Electronic Journal of Statistics
Vol. 10 (2016) 2066–2096
ISSN: 1935-7524
DOI: 10.1214/16-EJS1159

Asymptotic results for multivariate estimators of the mean density of random closed sets

Federico Camerlenghi*

Department of Decision Sciences
Bocconi University
Via Röntgen 1, 20136 Milano, Italy
e-mail: federico.camerlenghi@unibocconi.it

Claudio Macci

Dipartimento di Matematica
Università di Roma Tor Vergata
Via della Ricerca Scientifica, 00133 Rome, Italy
e-mail: macci@mat.uniroma2.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 problem of the evaluation and estimation of the mean density of random closed sets in $\mathbb{R}^d$ with integer Hausdorff dimension $0 < n < d$, is of great interest in many different scientific and technological fields. Among the estimators of the mean density available in literature, the so-called "Minkowski content"-based estimator reveals its benefits in applications in the non-stationary cases. We introduce here a multivariate version of such estimator, and we study its asymptotical properties by means of large and moderate deviation results. In particular we prove that the estimator is strongly consistent and asymptotically Normal. Furthermore we also provide confidence regions for the mean density of the involved random closed set in $m \geq 1$ distinct points $x_1, \ldots, x_m \in \mathbb{R}^d$.

MSC 2010 subject classifications: Primary 62F12, 60F10, 60D05.
Keywords and phrases: Minkowski content, large deviations, moderate deviations, random closed sets, confidence regions.

Received January 2016.

Contents

*Also affiliated to the Collegio Carlo Alberto, Moncalieri, Italy.
†Corresponding author.

1. Introduction

A random closed set Θ_n of locally finite n-dimensional Hausdorff measure $\mathcal{H}^n$ induces a random measure $\mu_{\Theta_n}(A) := \mathcal{H}^n(\Theta_n \cap A)$, $A \in \mathcal{B}_{\mathbb{R}^d}$, and the corresponding expected measure is defined as $\mathbb{E}[\mu_{\Theta_n}](A) := \mathbb{E}[\mathcal{H}^n(\Theta_n \cap A)]$, $A \in \mathcal{B}_{\mathbb{R}^d}$, where $\mathcal{B}_{\mathbb{R}^d}$ is the Borel σ-algebra of $\mathbb{R}^d$. (For a discussion of the measurability of the random variables $\mu_{\Theta_n}(A)$, we refer to [5, 27].) Whenever the measure $\mathbb{E}[\mu_{\Theta_n}]$ is absolutely continuous with respect to the d-dimensional Hausdorff measure $\mathcal{H}^d$, its density (i.e. its Radon-Nikodym derivative) with respect to $\mathcal{H}^d$ is called *mean density* of Θ_n, and denoted by λ_{Θ_n}.

Examples of random closed sets with integer Hausdorff dimension n less than d are fiber processes, boundaries of germ-grain models, n-facets of random tessellations, and surfaces of full dimensional random sets. The problem of the evaluation and estimation of the mean density λ_{Θ_n} has been of great interest in many different scientific and technological fields over the last decades. Recent areas of interest include pattern recognition and image analysis [23, 18], computer vision [24], medicine [1, 12, 13, 14], material science [11], etc. (see [8] for additional references).

The estimation of the mean density of non-stationary random sets turns out to be much more difficult both from the theoretical and the applied point of view. With regard to this, an explicit formula for $\lambda_{\Theta_n}(x)$ (see Eq. (2)) and different kinds of estimators have been proposed in recent literature (see [26, 8, 9] and references therein). One of these, named *"Minkowski content"-based estimator* and denoted by $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$ (see Eq. (4)), turns out to be asymptotically unbiased and weakly consistent, and it reveals its benefits in applications in the non-stationary cases. Indeed, the evaluation of the "Minkowski content"-based estimator in a generic point $x \in \mathbb{R}^d$ does not require any particular calculation, except for counting how many elements of the random sample of Θ_n have no void intersection with the ball centered at x. Whenever a random sample for the involved random closed set Θ_n is available, the feasibility of such an estimator is now apparent: in fact its computation reduces to check whether any pixel corresponding to the ball belongs to the sample of Θ_n or not in its digital image. This is the reason why such an estimator deserves to be studied in great detail.

In this paper we consider a multivariate version of the estimator in [8, 10], and we study its rate of convergence to the theoretical mean density in the fashion of the large deviations theory. We remind that such a theory gives an asymptotic computation of small probabilities on an exponential scale (see e.g. [17] as a reference on this topic); here, among the references in the literature, we remind that some large deviation results for the empirical volume fraction for stationary Poisson grain models can be found in [19].

We consider multivariate estimators, i.e. estimators of the density in m points $x_1, \dots, x_m$, and this allows to have a richer asymptotic analysis for the estimation of the density. In fact finite sets of points can be used to estimate the entire density function $\lambda_{\Theta_n}(x)$. To better explain this one could try to adapt some large deviation techniques (see for instance Dawson Gärtner Theorem, i.e. Theorem 4.6.1 in [17]) which allow to lift a collection of large deviation principles in "small" spaces into a large deviation principle on a "large" space.

We remark that in [10] the authors proved the weak consistency of the "Minkowski content"-based estimator in the univariate case. In this paper we prove the strong consistency and the asymptotic Normality for the multivariate estimator. These two issues will be a consequence of some standard arguments in large deviations: see Remark 2 for the strong convergence of sequences of random variables which satisfy the large deviation principle, and Remark 11 which illustrates how the proof of a moderate deviation result can be adapted to prove the weak convergence to a centered Normal distribution.

More specifically we will consider the *multivariate "Minkowski content"-based estimator*

$$\widehat{\boldsymbol{\lambda}}^{\mu,N} := (\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_1), \dots, \widehat{\lambda}^{\mu,N}_{\Theta_n}(x_m)),$$

of the mean density $\boldsymbol{\lambda}_{\Theta_n} := (\lambda_{\Theta_n}(x_1), \dots, \lambda_{\Theta_n}(x_m))$ in $m \geq 1$ distinct points $x_1, \dots x_m$ of $\mathbb{R}^d$. We remark that, given an i.i.d. random sample $\Theta_n^{(1)}, \dots, \Theta_n^{(N)}$ for Θ_n, the estimator $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$ in [8, 10] is of the type $\frac{1}{w_N}\sum_{i=1}^N Y_{i,N}$ where $Y_{i,N} = \mathbf{1}_{\Theta_n^{(i)} \cap B_{r_N(x)} \neq \emptyset}$, $B_{r_N}(x)$ is the ball centered in x with radius r_N, and w_N is a suitable normalization which depends on the *bandwidth* r_N. Thus, as far as the multivariate estimator is concerned, in general two distinct random variables $\mathbf{1}_{\Theta_n^{(i)} \cap B_{r_N(x_j)} \neq \emptyset}$ and $\mathbf{1}_{\Theta_n^{(i)} \cap B_{r_N(x_k)} \neq \emptyset}$ (with $j \neq k$) are not independent, and therefore we have a m-dimensional estimator with (possibly) dependent components. We point out that even in the simpler case $m = 1$, results on the strong consistency and the convergence in law of the "Minkowski content"-based estimator are not still available in literature. Here we directly deal with the general multivariate vector of dimension m in order to derive also confidence regions for the whole vector (see Section 4.2); of course the univariate case $m = 1$ will be seen as a particular case.

In Section 3 we present the asymptotic results for $\boldsymbol{S}_N := \sum_{i=1}^N \boldsymbol{Y}_{i,N}$ where

$$\boldsymbol{Y}_{i,N} \overset{\text{iid}}{\sim} \boldsymbol{Y}_N := (Y_N^{(1)}, \dots, Y_N^{(m)})$$

and $\boldsymbol{Y}_N$ is a random vector with (possibly) dependent Bernoulli distributed components. Then, for some arbitrary deterministic quantity w_N such that $w_N \to \infty$

as $N \to \infty$, we present a condition (see Eq. (10)) on the bivariate joint distributions of pairs of components of $\boldsymbol{Y}_N$ which allows us to prove large and moderate deviations for $\boldsymbol{S}_N/w_N$ (as $N \to \infty$). As a byproduct we get an asymptotic Normality result (Theorem 10).

In Section 4 the results are applied to the multivariate estimator $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ under quite general sufficient conditions on Θ_n (obviously in these applications w_N is chosen in terms of the bandwidth r_N as we said above). In particular we shall see that the sequence $\{\widehat{\boldsymbol{\lambda}}^{\mu,N} : N \geq 1\}$ satisfies a large deviation principle, from which it is possible to gain information on its convergence rate and then to deduce a strong consistency result for $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ (see Corollary 15). In Section 4.2 we also find a confidence region for $\boldsymbol{\lambda}_{\Theta_n}$; this will be done by considering the asymptotical distribution of $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ when an optimal bandwidth r_N is chosen. This result will play a pivotal role to get confidence regions for $\boldsymbol{\lambda}_{\Theta_n}$.

2. Preliminaries

In this section we recall some preliminaries on large deviations and on stochastic geometry, useful for the sequel. We shall refer to literature and previous works for a more exhaustive treatment.

2.1. Large deviations

We start with some basic definitions and we refer to [17] for additional details. Let $\mathcal{Z}$ be a topological space. Then a sequence of $\mathcal{Z}$-valued random variables $\{Z_N : N \geq 1\}$ satisfies the large deviation principle (LDP for short) with speed v_N and rate function I if $\lim_{N\to\infty} v_N = \infty$, $I : \mathcal{Z} \to [0,\infty]$ is a lower semicontinuous function,

$$\liminf_{N\to\infty} \frac{1}{v_N} \log P(Z_N \in O) \geq -\inf_{z\in O} I(z)$$

for all open sets O, and

$$\limsup_{N\to\infty} \frac{1}{v_N} \log P(Z_N \in C) \leq -\inf_{z\in C} I(z)$$

for all closed sets C. A rate function I is said to be good if all its level sets $\{\{z \in \mathcal{Z} : I(z) \leq \eta\} : \eta \geq 0\}$ are compact.

The main large deviation tool used in the proofs of this paper is the Gärtner Ellis Theorem (see e.g. Theorem 2.3.6 in [17]). In this theorem we have $\mathcal{Z} = \mathbb{R}^m$ for some $m \geq 1$; thus, from now on, we set $\boldsymbol{a} \cdot \boldsymbol{b} := \sum_{j=1}^m a_j b_j$ for two generic vectors $\boldsymbol{a} = (a_1, \ldots, a_m)$ and $\boldsymbol{b} = (b_1, \ldots, b_m)$ of $\mathbb{R}^m$. Before recalling its statement we 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 [17]) if the interior $\mathcal{D}_f^\circ$ of $\mathcal{D}_f := \{\boldsymbol{\gamma} \in \mathbb{R}^m : f(\boldsymbol{\gamma}) < \infty\}$ is non-empty, f is differentiable throughout $\mathcal{D}_\Lambda^\circ$, and f is steep, i.e. $\lim_{h\to\infty} \|\nabla f(\boldsymbol{\gamma}_h)\| = \infty$ whenever $\{\boldsymbol{\gamma}_h : h \geq 1\}$ is a sequence in $\mathcal{D}_f^\circ$ converging to some boundary point of $\mathcal{D}_f^\circ$.

Theorem 1 (Gärtner Ellis Theorem)**.** *Let $\{\boldsymbol{Z}_N : N \geq 1\}$ be a sequence of $\mathbb{R}^m$-valued random variables such that there exists the function $\Lambda : \mathbb{R}^m \to [-\infty, \infty]$ defined by*

$$\Lambda(\boldsymbol{\gamma}) := \lim_{N\to\infty} \frac{1}{v_N} \log \mathbb{E}\left[e^{v_N \boldsymbol{\gamma}\cdot \boldsymbol{Z}_N}\right] \text{ for all } \boldsymbol{\gamma} := (\gamma_1, \ldots, \gamma_m) \in \mathbb{R}^m.$$

Assume that the origin $\mathbf{0} = (0, \ldots, 0) \in \mathbb{R}^m$ belongs to the interior of $\mathcal{D}_\Lambda := \{\boldsymbol{\gamma} \in \mathbb{R}^m : \Lambda(\boldsymbol{\gamma}) < \infty\}$. Then, if Λ is essentially smooth and lower semi-continuous, then $\{\boldsymbol{Z}_N : N \geq 1\}$ satisfies the LDP with speed v_N and good rate function Λ^* defined by $\Lambda^*(\boldsymbol{y}) := \sup_{\boldsymbol{\gamma}\in\mathbb{R}^m}\{\boldsymbol{\gamma}\cdot\boldsymbol{y} - \Lambda(\boldsymbol{\gamma})\}$.

