k \Big] \\ &\le \sum_{k\ge 3} \frac{1}{k!} \Big(\frac{|t|b_N}{Nr_N^d} \Big)^k 2^k \mathbb{E}\Big[\Big(\int_{\Theta_n} \kappa\Big(\frac{x-y}{r_N}\Big) \mathscr{H}^n(\mathrm{d}y) \Big)^k \Big]. \end{aligned}$$

For N sufficiently big, the second condition in (9) implies that

$$\frac{2|t|b_N}{Nr_N^{d-n}} \le t_0$$

for some $t_0 > 0$, so that

$$\Big(\frac{2|t|b_N}{Nr_N^d}\Big)^k = \Big(\frac{2|t|b_N}{Nr_N^{d-n}t_0}\Big)^k\Big(\frac{t_0}{r_N^n}\Big)^k \leq \Big(\frac{2|t|b_N}{Nr_N^{d-n}t_0}\Big)^3\Big(\frac{t_0}{r_N^n}\Big)^k \quad \forall k \geq 3;$$

hence we can bound $R(N)$ as

$$\begin{aligned} R(N) &\leq \Big(\frac{2|t|b_N}{Nr_N^{d-n}t_0}\Big)^3 \sum_{k\geq 3}\frac{t_0^k}{k!}\mathbb{E}\Big[\Big(\frac{1}{r_N^n}\int_{\Theta_n}\kappa\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big)^k\Big] \\ &\overset{(36)}{=} O\Big(r_N^{d-n}\Big(\frac{b_N}{Nr_N^{d-n}}\Big)^3\Big). \end{aligned}$$

Hence we obtain

$$\begin{aligned} H_N(t) = \Big(1 + \frac{1}{2}\Big(\frac{tb_N}{Nr_N^d}\Big)^2 Var\Big(\int_{\Theta_n} k\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big) \\ + O\Big(r_N^{d-n}\Big(\frac{b_N}{Nr_N^{d-n}}\Big)^3\Big)\Big)^N. \end{aligned} \tag{41}$$

By assumption $\{\Theta_n^{(i)}\}_{i\in\mathbb{N}}$ is a sequence of i.i.d. random closed sets as Θ_n; therefore

$$\begin{aligned} Var\Big(\int_{\Theta_n} k\Big(\frac{x-y}{r_N}\Big)\mathscr{H}^n(\mathrm{d}y)\Big) &= Nr_N^{2d}Var(\widehat{\lambda}_{\Theta_n}^{\kappa,N}(x)) \\ &\overset{(12)}{=} Nr_N^{2d}\Big(\frac{C_{Var}(x)}{Nr_N^{d-n}} + o(\frac{1}{Nr_N^{d-n}})\Big) \\ &= r_N^{d+n}(C_{Var}(x) + o(1)), \quad \text{as } N \to +\infty. \end{aligned} \tag{42}$$

Thus, we conclude that

$$\begin{aligned} J_x(t) &\overset{(39)}{=} \lim_{N\to+\infty}\frac{Nr_N^{d-n}}{b_N^2}\log H_N(t) \\ &\overset{(41),(42)}{=} \frac{N^2r_N^{d-n}}{b_N^2}\log\Big(1 + \frac{1}{2}\Big(\frac{tb_N}{Nr_N^d}\Big)^2 r_N^{d+n}(C_{Var}(x) + o(1)) \\ &\qquad + O\Big(r_N^{d-n}\Big(\frac{b_N}{Nr_N^{d-n}}\Big)^3\Big)\Big) \\ &= \frac{N^2r_N^{d-n}}{b_N^2}\log\Big(1 + \frac{t^2b_N^2}{2N^2r_N^{d-n}}C_{Var}(x) + o\Big(\frac{b_N^2}{N^2r_N^{d-n}}\Big)\Big) \\ &\overset{(9)}{=} \frac{t^2}{2}C_{Var}(x). \end{aligned}$$

As a consequence the rate function is given by

$$J_x^*(y) = \sup_{t\in\mathbb{R}}\Big\{ty - J_x(t)\Big\} = \sup_{t\in\mathbb{R}}\Big\{ty - \frac{t^2}{2}C_{Var}(x)\Big\} = \frac{y^2}{2C_{Var}(x)}$$

and the assertion follows. □

Acknowledgements

We thank the anonymous referee for the accurate reading of the paper, and the valuable comments and suggestions which improved significantly the quality of this paper.