In our applications we always have $\mathcal{D}_\Lambda = \mathbb{R}^m$, and therefore the steepness always holds vacuously. In the following remark we briefly discuss the convergence of $\{\boldsymbol{Z}_N : N \geq 1\}$ in Theorem 1.

Remark 2. *One can check that, when we can apply the Gärtner Ellis Theorem, the rate function $\Lambda^*(\boldsymbol{y})$ uniquely vanishes at $\boldsymbol{y} = \boldsymbol{y}_0$, where $\boldsymbol{y}_0 := \nabla\Lambda(\mathbf{0})$. Then, if we consider the notation*

$$B_\delta(\boldsymbol{y}_0) := \{\boldsymbol{y} \in \mathbb{R}^m : \|\boldsymbol{y} - \boldsymbol{y}_0\| < \delta\} \text{ for } \delta > 0,$$

we have $\Lambda^*(B_\delta^c(\boldsymbol{y}_0)) := \inf_{y\in B_\delta^c(\mathbf{y_0})} \Lambda^*(\mathbf{y}) > \mathbf{0}$ and, for all η such that $0 < \eta < \Lambda^*(B_\delta^c(\boldsymbol{y}_0))$, there exists N_0 such that

$$P\left(\|\boldsymbol{Z}_N - \boldsymbol{y}_0\| \geq \delta\right) \leq \exp\left(-v_N(\Lambda^*(B_\delta^c(\boldsymbol{y}_0)) - \eta)\right) \text{ for all } N \geq N_0$$

(this is a consequence of the large deviation upper bound for the closed set $C = B_\delta^c(\boldsymbol{y}_0)$). Thus we can say that $\boldsymbol{Z}_N$ converges in probability to $\boldsymbol{y}_0$ as $N \to \infty$; moreover the convergence is almost sure if $\sum_{N\geq 1} \exp\left(-v_N(\Lambda^*(B_\delta^c(\boldsymbol{y}_0)) - \eta)\right) < \infty$ by a standard application of the Borel Cantelli lemma.

2.2. Random closed sets

To lighten the presentation we shall use similar notation to previous works [8, 25, 26]; in particular, for the reader's convenience, we refer to [26, Section 2] (and references therein) for the mathematical background and more details on the Minkowski content notion and marked point process theory.

We remind here that, given a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$, a *random closed set* Θ in $\mathbb{R}^d$ is a measurable map

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

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

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

where $\mathcal{G}$ and $\mathcal{C}$ are the system of the open and compact subsets of $\mathbb{R}^d$, respectively, while $\mathcal{F}_G := \{F \in \mathbb{F} : F \cap G \neq \emptyset\}$ and $\mathcal{F}^C := \{F \in \mathbb{F} : F \cap C = \emptyset\}$ (e.g., see [22]).

By means of marked point processes, every random closed set Θ in $\mathbb{R}^d$ can be represented as a *germ-grain model* as follows

$$\Theta(\omega) = \bigcup_{(x_i,s_i)\in\Psi(\omega)} x_i + Z(s_i), \qquad \omega \in \Omega, \tag{1}$$

where $\Psi = \{(\xi_i, S_i)\}_{i\in\mathbb{N}}$ is marked point processes in $\mathbb{R}^d$ with marks in a suitable mark space $\mathbf{K}$ so that $Z_i = Z(S_i)$, $i \in \mathbb{N}$, is a random set containing the origin (i.e., $Z : \mathbf{K} \to \mathbb{F}$).

Throughout the paper we assume that Ψ has intensity measure $\Lambda(\mathrm{d}(x,s)) = f(x,s)\mathrm{d}xQ(\mathrm{d}s)$, where Q is a probability measure on $\mathbf{K}$, and with second factorial moment measure $\nu_{[2]}(\mathrm{d}(x,s,y,t)) = g(x,s,y,t)\mathrm{d}x\mathrm{d}yQ_{[2]}(\mathrm{d}(s,t))$, where $Q_{[2]}$ is a probability measure on $\mathbf{K}^2$ (see [16, 21, 15] for general theory of point processes).

To set the notation, we denote by Θ_n any random closed set in $\mathbb{R}^d$ with Hausdorff dimension n, by $\mathrm{disc}f$ the set of the discontinuity points of any function f, by b_n the volume of the unit ball in $\mathbb{R}^n$. We also recall that the parallel set (or, equivalently, the Minkowski enlargement) of $A \subset \mathbb{R}^d$ at distance $r > 0$ is the set defined as $A_{\oplus r} := \{x \in \mathbb{R}^d : \mathrm{dist}(x, A) \leq r\}$. Moreover, we remind that a compact set $A \subset \mathbb{R}^d$ is called *countably $\mathcal{H}^n$-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 $\mathcal{H}^n$-negligible (e.g., see [3] and references therein for a more exhaustive treatment).

As mentioned in the Introduction, whenever the measure $\mathbb{E}[\mathcal{H}^n(\Theta_n \cap \cdot)]$ on $\mathbb{R}^d$ is absolutely continuous with respect to $\mathcal{H}^d$, we denote by λ_{Θ_n} its density, and we call it the *mean density of* Θ_n. It has been proved [26, Proposition 5] that any random closed set Θ_n in $\mathbb{R}^d$ with Hausdorff dimension $n < d$ as in (1) has mean density λ_{Θ_n} given in terms of the intensity measure Λ of its associated marked point process Ψ as follows:

$$\lambda_{\Theta_n}(x) = \int_{\mathbf{K}} \int_{x - Z(s)} f(y,s)\mathcal{H}^n(\mathrm{d}y)Q(\mathrm{d}s), \tag{2}$$

for $\mathcal{H}^d$-a.e $x \in \mathbb{R}^d$, where $-Z(s)$ is the reflection of $Z(s)$ at the origin.

In the sequel we will assume that an i.i.d. random sample $\Theta_n^{(1)}, \ldots, \Theta_n^{(N)}$ is available for the random closed set Θ_n. An approximation of the mean density based on the $\mathcal{H}^d$-measure of the Minkowski enlargement of the random set in question has been provided in [2, 26] under quite general regularity conditions on the grains of Θ_n and on the functions f and g introduced above:

Theorem 3 ([26, Theorem 7])**.** *Let Θ_n be as in* (1) *such that the following assumptions are satisfied:*

(A1) for any $(y,s) \in \mathbb{R}^d \times \mathbf{K}$, $y + Z(s)$ is a countably $\mathcal{H}^n$-rectifiable and compact subset of $\mathbb{R}^d$, such that there exists a closed set $\Xi(s) \supseteq Z(s)$ such that

$\int_{\mathbf{K}} \mathcal{H}^n(\Xi(s))Q(\mathrm{d}s) < \infty$ and

$$\mathcal{H}^n(\Xi(s) \cap B_r(x)) \geq \gamma r^n \qquad \forall x \in Z(s),\ \forall r \in (0,1)$$

for some $\gamma > 0$ independent of y and s;

(A2) for any $s \in \mathbf{K}$, $\mathcal{H}^n(\mathrm{disc}(f(\cdot, s))) = 0$ and $f(\cdot, s)$ is locally bounded such that for any compact $K \subset \mathbb{R}^d$

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

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

(A3) for any $(s, y, t) \in \mathbf{K} \times \mathbb{R}^d \times \mathbf{K}$, $\mathcal{H}^n(\mathrm{disc}(g(\cdot, s, y, t))) = 0$ and $g(\cdot, s, y, t)$ is locally bounded such that for any compact $K \subset \mathbb{R}^d$ and $a \in \mathbb{R}^d$,

$$\mathbf{1}_{(a - Z(t))_{\oplus 1}}(y) \sup_{x \in K_{\oplus \mathrm{diam}(Z(s))}} g(x, s, y, t) \leq \xi_{a,K}(s, y, t)$$

for some $\xi_{a,K}(s, y, t)$ with

$$\int_{\mathbb{R}^d \times \mathbf{K}^2} \mathcal{H}^n(\Xi(s))\xi_{a,K}(s, y, t)\mathrm{d}y Q_{[2]}(\mathrm{d}s, \mathrm{d}t) < \infty.$$

Then

$$\lambda_{\Theta_n}(x) = \lim_{r \downarrow 0} \frac{\mathbb{P}(x \in \Theta_{n_{\oplus r}})}{b_{d-n} r^{d-n}}, \qquad \mathcal{H}^d\text{-a.e. } x \in \mathbb{R}^d. \tag{3}$$

Remark 4. *The measure $\mathcal{H}^n(\Xi(s) \cap \cdot)$ in (A1) plays the same role as the measure ν of Theorem 2.104 in [3]; indeed (A1) might be seen as its stochastic version, and it is often fulfilled with $\Xi(s) = \partial Z(s)$ or $\Xi(s) = \partial Z(s) \cup \widetilde{A}$ for some sufficiently regular random closed set $\widetilde{A}$ (see also [25, Remark 3.6], and [26, Example 1]). Roughly speaking, such an assumption tells us that each possible grain associated to any point of the underling point process $\{\xi_i, \mathbf{K}\}_{i \in \mathbb{N}}$ is sufficiently regular, so that it admits n-dimensional Minkowski content; this explains also why requiring the existence of a constant γ as in (A1) independent on y and s is not too restrictive. Note that the condition $\int_{\mathbf{K}} \mathcal{H}^n(\Xi(s))Q(\mathrm{d}s) < \infty$ means that the $\mathcal{H}^n$-measure of the grains is finite in mean.*

The role of Assumptions (A2) and (A3) is more technical, in order to guarantee an application of the dominated convergence theorem in the proof of the theorem. We may also notice that if $Z(s)$ has a bounded diameter (i.e., $\mathrm{diam}(Z(s)) \leq C \in \mathbb{R}$ for Q-a.e. $s \in \mathbf{K}$), or if f and g are bounded, then Assumptions (A2) and (A3) simplify (see also Remark 9 in [26]).

The above assumptions imply then Eq. (3), which is obtained by a stochastic local version of the n-dimensional Minkowski-content of Θ_n. For the interested reader we refer to [26] for a more detailed discussion on this.

As a byproduct, given an i.i.d. random sample $\{\Theta_n^{(i)}\}_{i \in \mathbb{N}}$ of Θ_n, the following *"Minkowski content"-based estimator* of $\lambda_{\Theta_n}(x)$ has been proposed:

$$\widehat{\lambda}^{\mu,N}_{\Theta_n}(x) := \frac{\sum_{i=1}^{N} \mathbf{1}_{\Theta_n^{(i)} \cap B_{r_N}(x) \neq \emptyset}}{N b_{d-n} r_N^{d-n}}, \tag{4}$$

where r_N is the bandwidth, which depends on sample size N. It is easily checked that $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$ is asymptotically unbiased and weakly consistent, for $\mathcal{H}^d$-a.e. $x \in \mathbb{R}^d$, if r_N is such that

$$\lim_{N\to\infty} r_N = 0 \quad \text{and} \qquad \lim_{N\to\infty} N r_N^{d-n} = \infty. \tag{5}$$

3. General results for Bernoulli variables

We consider a triangular array $\{\boldsymbol{Y}_{i,N} : 1 \le i \le N\}$ of random variables defined on some probability space $(\Omega, \mathfrak{F}, \mathbb{P})$ and taking values in $(\mathbb{R}^m, \mathcal{B}(\mathbb{R}^m))$, with $m \ge 1$. We assume that each random vector $\boldsymbol{Y}_{i,N}$ consists of m components, namely $\boldsymbol{Y}_{i,N} = (Y^{(1)}_{i,N}, \dots, Y^{(m)}_{i,N})$, and that

$$\boldsymbol{Y}_{i,N} \overset{\text{iid}}{\sim} \boldsymbol{Y}_N \quad (\text{for } i \in \{1, \dots, N\}) \tag{6}$$

where the components of $\boldsymbol{Y}_N := (Y^{(1)}_N, \dots, Y^{(m)}_N)$ are (possibly) dependent Bernoulli distributed random variables; more precisely we mean $Y^{(j)}_N \sim \text{Bern}(p^{(j)}_N)$, i.e.

$$P(Y^{(j)}_N = 1) = 1 - P(Y^{(j)}_N = 0) = p^{(j)}_N \quad (\text{for } j \in \{1, \dots, m\}).$$

The aim is to prove asymptotic results for the sequence

$$\left\{ \frac{1}{w_N} \sum_{i=1}^N \boldsymbol{Y}_{i,N} : N \ge 1 \right\} \tag{7}$$

where $\{w_N : N \ge 1\}$ is a sequence of positive numbers such that:

$$\lim_{N\to\infty} w_N = \infty \text{ and } \lim_{N\to\infty} \frac{w_N}{N} = 0; \tag{8}$$

$$\lim_{N\to\infty} \frac{N p^{(j)}_N}{w_N} = \lambda_j \text{ for some } \lambda_j \in (0,\infty), \text{ for all } j \in \{1, \dots, m\}; \tag{9}$$

$$\lim_{N\to\infty} \frac{N\mathbb{P}(Y^{(j)}_N = 1, Y^{(k)}_N = 1)}{w_N} = 0 \quad \forall\, j,k \in \{1, \dots, m\} \text{ such that } j \ne k. \tag{10}$$

Remark 5. *(i) If $m = 1$, Eq. (10) can be neglected and some parts of proofs presented below are simplified.*

(ii) If $\boldsymbol{Y}_N$ has independent components, for $j,k \in \{1, \dots, m\}$ with $j \ne k$ we have

$$\frac{N\mathbb{P}(Y^{(j)}_N = 1, Y^{(k)}_N = 1)}{w_N} = \frac{N p^{(j)}_N p^{(k)}_N}{w_N} = p^{(k)}_N \frac{N p^{(j)}_N}{w_N};$$

thus condition (10) is clearly satisfied by taking into account condition (9) and that $\lim_{n\to\infty} p^{(k)}_N = 0$ (this limit is a consequence of condition (9) and the second limit in (8)).

In view of the following results we introduce some further notation: $\boldsymbol{S}_N := \sum_{i=1}^N \boldsymbol{Y}_{i,N}$, and therefore $\{\boldsymbol{S}_N/w_N : N \geq 1\}$ coincides with the sequence in (7);

$$\mathcal{I}(y;\lambda) := \begin{cases} y \log \frac{y}{\lambda} - y + \lambda & \text{if } y \geq 0 \\ \infty & \text{otherwise,} \end{cases} \quad \text{for } \lambda > 0. \tag{11}$$

We start with the large deviation principle for the sequence in (7).

Theorem 6. *Let $\{\boldsymbol{Y}_{i,N} : 1 \leq i \leq N\}$ be a triangular array of random vectors as above (thus (6), (8), (9) and (10) hold). Then $\{\boldsymbol{S}_N/w_N : N \geq 1\}$ satisfies the LDP with speed function $v_N = w_N$ and good rate function $I_m(\cdot;\boldsymbol{\lambda})$ defined by $I_m(\boldsymbol{y};\boldsymbol{\lambda}) := \sum_{j=1}^m \mathcal{I}(y_j;\lambda_j)$, where $\mathcal{I}(y_j;\lambda_j)$ is as in (11).*

Proof. We want to apply Gärtner Ellis Theorem and we have to prove that

$$\lim_{N\to\infty} \frac{1}{w_N} \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\boldsymbol{S}_N}\right] = \sum_{j=1}^m \lambda_j(e^{\gamma_j} - 1) \text{ for all } \boldsymbol{\gamma} := (\gamma_1, \dots, \gamma_m) \in \mathbb{R}^m. \tag{12}$$

In fact (12) yields the desired LDP with good rate function J_m defined by

$$J_m(\boldsymbol{y}) := \sup_{\boldsymbol{\gamma}\in\mathbb{R}^m} \left\{ \boldsymbol{\gamma}\cdot\boldsymbol{y} - \sum_{j=1}^m \lambda_j(e^{\gamma_j} - 1) \right\}$$

and, as we see, this function coincides with $I_m(\cdot;\boldsymbol{\lambda})$ in the statement of the proposition: for $m = 1$ we can easily check that $J_1(y_1) = \mathcal{I}(y_1;\lambda_1)$; for $m \geq 2$ we have

$$\begin{aligned} J_m(\boldsymbol{y}) &= \sup_{\boldsymbol{\gamma}\in\mathbb{R}^m} \left\{ \sum_{j=1}^m \gamma_j y_j - \sum_{j=1}^m \lambda_j(e^{\gamma_j} - 1) \right\} \\ &\qquad \leq \sum_{j=1}^m \sup_{\gamma_j\in\mathbb{R}} \{\gamma_j y_j - \lambda_j(e^{\gamma_j} - 1)\} = \sum_{i=1}^m \mathcal{I}(y_j;\lambda_j) \end{aligned}$$

and, since we can take m sequences $\{\gamma_j^{(h)} : h \geq 1\}$ such that $\lim_{h\to\infty}\{\gamma_j^{(h)} y_j - \lambda_j(e^{\gamma_j^{(h)}} - 1)\} = \mathcal{I}(y_j;\lambda_j)$ (for $j \in \{1, \dots, m\}$), we also have the inverse inequality $J_m(\boldsymbol{y}) \geq \sum_{i=1}^m \mathcal{I}(y_j;\lambda_j)$ noting that

$$J_m(\boldsymbol{y}) \geq \sum_{j=1}^m \gamma_j^{(h)} y_j - \sum_{j=1}^m \lambda_j(e^{\gamma_j^{(h)}} - 1)$$

and by letting $h \to \infty$.

In the remaining part of the proof we show the validity of (12). By (6) we have

$$\mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\boldsymbol{S}_N}\right] = (M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma}))^N, \tag{13}$$

where $M_{\boldsymbol{Y}_N}$ is the moment generating function of $\boldsymbol{Y}_N$; moreover, if we set $A_{j,N} := \{Y_N^{(j)} = 0\}$ (for $N \geq 1$ and $j \in \{1, \dots, m\}$), we have

$$\begin{aligned} M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma}) = \mathbb{P}\Big(\bigcap_{j=1}^m A_{j,N}\Big) &+ \sum_{j=1}^m e^{\gamma_j}\mathbb{P}\Big(A_{j,N}^c \cap \bigcap_{k\neq j} A_{k,N}\Big) \\ &+ \sum_{j_1<j_2} e^{\gamma_{j_1}+\gamma_{j_2}}\mathbb{P}\Big(A_{j_1,N}^c \cap A_{j_2,N}^c \cap \bigcap_{k\neq j_1,j_2} A_{k,N}\Big) \\ &+ \cdots + e^{\gamma_1+\cdots+\gamma_m}\mathbb{P}\Big(\bigcap_{j=1}^m A_{j,N}^c\Big). \end{aligned} \tag{14}$$

We remark that $\lim_{N\to\infty} M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma}) = 1$; in fact

$$\lim_{N\to\infty} \mathbb{P}\Big(\bigcap_{j=1}^m A_{j,N}\Big) \geq \lim_{N\to\infty}\Big(1 - \sum_{j=1}^m \mathbb{P}(A_{j,N}^c)\Big) = 1 - \lim_{N\to\infty}\sum_{j=1}^m p_N^{(j)} = 1$$

and, for all $\ell \geq 1$, the terms of the type

$$\mathbb{P}\Big(A_{j_1,N}^c \cap \cdots \cap A_{j_\ell,N}^c \cap \bigcap_{k\neq j_1,\dots,j_\ell} A_{k,N}\Big)$$

converge to zero as $N \to \infty$ because

$$\mathbb{P}\Big(A_{j_1,N}^c \cap \cdots \cap A_{j_\ell,N}^c \cap \bigcap_{k\neq j_1,\dots,j_\ell} A_{k,N}\Big) \leq \mathbb{P}(A_{j_1,N}^c) = p_N^{(j_1)} \longrightarrow 0.$$

Therefore we may claim that

$$\frac{1}{w_N}\log\mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\boldsymbol{S}_N}\right] \overset{(13)}{=} \frac{N\log M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma})}{w_N} = \frac{M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma}) - 1 + o(M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma}) - 1)}{w_N/N}. \tag{15}$$

Let us observe now that, for $\ell \geq 2$,

$$\begin{aligned} \lim_{N\to\infty} &\frac{\mathbb{P}\Big(A_{j_1,N}^c \cap \cdots \cap A_{j_\ell,N}^c \cap \bigcap_{k\neq j_1,\dots,j_\ell} A_{k,N}\Big)}{w_N/N} \\ &\leq \lim_{N\to\infty}\frac{\mathbb{P}(A_{j_1,N}^c \cap A_{j_2,N}^c)}{w_N/N} \overset{(10)}{=} 0; \end{aligned} \tag{16}$$

moreover

$$\begin{aligned} 1 - \mathbb{P}\Big(\bigcap_{j=1}^m A_{j,N}\Big) &= \mathbb{P}(A_{1,N}^c \cup \cdots \cup A_{m,N}^c) \\ = \sum_{j=1}^m \mathbb{P}(A_{j,N}^c) &- \sum_{j_1<j_2}\mathbb{P}(A_{j_1,N}^c \cap A_{j_2,N}^c) + \cdots + (-1)^{m+1}\mathbb{P}(A_{1,N}^c \cap \cdots \cap A_{m,N}^c) \end{aligned}$$

by the inclusion–exclusion formula (see [7, pg. 24]), and therefore

$$\lim_{N\to\infty} \frac{1-\mathbb{P}\left(\bigcap_{j=1}^m A_{j,N}\right)}{w_N/N} \overset{(16)}{=} \lim_{N\to\infty} \frac{\sum_{j=1}^m P(A_{j,N}^c)}{w_N/N} \overset{(9)}{=} \sum_{j=1}^m \lambda_j. \tag{17}$$

Note that, with some slight changes, one can also prove that

$$\lim_{N\to\infty} \frac{1-\mathbb{P}\left(\bigcap_{k\neq j} A_{k,N}\right)}{w_N/N} = \sum_{k\neq j} \lambda_k \text{ for all } j \in \{1,\dots,m\}. \tag{18}$$

Finally, by substituting in (15) the expression of $M_{\mathbf{Y}_N}$ in (14), and by taking into account (16), we have

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{w_N} \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\boldsymbol{S}_N}\right] &= \lim_{N\to\infty} \frac{\mathbb{P}\left(\bigcap_{j=1}^m A_{j,N}\right)-1}{w_N/N} \\ &\quad + \lim_{N\to\infty} \frac{\sum_{j=1}^m e^{\gamma_j}\mathbb{P}\left(A_{j,N}^c \cap \bigcap_{k\neq j} A_{k,N}\right)}{w_N/N},\end{aligned}$$

where

$$\lim_{N\to\infty} \frac{\mathbb{P}\left(\bigcap_{j=1}^m A_{j,N}\right)-1}{w_N/N} = -\sum_{j=1}^m \lambda_j$$

by (17), and

$$\begin{aligned}&\frac{\sum_{j=1}^m e^{\gamma_j}\mathbb{P}\left(A_{j,N}^c \cap \bigcap_{k\neq j} A_{k,N}\right)}{w_N/N} \\ &\quad = \frac{\sum_{j=1}^m e^{\gamma_j}\mathbb{P}\left(\bigcap_{k\neq j} A_{k,N}\right) - \sum_{j=1}^m e^{\gamma_j}\mathbb{P}\left(\bigcap_{k=1}^m A_{k,N}\right)}{w_N/N} \\ &\quad = \frac{\sum_{j=1}^m e^{\gamma_j}\left(\mathbb{P}\left(\bigcap_{k\neq j} A_{k,N}\right)-1\right) - \sum_{j=1}^m e^{\gamma_j}\left(\mathbb{P}\left(\bigcap_{k=1}^m A_{k,N}\right)-1\right)}{w_N/N} \\ &\quad \to \sum_{j=1}^m e^{\gamma_j}\left(-\sum_{k\neq j}\lambda_k + \sum_{k=1}^m \lambda_k\right) = \sum_{j=1}^m e^{\gamma_j}\lambda_j \text{ (as } N\to\infty)\end{aligned}$$

by (18) and (17). □

Remark 7. *Assume that $m=1$ for simplicity, and let us denote the Poisson distribution with mean λ by $\mathrm{P}(\lambda)$. It is well-known that, if $\lim_{N\to\infty} Np_N^{(1)} = \lambda_1$ for some $\lambda_1 \in (0,\infty)$, then $S_N^{(1)} = \sum_{i=1}^N Y_{i,N}^{(1)}$ converges weakly to $\mathrm{P}(\lambda_1)$ (as $N\to\infty$). Then Theorem 6 (for $m=1$) describes what happens if we divide both $Np_N^{(1)}$ (in the limit above) and the sums $\{S_N^{(1)} : N \geq 1\}$ by w_N; in fact*

$S_N^{(1)}/w_N$ converges to λ_1 (as $N \to \infty$) in probability (see Remark 2) and, for the rate function $I_1(\cdot;\lambda_1)$ in Theorem 6, we can say that $I_1(y_1;\lambda_1)$ is the relative entropy of $\mathrm{P}(y_1)$ with respect to $\mathrm{P}(\lambda_1)$ (when $y_1 \geq 0$).