We also thank Prof. Claudio Macci (Department of Mathematics, Università di Roma Tor Vergata) for fruitful discussions.

The authors are members of the Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INdAM).

References

[1] Aitchison, J., Kay, J.W. and Lauder, I.J. (2004). *Statistical Concepts and Applications in Clinical Medicine*. Chapman and Hall/CRC, London.

[2] Ambrosio, L., Capasso, V. and Villa, E. (2009). On the approximation of mean densities of random closed sets. *Bernoulli*, **15**, 1222–1242. MR2107677

[3] Ambrosio, L., Fusco, N. and Pallara, D. (2000). *Functions of Bounded Variation and Free Discontinuity Problems*. Clarendon Press, Oxford. MR2597590

[4] Baddeley, A., Barany, I., Schneider, R. and Weil, W. (2007). *Stochastic Geometry*. Lecture Notes in Mathematics **1982**, Springer, Berlin. MR1857292

[5] Baddeley, A. and Molchanov, I.S. (1997). On the expected measure of a random set. In: *Proceedings of the International Symposium on Advances in Theory and Applications of Random Sets (Fontainebleau, 1996)*, River Edge, NJ, World Sci. Publishing, 3–20. MR2327289

[6] Beneš, V. and Rataj, J. (2004). *Stochastic Geometry: Selected Topics*. Kluwer, Dordrecht. MR1654394

[7] Billingsley, P. (1995). *Probability and Measure*, 3rd edition. John Wiley & Sons. MR2068607

[8] Bonilla, L.L., Capasso, V., Alvaro, M., Carretero, and Terragni, F. (2017). On the mathematical modelling of tumor–induced angiogenesis. *Math. Biosci. Eng.*, **14**, 45–66. MR1324786

[9] Bryc, W. (1993) A remark on the connection between the large deviation principle and the central limit theorem. *Statist. Probab. Lett.*, **18**, 253–256. MR3562897

[10] Camerlenghi, F., Capasso, V. and Villa, E. (2014). On the estimation of the mean density of random closed sets. *J. Multivariate Anal.*, **125**, 65–88. MR1245694

[11] Camerlenghi, F., Capasso, V. and Villa, E. (2014). Numerical experiments for the estimation of mean densities of random sets. In: Proceedings of the 11th European Congress of Stereology and Image Analysis. *Image Anal. Stereol.*, **33**, 83–94. MR3163831