Now we prove the moderate deviation result for the sequence in (7); a brief explanation of this terminology is given in Remark 11.

Theorem 8. *Consider the same assumptions of Theorem 6. Then, for any sequence of positive numbers $\{a_N : N \geq 1\}$ such that $\lim_{N\to\infty} a_N = 0$ and $\lim_{N\to\infty} w_N a_N = \infty$, $\left\{\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N/a_N}} : N \geq 1\right\}$ satisfies the LDP with speed function $v_N = 1/a_N$ and good rate function $\tilde{I}_m(\cdot;\boldsymbol{\lambda})$ defined by $\tilde{I}_m(\boldsymbol{y};\boldsymbol{\lambda}) := \sum_{j=1}^m \frac{y_j^2}{2\lambda_j}$.*

Proof. We want to apply Gärtner Ellis Theorem and we have to prove that

$$\lim_{N\to\infty} a_N \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N a_N}}}\right] = \sum_{j=1}^m \lambda_j \frac{\gamma_j^2}{2} \text{ for all } \boldsymbol{\gamma} := (\gamma_1, \dots, \gamma_m) \in \mathbb{R}^m. \quad (19)$$

In fact (19) yields the desired LDP with good rate function $\tilde{J}_m$ defined by

$$\tilde{J}_m(\boldsymbol{y}) := \sup_{\boldsymbol{\gamma}\in\mathbb{R}^m} \left\{\boldsymbol{\gamma}\cdot\boldsymbol{y} - \sum_{j=1}^m \lambda_j \frac{\gamma_j^2}{2}\right\}$$

and, as we see, this function coincides with $\tilde{I}_m(\cdot;\boldsymbol{\lambda})$ in the statement of the proposition: for $m = 1$ we can easily check that $\tilde{J}_1(y_1) = \frac{y_1^2}{2\lambda_1}$; for $m \geq 2$ we have

$$\tilde{J}_m(\boldsymbol{y}) = \sup_{\boldsymbol{\gamma}\in\mathbb{R}^m} \left\{\sum_{j=1}^m \gamma_j y_j - \sum_{j=1}^m \lambda_j \frac{\gamma_j^2}{2}\right\} \leq \sum_{j=1}^m \sup_{\gamma_j\in\mathbb{R}} \left\{\gamma_j y_j - \lambda_j \frac{\gamma_j^2}{2}\right\} = \sum_{i=1}^m \frac{y_j^2}{2\lambda_j}$$

and, since we can take m sequences $\{\gamma_j^{(h)} : h \geq 1\}$ such that $\lim_{h\to\infty}\{\gamma_j^{(h)} y_j - \lambda_j \frac{(\gamma_j^{(h)})^2}{2}\} = \frac{y_j^2}{2\lambda_j}$ (for $j \in \{1,\dots,m\}$), we also have the inverse inequality $\tilde{J}_m(\boldsymbol{y}) \geq \sum_{i=1}^m \frac{y_j^2}{2\lambda_j}$ noting that

$$\tilde{J}_m(\boldsymbol{y}) \geq \sum_{j=1}^m \gamma_j^{(h)} y_j - \sum_{j=1}^m \lambda_j \frac{(\gamma_j^{(h)})^2}{2}$$

and by letting $h \to \infty$.

In the remaining part of the proof we show that (19) holds. In what follows we consider some notation already introduced in the proof of Theorem 6, i.e. $M_{\boldsymbol{Y}_N}$ is the moment generating function of $\boldsymbol{Y}_N$ and we set $A_{j,N} := \{Y_N^{(j)} = 0\}$ (for $N \geq 1$ and $j \in \{1,\dots,m\}$). By (6) and after some computations we get

$$\log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N a_N}}}\right] = N\left(\log M_{\boldsymbol{Y}_N}(\boldsymbol{\gamma}/\sqrt{w_N a_N}) - \boldsymbol{\gamma}\cdot\frac{\boldsymbol{p}_N}{\sqrt{w_N a_N}}\right),$$

where $\boldsymbol{p}_N = (p_N^{(1)}, \dots, p_N^{(m)})$; thus

$$\begin{aligned} & a_N \log \mathbb{E}\left[e^{\boldsymbol{\gamma} \cdot \frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N a_N}}} \right] \\ & \qquad = a_N N \left(\log \mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}} \sum_{j=1}^m \gamma_j Y_N^{(j)} \right) \right] - \sum_{j=1}^m \frac{\gamma_j p_N^{(j)}}{\sqrt{w_N a_N}} \right), \end{aligned} \tag{20}$$

where

$$\begin{aligned} \mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}} \sum_{j=1}^m \gamma_j Y_N^{(j)} \right) \right] &= \mathbb{E}\left[\prod_{j=1}^m \exp\left(\frac{1}{\sqrt{w_N a_N}} \gamma_j \mathbb{1}_{A_{j,N}^c} \right) \right] \\ &= \mathbb{E}\left[\prod_{j=1}^m \left(\mathbb{1}_{A_{j,N}^c} (e^{\gamma_j/\sqrt{w_N a_N}} - 1) + 1 \right) \right]. \end{aligned}$$

The product inside the expected value in the above equation may be rewritten as

$$\begin{aligned} & \left(\mathbb{1}_{A_{1,N}^c} (e^{\gamma_1/\sqrt{w_N a_N}} - 1) + 1 \right) \cdots \left(\mathbb{1}_{A_{m,N}^c} (e^{\gamma_m/\sqrt{w_N a_N}} - 1) + 1 \right) \\ & = 1 + \sum_{j=1}^m \mathbb{1}_{A_{j,N}^c} (e^{\gamma_j/\sqrt{w_N a_N}} - 1) + \sum_{j_1 < j_2} \mathbb{1}_{A_{j_1,N}^c} \mathbb{1}_{A_{j_2,N}^c} \prod_{i=1}^2 (e^{\gamma_{j_i}/\sqrt{w_N a_N}} - 1) \\ & \qquad + \sum_{j_1 < j_2 < j_3} \mathbb{1}_{A_{j_1,N}^c} \mathbb{1}_{A_{j_2,N}^c} \mathbb{1}_{A_{j_3,N}^c} \prod_{i=1}^3 (e^{\gamma_{j_i}/\sqrt{w_N a_N}} - 1) \\ & \qquad\qquad + \cdots + \prod_{j=1}^m \mathbb{1}_{A_{j,N}^c} (e^{\gamma_j/\sqrt{w_N a_N}} - 1). \end{aligned}$$

Now, by taking the expectation of the previous expression and by substituting $e^{\gamma_{j_i}/\sqrt{w_N a_N}} - 1$ with its asymptotic expansion, we get

$$\begin{aligned} & \mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}} \sum_{j=1}^m \gamma_j Y_N^{(j)} \right) \right] \\ & = 1 + \sum_{j=1}^m \mathbb{P}(A_{j,N}^c) \left(\frac{\gamma_j}{\sqrt{w_N a_N}} + \frac{1}{2} \frac{\gamma_j^2}{w_N a_N} + o\left(\frac{1}{w_N a_N} \right) \right) \\ & + \sum_{j_1 < j_2} \mathbb{P}(A_{j_1,N}^c \cap A_{j_2,N}^c) \prod_{i=1}^2 \left(\frac{\gamma_{j_i}}{\sqrt{w_N a_N}} + \frac{1}{2} \frac{\gamma_{j_i}^2}{w_N a_N} + o\left(\frac{1}{w_N a_N} \right) \right) + \\ & \cdots + \mathbb{P}(A_{1,N}^c \cap \cdots \cap A_{m,N}^c) \prod_{j=1}^m \left(\frac{\gamma_j}{\sqrt{w_N a_N}} + \frac{1}{2} \frac{\gamma_j^2}{w_N a_N} + o\left(\frac{1}{w_N a_N} \right) \right). \end{aligned} \tag{21}$$

Then we may claim that

$$\lim_{N\to\infty} \mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}}\sum_{j=1}^m \gamma_j Y_N^{(j)}\right)\right] = 1$$

because, as $N \to \infty$, we have

$$\mathbb{P}(A_{j_1,N}^c \cap \cdots \cap A_{j_\ell,N}^c) \leq \mathbb{P}(A_{j_1,N}^c) = p_N^{(j_1)} \to 0$$

for $\ell \geq 1$, and $w_N a_N \to \infty$.

Now, returning to consider the expression in (20), and by considering a first order Taylor expansion of $\log\left(1 + \left(\mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}}\sum_{j=1}^m \gamma_j Y_N^{(j)}\right)\right] - 1\right)\right)$, we obtain

$$\begin{aligned} &a_N \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N a_N}}}\right] \\ &\quad = a_N N\left(\mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}}\sum_{j=1}^m \gamma_j Y_N^{(j)}\right)\right] - 1 + R_N - \sum_{j=1}^m \frac{\gamma_j p_N^{(j)}}{\sqrt{w_N a_N}}\right) \end{aligned}$$

where

$$R_N = o\left(\mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}}\sum_{j=1}^m \gamma_j Y_N^{(j)}\right)\right] - 1\right).$$

By replacing now the expression given in (21) it follows

$$\begin{aligned} a_N \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N a_N}}}\right] &= a_N N\left(\sum_{j=1}^m \mathbb{P}(A_{j,N}^c)\left(\frac{1}{2}\frac{\gamma_j^2}{w_N a_N} + o\left(\frac{1}{w_N a_N}\right)\right)\right. \\ &+ \sum_{j_1<j_2} \mathbb{P}(A_{j_1,N}^c \cap A_{j_2,N}^c)\prod_{i=1}^2\left(\frac{\gamma_{j_i}}{\sqrt{w_N a_N}} + \frac{1}{2}\frac{\gamma_{j_i}^2}{w_N a_N} + o\left(\frac{1}{w_N a_N}\right)\right) + \cdots \\ &\cdots + \mathbb{P}(A_{1,N}^c \cap \cdots \cap A_{m,N}^c)\prod_{j=1}^m\left(\frac{\gamma_j}{\sqrt{w_N a_N}} + \frac{1}{2}\frac{\gamma_j^2}{w_N a_N} + o\left(\frac{1}{w_N a_N}\right)\right) + R_N\left.\vphantom{\sum_{j=1}^m}\right). \end{aligned}$$

Now we remark that the first terms discarded for R_N, which concern $\left(\mathbb{E}\left[\exp\left(\frac{1}{\sqrt{w_N a_N}}\sum_{j=1}^m \gamma_j Y_N^{(j)}\right)\right] - 1\right)^2$, are

$$\frac{\mathbb{P}(A_{j_1,N}^c)\mathbb{P}(A_{j_2,N}^c)}{w_N a_N}\gamma_{j_1}\gamma_{j_2} = \frac{N p_N^{(j_1)} p_N^{(j_2)}}{w_N a_N N}\gamma_{j_1}\gamma_{j_2} \sim \frac{\lambda_{j_1} p_N^{(j_2)}}{a_N N}\gamma_{j_1}\gamma_{j_2} = o\left(\frac{1}{a_N N}\right),$$

as $N \to \infty$, for all $j_1, j_2 \in \{1, \ldots, m\}$; therefore $R_N = o\left(\frac{1}{a_N N}\right)$. Moreover, by (10), for all $\ell \geq 2$ we have

$$\mathbb{P}(A_{j_1,N}^c \cap \cdots \cap A_{j_\ell,N}^c) \leq \mathbb{P}(A_{j_1,N}^c \cap A_{j_2,N}^c) = o\left(\frac{w_N}{N}\right),$$

and then $o\left(\frac{w_N}{N}\right)\frac{1}{w_N a_N} = o\left(\frac{1}{a_N N}\right)$. In conclusion we obtain

$$\begin{aligned}
&a_N \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N a_N}}}\right] \\
&= a_N N \left(\sum_{j=1}^m \mathbb{P}(A_{j,N}^c)\left(\frac{1}{2}\frac{\gamma_j^2}{w_N a_N} + o\left(\frac{1}{w_N a_N}\right)\right) + o\left(\frac{1}{a_N N}\right)\right) \\
&= \sum_{j=1}^m \frac{N p_N^{(j)}}{w_N}\cdot\frac{\gamma_j^2}{2} + \sum_{j=1}^m \frac{N p_N^{(j)}}{w_N}\cdot w_N a_N \cdot o\left(\frac{1}{w_N a_N}\right) \\
&\qquad + a_N N \cdot o\left(\frac{1}{a_N N}\right) \to \sum_{j=1}^m \lambda_j \frac{\gamma_j^2}{2}
\end{aligned}$$

as $N \to \infty$, by (9). □

Remark 9. *The hypothesis $a_N \to 0$ is required to have $v_N \to \infty$ as $N \to \infty$. On the other hand this condition is not needed to prove the limit (19); in fact, as it happens in the proof of Theorem 10 below, the limit (19) holds even if $a_N = 1$ for all $N \geq 1$.*

We conclude with an asymptotic Normality result. In view of this we use the symbol $\mathrm{N}_m(\mathbf{0}, \Sigma)$ for the centered m-variate Normal distribution with covariance matrix Σ, where Σ is the diagonal matrix with entries $\lambda_1, \dots, \lambda_m$, i.e.

$$\Sigma := \begin{pmatrix} \lambda_1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & \lambda_{m-1} & 0 \\ 0 & \dots & 0 & 0 & \lambda_m \end{pmatrix}. \tag{22}$$

Theorem 10. *Consider the same assumptions of Theorem 6. Then $\left\{\frac{(\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N])}{\sqrt{w_N}} : N \geq 1\right\}$ converges weakly (as $N \to \infty$) to $\mathrm{N}_m(\mathbf{0}, \Sigma)$.*

Proof. The proof is a consequence of the following limit

$$\lim_{N\to\infty} \log \mathbb{E}\left[e^{\boldsymbol{\gamma}\cdot\frac{\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N]}{\sqrt{w_N}}}\right] = \sum_{j=1}^m \lambda_j^2 \frac{\gamma_j^2}{2} \text{ for all } \boldsymbol{\gamma} := (\gamma_1, \dots, \gamma_m) \in \mathbb{R}^m,$$

i.e. (19) with $a_N = 1$ for all $N \geq 1$. This limit holds because all the computations (to get (19) in the proof of Theorem 8) still work well even if $a_N = 1$ for all $N \geq 1$. □

Remark 11. *Typically moderate deviations fill the gap between a convergence to zero (or the null vector) of centered random variables, and a weak convergence to a (possibly multivariate) centered Normal distribution. This is what happens in Theorem 8 for $\left\{\frac{(\boldsymbol{S}_N - \mathbb{E}[\boldsymbol{S}_N])}{\sqrt{w_N/a_N}} : N \geq 1\right\}$: we mean the convergence to the null* *vector of* $\mathbb{R}^m$ *when* $a_N = 1/w_N$ *for all* $N \geq 1$ *(thus* $a_N \to 0$ *holds, but* $a_N w_N \to \infty$ *fails), which is yielded by Theorem 6, and the weak convergence to* $\mathrm{N}_m(\mathbf{0}, \Sigma)$ *when* $a_N = 1$ *for all* $N \geq 1$ *(thus* $a_N w_N \to \infty$ *holds, but* $a_N \to 0$ *fails) in Theorem 10.*

4. Applications in Stochastic Geometry

In this section we apply the general results showed in the previous one to a multivariate version of the "Minkowski content"-based estimator defined in (4), say $\widehat{\boldsymbol{\lambda}}^{\mu,N}$. Namely, in Section 4.1 we first provide a sufficient condition to guarantee Eq. (10); then a series of results on $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ will follow as corollaries of the theorems above, among them the strong consistency of $\widehat{\boldsymbol{\lambda}}^{\mu,N}$. In Section 4.2 we study the asymptotical distribution of $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ to get confidence regions for the m-dimensional vector $(\lambda_{\Theta_n}(x_1), \dots, \lambda_{\Theta_n}(x_m))$ of mean densities of Θ_n in $(x_1, \dots, x_m) \in (\mathbb{R}^d)^m$. It will emerge the importance of choosing a suitable optimal bandwidth r_N satisfying condition (5), the same for any component of the vector $(\lambda_{\Theta_n}(x_1), \dots, \lambda_{\Theta_n}(x_m))$.

4.1. Statistical properties of the "Minkowski content"-based estimator

Let Θ_n be a random closed set with integer Hausdorff dimension $n < d$ in $\mathbb{R}^d$, satisfying the assumptions of Theorem 3, and let $\Theta_n^{(1)}, \dots, \Theta_n^{(N)}$ be an i.i.d. random sample for Θ_n. We consider the "Minkowski content"-based estimator $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$ for the mean density of Θ_n at a point $x \in \mathbb{R}^d$, defined in (4), with bandwidth r_N satisfying condition (5). Given m distinct points $x_1, \dots, x_m \in \mathbb{R}^d$, we can define the *multivariate "Minkowski content"-based estimator*

$$\widehat{\boldsymbol{\lambda}}^{\mu,N} := (\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_1), \dots, \widehat{\lambda}^{\mu,N}_{\Theta_n}(x_m)). \tag{23}$$

Since $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)$ is asymptotically unbiased and weakly consistent for $\lambda_{\Theta_n}(x_j)$, as $j \in \{1, \dots, m\}$, we have that $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ is a good estimator for

$$\boldsymbol{\lambda}_{\Theta_n} := (\lambda_{\Theta_n}(x_1), \dots, \lambda_{\Theta_n}(x_m)).$$

Moreover, observe that $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ can be rewritten in the following way

$$\widehat{\boldsymbol{\lambda}}^{\mu,N} = \frac{1}{w_N} \sum_{i=1}^{N} \boldsymbol{Y}_{i,N},$$

where we set

$$w_N = N b_{d-n} r_N^{d-n}$$

and

$$\boldsymbol{Y}_{i,N} = (\mathbb{1}_{\Theta^{(i)}_{n \oplus r_N}}(x_1), \dots, \mathbb{1}_{\Theta^{(i)}_{n \oplus r_N}}(x_m)).$$

Therefore the multivariate "Minkowski content"-based estimator $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ is of type (7).

In particular $\boldsymbol{Y}_{i,N} \overset{\text{iid}}{\sim} \boldsymbol{Y}_N = (\mathbb{1}_{\Theta_{n_{\oplus r_N}}}(x_1), \dots, \mathbb{1}_{\Theta_{n_{\oplus r_N}}}(x_m))$ with $\mathbb{1}_{\Theta_{n_{\oplus r_N}}}(x_j) \sim \text{Bern}(p_N^{(j)})$, and $p_N^{(j)} := \mathbb{P}(x_j \in \Theta_{n_{\oplus r_N}})$. Observe that condition (8) immediately follows by (5), whereas (9) is fulfilled for a.e. $x_j \in \mathbb{R}^d$ by replacing x with x_j in Eq. (3). Then, in order to apply the results proved in Section 3 with the aim to infer on $\widehat{\boldsymbol{\lambda}}^{\mu,N}$, it remains to show that also assumption (10) is satisfied.

The next lemma provides a sufficient condition on the random closed set Θ_n for the validity of (10). Note that the condition (24) in the statement is satisfied if the points $x_i, \dots, x_m$ and Q-almost every realization $Z(s)$ of the grains of Θ are such the $(x_i - Z(s)) \cap (x_j - Z(s))$ has null $\mathcal{H}^n$-measure. For instance, in $\mathbb{R}^2$ with $n = 1$, if $x_1 = (0,0)$, $x_2 = (1,0)$ such a condition is not satisfied if $Z(s) = \partial[0,1]^2$, but it is if $Z(s)$ is the boundary of any ball. By taking into account that usually Q is assumed to be continuous in applications, it is quite intuitive that the condition (24) is generally fulfilled for any fixed m-tuple of points $(x_1, \dots, x_m) \in (\mathbb{R}^d)^m$.

Lemma 12. *Let Θ_n be a random closed set as above. If for a m-tuple of points $(x_1, \dots, x_m) \in (\mathbb{R}^d)^m$*

$$\mathcal{H}^n((x_i - Z(s)) \cap (x_j - Z(s))) = 0 \quad \forall i \neq j, \quad \textit{for } Q\textit{-almost all } s \in \boldsymbol{K} \tag{24}$$

then (10) *is satisfied.*

Proof. Define the set $A_j := \{x_j \notin \Theta_{n_{\oplus r}}\}$ for every $j = 1, \dots, m$; then we have to prove that

$$\lim_{r \to 0} \frac{\mathbb{P}(A_j^c \cap A_k^c)}{b_{d-n} r^{d-n}} = 0 \qquad \text{for any } j \neq k. \tag{25}$$

We introduce the random variables $W_r^{(j)}$ counting the number of enlarged grains which cover the point x_j, namely

$$W_r^{(j)} := \sum_{(y_i, s_i) \in \Psi} \mathbb{1}_{(y_i + Z(s_i))_{\oplus r}}(x_j).$$

To lighten the notation, without loss of generalization, we will prove (25) with $j = 1$ and $k = 2$. To do this, let us observe that

$$\begin{aligned} \mathbb{P}(A_1^c \cap A_2^c) &= \mathbb{P}(x_1 \in \Theta_{n_{\oplus r}}, x_2 \in \Theta_{n_{\oplus r}}) = \mathbb{P}(W_r^{(1)} > 0, W_r^{(2)} > 0) \\ &= \mathbb{P}(W_r^{(1)} W_r^{(2)} > 0) = \sum_{i \geq 1} \mathbb{P}(W_r^{(1)} W_r^{(2)} = i) \leq \mathbb{E}[W_r^{(1)} W_r^{(2)}]. \end{aligned} \tag{26}$$

Since the marginal process of Φ is *simple* (i.e., every point y_i has multiplicity one), the expectation on the right hand side of (26) can be written as follows:

$$\mathbb{E}[W_r^{(1)} W_r^{(2)}] = \mathbb{E}\Big[\sum_{(y_i, s_i), (y_j, s_j) \in \Phi} \mathbb{1}_{(y_i + Z(s_i))_{\oplus r}}(x_1) \mathbb{1}_{(y_j + Z(s_j))_{\oplus r}}(x_2) \Big]$$

$$
\begin{aligned}
&= \underbrace{\mathbb{E}\Big[\sum_{(y_i,s_i)\in\Phi} \mathbb{1}_{(y_i+Z(s_i))_{\oplus r}}(x_1)\mathbb{1}_{(y_i+Z(s_i))_{\oplus r}}(x_2)\Big]}_{=:\mathcal{E}_1} \\
&\quad + \underbrace{\mathbb{E}\Big[\sum_{\substack{(y_i,s_i),(y_j,s_j)\in\Phi,\\ y_i\neq y_j}} \mathbb{1}_{(y_i+Z(s_i))_{\oplus r}}(x_1)\mathbb{1}_{(y_j+Z(s_j))_{\oplus r}}(x_2)\Big]}_{=:\mathcal{E}_2}. \qquad (27)
\end{aligned}
$$