[12] Camerlenghi, F., Macci, C. and Villa, E. (2016). Asymptotic results for the estimation of the mean density of random closed sets. *Electron. J. Stat.*, **10**, 2066–2096. MR3241508
[13] Camerlenghi, F. and Villa, E. (2015). Optimal bandwidth of the "Minkowski content"–based estimator of the mean density of random closed sets: theoretical results and numerical experiments. *J. Math. Imaging Vision*, **53**, 264–287. MR3522669
[14] Mathematical Modelling for Polymer Processing. Polymerization, Crystallization, Manufacturing (V. Capasso, Editor). ECMI Series on Mathematics in Industry Vol 2, Springer Verlag, Heidelberg 2003. MR3397099
[15] Capasso, V., Dejana, E. and Micheletti, A. (2008). Methods of stochastic geometry, and related statistical problems in the analysis and therapy of tumour growth and tumour-driven angiogenesis. In: *Selected Topics on Cancer Modelling*, (N. Bellomo et al. Eds.), Birkhauser, Boston, 299–335. MR1964653
[16] Capasso, V. and Micheletti, A. (2006). Stochastic geometry and related statistical problems in biomedicine. In: *Complex System in Biomedicine*, (A. Quarteroni et al. Eds.), Springer, Milano, 35–69. MR2477136
[17] Capasso, V., Micheletti, A. and Morale, D. (2008), Stochastic geometric models and related statistical issues in tumour-induced angiogenesis. *Math. Biosci.*, **214**, 20–31. MR2487997
[18] Capasso, V. and Villa, E. (2006). On the continuity and absolute continuity of random closed sets. *Stoch. An. Appl.*, **24**, 381–397. MR2446608
[19] Capasso, V. and Villa, E. (2007). On mean densities of inhomogeneous geometric processes arising in material sciences and medicine. *Image Anal. Setreol.*, **26**, 23–36. MR2204719
[20] Capasso, V. and Villa, E. (2008). On the geometric densities of random closed sets. *Stoch. An. Appl.*, **26**, 784–808. MR2337324
[21] Charalambides, C.A. (2002). *Enumerative combinatorics.* CRC Press Series on Discrete Mathematics and its Applications. Chapman & Hall/CRC, Boca Raton, FL. MR2185580
[22] Charalambides, C.A. (2005). *Combinatorial methods in discrete distributions.* Hoboken, NJ: Wiley.
[23] Chiu, S.N., Stoyan, D., Kendall, W.S. and Mecke, J. (2013). *Stochastic Geometry and its Applications*, 3rd edition, John Wiley & Sons, Chichcester.
[24] Daley, D.J. and Vere-Jones, D. (2003). *An introduction to the theory of point processes. Vol. I. 2nd edition.* Springer, New York.
[25] Daley, D.J. and Vere-Jones, D. (2008). *An introduction to the theory of point processes. Vol. II. 2nd edition.* Springer, New York.
[26] Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. 2nd edition, Springer.
[27] Devroye, L. and Györfi, L. (1985). *Nonparametric density estimation: the L_1 view.* Wiley, New York.
[28] Devroye, L., Györfi, L. and Lugosi, G. (1996). *A Probabilistic Theory*

of Pattern Recognition. Springer Series in Stochastic Modelling and Applied Probability, New York.

[29] Falconer, K.J. (1986). *The Geometry of Fractal Sets*. Cambridge University Press, Cambridge.

[30] Federer, H. (1969). *Geometric Measure Theory*. Springer, Berlin.

[31] Gao, F. (2003). Moderate deviations and large deviations for kernel density estimators. *J. Theoret. Probab.*, **16**, 401–418.

[32] Härdle, W. (1991). *Smoothing Techniques with Implementation in **S***. Springer-Verlag, New York.

[33] Hug, D. and Last, G. (2000). On support measures in Minkowski spaces and contact distributions in stochastic geometry. *Ann. Prob.*, **28**, 796–850.

[34] Kingman, J.F.C. (1993). *Poisson processes.* Oxford University Press, Oxford.

[35] Louani, D. (1998). Large deviations limit theorems for the kernel density estimator. *Scand. J. Statist.*, **25**, 243–253.

[36] Matheron, G. (1975). *Random Sets and Integral Geometry*. John Wiley & Sons, New York.

[37] Molchanov, I. (1997). *Statistics of the Boolean Model for Practitioners and Mathematicians*. Chichester, Wiley.

[38] Molchanov, I. (2005). *Theory of Random Sets*. Springer-Verlag, London.

[39] Schneider, R. and Weil, W. (2008). *Stochastic and Integral Geometry*. Springer-Verlag, Berlin Heidelberg.

[40] Sheather, S.J. (2004). Density estimation, *Statist. Sci.*, **19**, 588–597.

[41] Silverman, B.W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, London.

[42] Szeliski, R. (2011). *Computer Vision. Algorithms and Applications*. Springer-Verlag, London.

[43] Villa E. (2010). Mean densities and spherical contact distribution function of inhomogeneous Boolean models. *Stoch. An. Appl.*, **28**, 480–504.

[44] Villa E. (2014). On the local approximation of mean densities of random closed sets. *Bernoulli*, **20**, 1–27.

[45] Zähle, M. (1982). Random processes of Hausdorff rectifiable closed sets. *Math. Nachr.*, **108**, 49–72.