A direct application of Campbell's theorem (e.g., see [4, p. 28]) implies

$$
\begin{aligned}
\mathcal{E}_1 &= \int_{\boldsymbol{K}}\int_{\mathbb{R}^d} \mathbb{1}_{(y+Z(s))_{\oplus r}}(x_1)\mathbb{1}_{(y+Z(s))_{\oplus r}}(x_2) f(y,s)\mathrm{d}yQ(\mathrm{d}s) \\
&= \int_{\boldsymbol{K}}\int_{\mathbb{R}^d} \mathbb{1}_{(x_1-Z(s))_{\oplus r}\cap(x_2-Z(s))_{\oplus r}}(y) f(y,s)\mathrm{d}yQ(\mathrm{d}s) \\
&= \int_{\boldsymbol{K}}\int_{\mathbb{R}^d} \Big\{\mathbb{1}_{(x_1-Z(s))_{\oplus r}}(y) + \mathbb{1}_{(x_2-Z(s))_{\oplus r}}(y) \\
&\qquad\qquad - \mathbb{1}_{(x_1-Z(s))_{\oplus r}\cup(x_2-Z(s))_{\oplus r}}(y)\Big\} f(y,s)\mathrm{d}yQ(\mathrm{d}s) \\
&= \sum_{j=1}^{2}\int_{\boldsymbol{K}}\int_{(x_j-Z(s))_{\oplus r}} f(y,s)\mathrm{d}yQ(\mathrm{d}s) \\
&\qquad\qquad - \int_{\boldsymbol{K}}\int_{[(x_1-Z(s))\cup(x_2-Z(s))]_{\oplus r}} f(y,s)\mathrm{d}yQ(\mathrm{d}s).
\end{aligned}
$$

Theorem 3.5 in [25] applies to both $(x_j - Z(s))_{\oplus r}$ and $[(x_1 - Z(s)) \cup (x_2 - Z(s))]_{\oplus r}$; moreover Assumptions (A1) and (A2) (see Theorem 3), together with the dominated convergence theorem, allow us to claim that

$$
\begin{aligned}
\lim_{r\to 0}\frac{\mathcal{E}_1}{b_{d-n}r^{d-n}} &= \sum_{j=1}^{2}\int_{\boldsymbol{K}}\int_{(x_j-Z(s))} f(y,s)\mathcal{H}^n(\mathrm{d}y)Q(\mathrm{d}s) \\
&\qquad - \int_{\boldsymbol{K}}\int_{[(x_1-Z(s))\cup(x_2-Z(s))]} f(y,s)\mathcal{H}^n(\mathrm{d}y)Q(\mathrm{d}s) = 0,
\end{aligned}
$$

because the assumption $\mathcal{H}^n((x_1 - Z(s)) \cap (x_2 - Z(s))) = 0$ for Q-almost all $s \in \boldsymbol{K}$ implies that

$$
\begin{aligned}
&\int_{\boldsymbol{K}}\int_{[(x_1-Z(s))\cup(x_2-Z(s))]} f(y,s)\mathcal{H}^n(\mathrm{d}y)Q(\mathrm{d}s) \\
&\qquad\qquad = \sum_{j=1}^{2}\int_{\boldsymbol{K}}\int_{(x_j-Z(s))} f(y,s)\mathcal{H}^n(\mathrm{d}y)Q(\mathrm{d}s).
\end{aligned}
$$

We argue similarly for the term $\mathcal{E}_2$ in (27). By the definition of second factorial moment measure (e.g., see [15]) we have

$$\begin{aligned}\frac{\mathcal{E}_2}{b_{d-n}r^{d-n}} &= \frac{1}{b_{d-n}r^{d-n}}\int_{\boldsymbol{K}^2}\int_{\mathbb{R}^{2d}} \mathbb{1}_{(y_1+Z(s_1))_{\oplus r}}(x_1)\mathbb{1}_{(y_2+Z(s_2))_{\oplus r}}(x_2) \\ &\qquad\qquad\times g(y_1,s_1,y_2,s_2)\mathrm{d}y_1\mathrm{d}y_2 Q_{[2]}(\mathrm{d}s_1,\mathrm{d}s_2) \\ &= \int_{\boldsymbol{K}^2}\int_{\mathbb{R}^d}\frac{\mathbb{1}_{(x_2-Z(s_2))_{\oplus r}}(y_2)}{b_{d-n}r^{d-n}}\int_{\mathbb{R}^d}\mathbb{1}_{(x_1-Z(s_1))_{\oplus r}}(y_1) \\ &\qquad\qquad\times g(y_1,s_1,y_2,s_2)\mathrm{d}y_1\mathrm{d}y_2 Q_{[2]}(\mathrm{d}s_1,\mathrm{d}s_2)\end{aligned}$$

Let us observe that

$$\lim_{r\to 0}\mathbb{1}_{(x_2-Z(s_2))_{\oplus r}}(y_2)=0 \quad \text{for any } (y_2,s_2)\in\mathbb{R}^d\times\boldsymbol{K}$$

because $Z(s)$ is a lower dimensional set for any $s\in\boldsymbol{K}$ by (A1); besides, thanks to assumptions (A1) and (A3), Theorem 3.5 in [25] applies now with $\mu = g(\cdot,s,y,t)\mathcal{H}^d$, so that

$$\begin{aligned}\lim_{r\to 0}\frac{1}{b_{d-n}r^{d-n}}&\int_{\mathbb{R}^d}\mathbb{1}_{(x_1-Z(s_1))_{\oplus r}}(y_1)g(y_1,s_1,y_2,s_2)\mathrm{d}y_1 \\ &= \int_{(x_1-Z(s_1))} g(y_1,s_1,y_2,s_2)\mathcal{H}^n(\mathrm{d}y_1)\end{aligned}$$

for all $(s_1,s_2,y_2)\in\boldsymbol{K}^2\times\mathbb{R}^d$. Such a limit is finite being g locally bounded and $\mathcal{H}^n(Z(s))<\infty$ for any $s\in\boldsymbol{K}$ by (A1). Therefore, for all $(y_2,s_1,s_2)\in\mathbb{R}^d\times\boldsymbol{K}^2$,

$$\lim_{r\to 0}\frac{\mathbb{1}_{(x_2-Z(s_2))_{\oplus r}}(y_2)}{b_{d-n}r^{d-n}}\int_{\mathbb{R}^d}\mathbb{1}_{(x_1-Z(s_1))_{\oplus r}}(y_1)g(y_1,s_1,y_2,s_2)\mathrm{d}y_1 = 0$$

while, for all $(y_2,s_1,s_2)\in\mathbb{R}^d$ and $r\le 1$,

$$\begin{aligned}&\frac{\mathbb{1}_{(x_2-Z(s_2))_{\oplus r}}(y_2)}{b_{d-n}r^{d-n}}\int_{\mathbb{R}^d}\mathbb{1}_{(x_1-Z(s_1))_{\oplus r}}(y_1)g(y_1,s_1,y_2,s_2)\mathrm{d}y_1 \\ &\quad\le \frac{\mathbb{1}_{(x_2-Z(s_2))_{\oplus 1}}(y_2)}{b_{d-n}r^{d-n}}\sup_{y_1\in(x_1-Z(s_1))_{\oplus r}} g(y_1,s_1,y_2,s_2)\mathcal{H}^d((x_1-Z(s_1))_{\oplus 1}) \\ &\quad\le \mathbb{1}_{(x_2-Z(s_2))_{\oplus 1}}(y_2)\sup_{y_1\in(B_1(x_1))_{\oplus\mathrm{diam}(Z(s_1))}} g(y_1,s_1,y_2,s_2)\frac{\mathcal{H}^d((x_1-Z(s_1))_{\oplus 1})}{b_{d-n}r^{d-n}} \\ &\quad\le \mathbb{1}_{(x_2-Z(s_2))_{\oplus 1}}(y_2)\xi_{x_1,B_1(x_1)}(s_1,y_2,s_2)\mathcal{H}^n(\Xi(s_1))\frac{2^n4^db_d}{\gamma b_{d-n}}. \end{aligned}\tag{28}$$

The last inequality follows by Remark 4 in [26] which guarantees that

$$\mathcal{H}^d(Z(s)_{\oplus R})\le\mathcal{H}^n(\Xi(s))\gamma^{-1}2^n4^db_dR^{d-n} \qquad \text{if } R<2.$$

Since the dominating function in (28) has finite integral in $\boldsymbol{K}^2\times\mathbb{R}^d$ by (A3), the dominated convergence theorem can be applied to conclude that:

$$\lim_{r\to 0}\frac{\mathcal{E}_2}{b_{d-n}r^{d-n}}=0.$$

Then the assertion follows. $\square$

By combining now Lemma 12 and the results in Section 3, we easily get some asymptotic results for the multivariate "Minkowski content"-based estimator.

To lighten the notation, we shall write $\boldsymbol{\lambda}_{\Theta_n} = (\lambda_1, \ldots, \lambda_m)$ instead of $\boldsymbol{\lambda}_{\Theta_n} := (\lambda_{\Theta_n}(x_1), \ldots, \lambda_{\Theta_n}(x_m))$.

Corollary 13 (LDP). *Let Θ_n be a random closed set with integer Hausdorff dimension $n < d$ in $\mathbb{R}^d$ as above, satisfying (24) for some $(x_1, \ldots, x_m) \in (\mathbb{R}^d)^m$, and let $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ be the multivariate "Minkowski content"-based estimator defined in (23).*

Then the sequence $\{\widehat{\boldsymbol{\lambda}}^{\mu,N} : N \geq 1\}$ satisfies the LDP with speed function $v_N = N b_{d-n} r_N^{d-n}$ and good rate function $I_m(\,\cdot\,; \boldsymbol{\lambda}_{\Theta_n})$ defined by $I_m(\boldsymbol{y}; \boldsymbol{\lambda}_{\Theta_n}) := \sum_{j=1}^m \mathcal{I}(y_j; \lambda_j)$, where $\mathcal{I}(y_j; \lambda_j)$ is as in (11).

Proof. The assertion directly follows by Lemma 12 and Theorem 6. □

An estimate of the rate of convergence of $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ to $\boldsymbol{\lambda}_{\Theta_n}$ follows now as a direct consequence of the above corollary and of Remark 2.

Corollary 14 (Convergence rate). *Let Θ_n be as in the assumptions of Corollary 13, and $I_m(B_\delta^c(\boldsymbol{\lambda}_{\Theta_n}); \boldsymbol{\lambda}_{\Theta_n}) := \inf_{\boldsymbol{y} \in B_\delta^c(\boldsymbol{\lambda}_{\Theta_n})} I_m(\boldsymbol{y}; \boldsymbol{\lambda}_{\Theta_n})$. Then, for any $0 < \eta < I_m(B_\delta^c(\boldsymbol{\lambda}_{\Theta_n}); \lambda_{\Theta_n})$ there exists N_0 such that*

$$P\left(\left\|\widehat{\boldsymbol{\lambda}}^{\mu,N} - \boldsymbol{\lambda}_{\Theta_n}\right\| \geq \delta\right) \leq \exp\left(-N b_{d-n} r_N^{d-n} (I_m(B_\delta^c(\boldsymbol{\lambda}_{\Theta_n}); \boldsymbol{\lambda}_{\Theta_n}) - \eta)\right)$$

for all $N \geq N_0$.

Let us now denote by H the quantity, independent on N,

$$H := b_{d-n}(I_m(B_\delta^c(\boldsymbol{\lambda}_{\Theta_n}); \boldsymbol{\lambda}_{\Theta_n}) - \eta),$$

and let us observe that $\sum_{N \geq 1} \exp\left(-N r_N^{d-n} H\right) < \infty$ if $N r_N^{d-n} \sim N^\alpha$ for some $\alpha > 0$. From this and Remark 2 we can state the following

Corollary 15 (Strong consistency). *Let Θ_n be as in the assumptions of Corollary 13, with $r_N \to 0$ such that $N r_N^{d-n}/N^\alpha \to C$ and for some $C, \alpha > 0$ as $N \to \infty$. Then the multivariate estimator $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ of $\boldsymbol{\lambda}_{\Theta_n}$ is strongly consistent, i.e.*

$$\widehat{\boldsymbol{\lambda}}^{\mu,N} \overset{a.s.}{\to} \boldsymbol{\lambda}_{\Theta_n}, \quad \text{as } N \to \infty.$$

Finally, by remembering that in this section $w_N = N b_{d-n} r_N^{d-n}$, $\boldsymbol{p}_N = (\mathbb{P}(x_1 \in \Theta_{n_{\oplus r_N}}), \ldots, \mathbb{P}(x_m \in \Theta_{n_{\oplus r_N}}))$ and Σ is the diagonal covariance matrix defined in (22), the result on moderate deviations and on asymptotic Normality proved in Theorem 8 and in Theorem 10, respectively, may be stated for the multivariate "Minkowski content"-based estimator as follows.

Corollary 16 (Moderate deviations). *Let Θ_n be as in the assumptions of Corollary 13. Then, for any sequence of positive numbers $\{a_N : N \geq 1\}$ such that $\lim_{N\to\infty} a_N = 0$ and $\lim_{N\to\infty} w_N a_N = \infty$, the sequence $\left\{\frac{(w_N \widehat{\boldsymbol{\lambda}}^{\mu,N} - N \boldsymbol{p}_N)}{\sqrt{w_N/a_N}} : N \geq 1\right\}$ satisfies the LDP with speed function $v_N = 1/a_N$ and good rate function $\tilde{I}_m(\cdot;\boldsymbol{\lambda}_{\Theta_n})$ defined by $\tilde{I}_m(\boldsymbol{y}; \boldsymbol{\lambda}_{\Theta_n}) := \sum_{j=1}^m \frac{y_j^2}{2\lambda_j}$.*

Corollary 17 (Asymptotic Normality)**.** *Let Θ_n be as in the assumptions of Corollary 13. Then the sequence $\left\{\frac{1}{\sqrt{w_N}}(w_N\widehat{\boldsymbol{\lambda}}^{\mu,N}-N\boldsymbol{p}_N):N\geq 1\right\}$ converges weakly (as $N\to\infty$) to the centered m-variate Normal distribution $\mathrm{N}_m(\mathbf{0},\Sigma)$.*

4.2. Confidence regions for the mean density

We mentioned that the evaluation and the estimation of the mean density of a random closed set is a problem of great interest in Stochastic Geometry. After having showed that $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ is asymptotically unbiased and consistent for $\boldsymbol{\lambda}_{\Theta_n}$, a natural problem is now to find confidence regions for $\boldsymbol{\lambda}_{\Theta_n}$ at certain fixed level α, given an i.i.d. random sample $\Theta_n^{(1)},\dots,\Theta_n^{(N)}$ for Θ_n. The asymptotic Normality result derived in the previous section will help us in finding out a suitable statistics (see Theorem 21 below). It is worth noticing that it will be crucial to choose a *suitable* bandwidth r_N for $\widehat{\boldsymbol{\lambda}}^{\mu,N}$. In the univariate case ($m=1$) such a bandwidth turns out to be the *optimal bandwidth* $r_N^{o,AMSE}(x)$ defined in [10] as the value which minimizes the *asymptotic mean square error* (AMSE) of $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$:

$$r_N^{\mathrm{o,AMSE}}(x):=\arg\min_{r_N} AMSE(\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)).\tag{29}$$

This is the reason why, in order to propose confidence regions for its mean density, we shall introduce further regularity assumptions on the random set Θ_n and a *common optimal bandwith* (see (30)) associated to the multivariate estimator $\widehat{\boldsymbol{\lambda}}^{\mu,N}$. For the sake of completeness we recall some basic results on $r_N^{o,AMSE}(x)$ proved in [10], and we refer there for further details. For the reader's convenience, we shall use the same notation of [10].

The *mean square error* $MSE(\widehat{\lambda}^{\mu,N}_{\Theta_n}(x))$ of $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$ is defined, as usual, by

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

A Taylor series expansion for the bias and the variance of $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)$ provides an asymptotic approximation of the mean square error, and then, by the definition given in (29), an explicit formula for $r_N^{\mathrm{o,AMSE}}(x)$ is obtained (see Theorem 19 below). To fix the notation, in the sequel $\alpha:=(\alpha_1,\dots,\alpha_d)$ will be a multi-index of $\mathbb{N}_0^d$; we denote

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

Note that $D_y^\alpha f(y,s)=f$ if $|\alpha|=0$. Moreover, we denote by $\mathrm{reach}(A)$ the *reach* of a compact set $A\subset\mathbb{R}^d$, and by $\Phi_i(A\,\cdot)$, $i=0,\dots,n$ its *curvature measures* (we refer to [10, Appendix] and references therein for basic definitions and results on sets with positive reach and on curvature measures). An optimal bandwidth $r_N^{o,AMSE}(x)$ has been obtained for random sets as in (1) satisfying the following assumptions:

(R) For any $s \in \mathbf{K}$, $\operatorname{reach}(Z(s)) > R$, for some $R > 0$, such that there exists a closed set $\Xi(s) \supseteq Z(s)$ such that $\int_{\mathbf{K}} \mathcal{H}^n(\Xi(s))Q(\mathrm{d}s) < \infty$ and

$$\mathcal{H}^n(\Xi(s) \cap B_r(x)) \geq \gamma r^n \qquad \forall x \in Z(s),\ \forall r \in (0,1)$$

for some $\gamma > 0$ independent of s;

($M2$) For any $s \in \mathbf{K}$, $f(\cdot, s)$ is of class C^1 and, for any α with $|\alpha| \in \{0,1\}$, $D_x^\alpha f(\cdot, s)$ is locally bounded such that for any compact $K \subset \mathbb{R}^d$

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

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

$$\int_{\mathbf{K}} |\Phi_i|(\Xi(s))\widetilde{\xi}_K^{(\alpha)}(s)Q(\mathrm{d}s) < \infty \quad \forall i = 0, ..., n.$$

($\overline{A3}$) For any $(s,t) \in \mathbf{K}^2$, the function $g(\cdot, s, \cdot, t)$ is continuous and locally bounded such that for any compact sets $K, \overline{K} \subset \mathbb{R}^d$:

$$\sup_{y \in \overline{K}_{\oplus \operatorname{diam} Z(t)}} \sup_{x \in K_{\oplus \operatorname{diam} Z(s)}} g(x,s,y,t) \leq \xi_{K,\overline{K}}(s,t)$$

for some $\xi_{K,\overline{K}}(s,t)$ with

$$\int_{\mathbf{K}^2} \mathcal{H}^n(\Xi(s))\mathcal{H}^n(\Xi(t))\xi_{K,\overline{K}}(s,t)Q_{[2]}(\mathrm{d}(s,t)) < +\infty.$$

(M4) Ψ has third factorial moment measure

$$\begin{aligned}\nu_{[3]}&(\mathrm{d}(y_1,s_1,y_2,s_2,y_3,s_3)) \\ &= h(y_1,s_1,y_2,s_2,y_3,s_3)\mathrm{d}y_1\mathrm{d}y_2\mathrm{d}y_3 Q_{[3]}(\mathrm{d}(s_1,s_2,s_3))\end{aligned}$$

such that for any $(y_1, s_1, s_2, s_3) \in \mathbb{R}^d \times \mathbf{K}^3$, the function $h(y_1, s_1, \cdot, s_2, \cdot, s_3)$ is continuous and locally bounded such that for any compact sets $K, \overline{K} \subset \mathbb{R}^d$ and $a \in \mathbb{R}^d$:

$$\begin{aligned}\mathbf{1}_{(a-Z(s_1))_{\oplus 1}}(y_1) \sup_{y_2 \in K_{\oplus \operatorname{diam}(Z(s_2))}} \sup_{y_3 \in \overline{K}_{\oplus \operatorname{diam}(Z(s_3))}} & h(y_1,s_1,y_2,s_2,y_3,s_3) \\ &\leq \overline{\xi}_{a,K,\overline{K}}(s_1,y_1,s_2,s_3)\end{aligned}$$

for some $\overline{\xi}_{a,K,\overline{K}}(s_1,y_1,s_2,s_3)$ with

$$\begin{aligned}\int_{\mathbb{R}^d \times \mathbf{K}^3} & \mathcal{H}^n(\Xi(s_2))\mathcal{H}^n(\Xi(s_3)) \\ &\times \overline{\xi}_{a,K,\overline{K}}(s_1,y_1,s_2,s_3)\mathrm{d}y_1 Q_{[3]}(\mathrm{d}(s_1,s_2,s_3)) < +\infty.\end{aligned}$$

Remark 18. *The above assumption (R) plays here the role of assumption (A1) of Theorem 3; namely, it is known that a lower dimensional set with positive reach is locally the graph of a function of class C^1 (e.g., see [6, p. 204]), and so the rectifiability condition in (A1) is fulfilled. Moreover, the condition* $\mathrm{reach}(Z(s)) > R$ *plays a central role in the proof of the theorem below, where a Steiner type formula is applied. Referring to [10] for a more detailed discussion of the above assumptions, we point out here that $(M2)$ implies $(A2)$, while $(\overline{A3})$ together with $(A1)$ imply $(A3)$.*

Theorem 19. *[10] Denote by*

$$\mathcal{A}_1(x) := \frac{b_{d-n+1}}{b_{d-n}} \int_{\boldsymbol{K}} \int_{Z(s)} f(x-y,s)\Phi_{n-1}(Z(s);\mathrm{d}y)Q(\mathrm{d}s)$$

$$\mathcal{A}_2(x) := \frac{d-n}{d-n+1} \sum_{|\alpha|=1} \int_{\boldsymbol{K}} \int_{N(Z(s))} D_x^\alpha f(x-y,s)u^\alpha \mu_n(Z(s);\mathrm{d}(y,u))Q(\mathrm{d}s)$$

$$\mathcal{A}_3(x) := \int_{\boldsymbol{K}^2} \int_{(x-Z(s_1))} \int_{(x-Z(s_2))} g(y_1,s_1,y_2,s_2) \mathcal{H}^{d-1}(\mathrm{d}y_2)\mathcal{H}^{d-1}(\mathrm{d}y_1)Q_{[2]}(\mathrm{d}(s_1,s_2)).$$

- *Let Θ_n be as in (1) with $0 < n < d-1$, satisfying the assumptions (R), $(M2)$ and $(\overline{A3})$; then*

$$r_N^{o,AMSE}(x) = \Big(\frac{(d-n)\lambda_{\Theta_n}(x)}{2Nb_{d-n}(\mathcal{A}_1(x)-\mathcal{A}_2(x))^2}\Big)^{\frac{1}{d-n+2}}.$$

- *Let Θ_n be as in (1) with $0 < n = d-1$, satisfying the assumptions (R), $(M2)$, $(\overline{A3})$ and $(M4)$; then*

$$r_N^{o,AMSE}(x) = \Big(\frac{\lambda_{\Theta_{d-1}}(x)}{4N\big(\mathcal{A}_1(x)-\mathcal{A}_3(x)\big)^2}\Big)^{\frac{1}{3}}.$$

We point out that, in the definition of $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ given in (23), the bandwidth r_N is the same for each component of the vector. Therefore it emerges the need of defining a suitable *common optimal bandwidth* $r_N^{o,c}$. A possible solution might be to take the value which minimizes the usual asymptotic integrated mean square error; actually, since the m points $x_1,\dots,x_m \in \mathbb{R}^d$ are fixed, a more feasible solution is to define as common optimal bandwidth the following quantity:

$$r_N^{o,c} = r_N^{o,c}(x_1,\dots,x_m) := \arg\min_{r_N} \sum_{j=1}^m AMSE(\widehat{\lambda}_{\Theta_n}^{\mu,N}(x_j)). \tag{30}$$

By recalling that (see [10])

$$AMSE(\widehat{\lambda}_{\Theta_n}^{\mu,N}(x_j)) = a_j^2 r_N^2 + \frac{\lambda_j}{w_N},$$

where we set

$$a_j := \begin{cases} \mathcal{A}_1(x_j) - \mathcal{A}_2(x_j) & \text{if } (d-n) > 1 \\ \mathcal{A}_1(x_j) - \mathcal{A}_3(x_j) & \text{if } (d-n) = 1 \end{cases},$$

it follows that

$$\sum_{j=1}^{m} AMSE(\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)) = A^2 r_N^2 + \frac{\Lambda}{N b_{d-n} r_N^{d-n}},$$

with

$$A^2 := \sum_{j=1}^{m} a_j^2 \qquad \text{and} \qquad \Lambda := \sum_{j=1}^{m} \lambda_j.$$

Therefore we immediately obtain that

$$r_N^{o,c} \overset{(30)}{=} \left(\frac{(d-n)\Lambda}{2N b_{d-n} A^2} \right)^{1/(d-n+2)}. \tag{31}$$

Remark 20. *The optimal bandwidth of the multivariate "Minkowski content"-based estimator defined above satisfies the assumptions of Corollary 15, and so* $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ *with* $r_N = r_N^{o,c}$ *is a strongly consistent estimator for* $\boldsymbol{\lambda}_{\Theta_n}$.

Theorem 21. *Let* Θ_n *be a random closed set with integer Hausdorff dimension* $n < d$ *in* $\mathbb{R}^d$ *as in the assumptions of Theorem 19, and* $\widehat{\boldsymbol{\lambda}}^{\mu,N}$ *be the multivariate "Minkowski content"-based estimator defined in (23). Then, if* $r_N \equiv r_N^{o,c}$ *given in (31), the sequence*

$$\left\{ \left(\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_1) - \lambda_1}{\sqrt{\lambda_1 / w_N}}, \dots, \frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_m) - \lambda_m}{\sqrt{\lambda_m / w_N}} \right) : \; N \geq 1 \right\}$$

converges weakly (as $N \to \infty$*) to the* m*-variate Normal distribution* $\mathrm{N}_m(\boldsymbol{\nu}, \mathbb{1})$*, with mean*

$$\boldsymbol{\nu} := \sqrt{\frac{d-n}{2}} \Big(\frac{a_1}{A} \sqrt{\frac{\Lambda}{\lambda_1}}, \dots, \frac{a_m}{A} \sqrt{\frac{\Lambda}{\lambda_m}} \Big)$$

and the identity matrix $\mathbb{1}$ *as covariance matrix, for any* $(x_1, \dots, x_m) \in (\mathbb{R}^d)^m$ *such that (24) is fulfilled.*

Proof. To lighten the notation we denote $(v_j)_{j=1,\dots,m} := (v_1, \dots, v_m)$.

By remembering that $Y^{(j)}_{i,N} := \mathbb{1}_{\Theta^{(i)}_{n_{\oplus r_N}}}(x_j)$ for any $j = 1, \dots, m$, a direct calculation shows that

$$\begin{aligned} &\left(\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \lambda_j}{\sqrt{\lambda_j / w_N}} \right)_{j=1,\dots,m} \\ &\quad = \left(\frac{\frac{1}{w_N} \sum_{i=1}^{N} Y^{(j)}_{i,N} - \lambda_j}{\sqrt{\lambda_j / w_N}} \right)_{j=1,\dots,m} \end{aligned}$$

$$
\begin{aligned}
&= \left(\frac{1}{\sqrt{\lambda_j w_N}} \sum_{i=1}^{N} \left(Y_{i,N}^{(j)} - \frac{w_N \lambda_j}{N} \right) \right)_{j=1,\dots,m} \\
&= \left(\frac{1}{\sqrt{\lambda_j w_N}} \sum_{i=1}^{N} (Y_{i,N}^{(j)} - p_N^{(j)}) + \frac{N p_N^{(j)} - w_N \lambda_j}{\sqrt{\lambda_j w_N}} \right)_{j=1,\dots,m} \\
&= \left(\frac{1}{\sqrt{\lambda_j w_N}} \sum_{i=1}^{N} (Y_{i,N}^{(j)} - p_N^{(j)}) \right)_{j=1,\dots,m} + \left(\frac{N p_N^{(j)} - w_N \lambda_j}{\sqrt{\lambda_j w_N}} \right)_{j=1,\dots,m} . \qquad (32)
\end{aligned}
$$

We study the convergence in distribution of the two vectors in (32) separately. Let us consider the first one and remember that

$$
\widehat{\lambda}_{\Theta_n}^{\mu,N}(x_j) = \frac{1}{w_N} \sum_{i=1}^{N} Y_{i,N}^{(j)};
$$

hence

$$
\left(\frac{1}{\sqrt{\lambda_j w_N}} \sum_{i=1}^{N} (Y_{i,N}^{(j)} - p_N^{(j)}) \right)_{j=1,\dots,m} = \left(\frac{1}{\sqrt{\lambda_j}} \frac{w_N \widehat{\lambda}_{\Theta_n}^{\mu,N}(x_j) - N p_N^{(j)}}{\sqrt{w_N}} \right)_{j=1,\dots,m},
$$

which converges weakly (as $N \to \infty$) to a standard m-variate normal distribution $\mathrm{N}_m(\mathbf{0}, \mathbb{1})$ by Corollary 17.

Now we pass to consider the (non-random) vector in (32):

$$
\left(\frac{N p_N^{(j)} - w_N \lambda_j}{\sqrt{\lambda_j w_N}} \right)_{j=1,\dots,m} = \left(\sqrt{\frac{w_N}{\lambda_j}} \Big(\frac{N p_N^j}{w_N} - \lambda_j \Big) \right)_{j=1,\dots,m} .
$$

By remembering that $w_N = N b_{d-n} (r_N^{o,c})^{d-n}$ and that (see [10, Eq. (5)])

$$
Bias(\widehat{\lambda}_{\Theta_n}^{\mu,N}(x_j)) = \frac{\mathbb{P}(x_j \in \Theta_{n \oplus r_N})}{b_{d-n} r_N^{d-n}} - \lambda_{\Theta_n}(x_j),
$$

we get

$$
\frac{N p_N^{(j)}}{w_N} = \frac{\mathbb{P}(x_j \in \Theta_{n \oplus r_N})}{b_{d-n} (r_N^{o,c})^{d-n}} = Bias(\widehat{\lambda}_{\Theta_n}^{\mu,N}(x_j)) + \lambda_j = a_j r_N^{o,c} + o(r_N^{o,c}) + \lambda_j,
$$

where the last equation follows by Theorem 2 and Theorem 3 in [10].

By replacing now the expression (31) for $r_N^{o,c}$, the previous equality becomes

$$
\frac{N p_N^{(j)}}{w_N} \overset{(31)}{=} \lambda_j + a_j \left(\frac{(d-n)\Lambda}{2 N b_{d-n} A^2} \right)^{1/(d-n+2)} + o\Big(N^{-1/(d-n+2)} \Big);
$$

therefore we have

$$\left(\frac{Np_N^{(j)} - w_N\lambda_j}{\sqrt{\lambda_j w_N}}\right)_{j=1,\dots,m} = \left(\sqrt{\frac{Nb_{d-n}}{N^{(d-n)/(d-n+2)}\lambda_j}}\sqrt{\left(\frac{(d-n)\Lambda}{2b_{d-n}A^2}\right)^{(d-n)/(d-n+2)}}\right.$$
$$\left.\times\left(a_j\left(\frac{(d-n)\Lambda}{2Nb_{d-n}A^2}\right)^{1/(d-n+2)} + o\Big(N^{-1/(d-n+2)}\Big)\right)\right)_{j=1,\dots,m}$$
$$= \left(\frac{a_j}{A}\sqrt{\frac{\Lambda}{\lambda_j}}\sqrt{\frac{d-n}{2}} + o(1)\right)_{j=1,\dots,m}.$$

Hence

$$\lim_{N\to\infty}\left(\frac{Np_N^{(j)} - w_N\lambda_j}{\sqrt{\lambda_j w_N}}\right)_{j=1,\dots,m} = \sqrt{\frac{d-n}{2}}\left(\frac{a_j}{A}\sqrt{\frac{\Lambda}{\lambda_j}}\right)_{j=1,\dots,m}$$

and the assertion follows. □

Corollary 22 (Univariate case)**.** *For every $x \in \mathbb{R}^d$, if $r_N = r_N^{o,AMSE}(x)$, we have*

$$\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x) - \lambda_{\Theta_n}(x)}{\sqrt{\lambda_{\Theta_n}(x)/w_N}} \xrightarrow{d} Z \quad \textit{as} \ \ N \to \infty,$$

where $Z \sim \mathrm{N}(\sqrt{(d-n)/2}, 1)$.

Proof. In this case $r_N^{o,c} = r_N^{o,AMSE}(x)$ and the assertion is a direct consequence of Theorem 21 with $m = 1$. □

Remark 23. *Under the same assumptions of Theorem 21 we can equivalently say that the sequence*

$$\left\{\left(\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_1) - \lambda_1}{\sqrt{\lambda_1/w_N}} - \sqrt{\frac{d-n}{2}}\frac{a_1}{A}\sqrt{\frac{\Lambda}{\lambda_1}}, \dots,\right.\right.$$
$$\left.\left.\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_m) - \lambda_m}{\sqrt{\lambda_m/w_N}} - \sqrt{\frac{d-n}{2}}\frac{a_m}{A}\sqrt{\frac{\Lambda}{\lambda_m}}\right) : \ N \geq 1\right\}$$

converges weakly (as $N \to \infty$) to the m-variate Normal distribution $\mathrm{N}_m(\mathbf{0}, \mathbb{1})$, if $r_N \equiv r_N^{o,c}$ given in (31).

With the aim to find out a confidence region for the vector $\boldsymbol{\lambda}_{\Theta_n}$, we prove now the following result, which will easily follow by Theorem 21 and Slutsky's Theorem.

Proposition 24. *Let the assumptions of Theorem 21 be satisfied. Then, if $r_N \equiv r_N^{o,c}$ given in (31), denoted by*

$$\widehat{\Lambda}_N := \sum_{j=1}^{m} \widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j),$$

the sequence

$$\left\{ \left(\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \lambda_j}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)/w_N}} - \sqrt{\frac{d-n}{2}} \frac{a_j}{A} \sqrt{\frac{\widehat{\Lambda}_N}{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}} \right)_{j=1,\dots,m} : N \geq 1 \right\}$$

converges weakly (as $N \to \infty$) to the centered m-variate Normal distribution $\mathrm{N}_m(\mathbf{0}, \mathbb{1})$ *for any* $(x_1, \cdots, x_m) \in (\mathbb{R}^d)^m$ *such that* (24) *is fulfilled.*

Proof. Let us consider the random vector $\boldsymbol{X}_N = (X_N^{(1)}, \dots, X_N^{(m)})$, whose components are defined as

$$X_N^{(j)} := \frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \lambda_j}{\sqrt{\lambda_j/w_N}} - \sqrt{\frac{d-n}{2}} \frac{a_j}{A} \sqrt{\frac{\Lambda}{\lambda_j}}$$

for $j = 1, \dots, m$. By Remark 23 we know that $\boldsymbol{X}_N$ converges weakly (as $N \to \infty$) to the m-variate Normal distribution $\mathrm{N}_m(\mathbf{0}, \mathbb{1})$.

Let us also define the random vector $\boldsymbol{Y}_N = (Y_N^{(1)}, \dots, Y_N^{(m)})$, with components

$$Y_N^{(j)} := \frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \lambda_j}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)/w_N}} - \sqrt{\frac{d-n}{2}} \frac{a_j}{A} \sqrt{\frac{\widehat{\Lambda}_N}{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}} \qquad \forall j = 1, \dots, m.$$

We remember that $\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) \xrightarrow{\mathbb{P}} \lambda_j$ for every $j = 1, \dots, m$, and also $\widehat{\Lambda}_N \xrightarrow{\mathbb{P}} \Lambda$ by the continuous mapping theorem; thus it follows that

$$\begin{aligned} |X_N^{(j)} - Y_N^{(j)}| \leq & \left| \frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \lambda_j}{\sqrt{\lambda_j/w_N}} - \frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \lambda_j}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)/w_N}} \right| \\ & + \sqrt{\frac{d-n}{2}} \frac{a_j}{A} \left| \sqrt{\frac{\Lambda}{\lambda_j}} - \sqrt{\frac{\widehat{\Lambda}_N}{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}} \right| \xrightarrow{\mathbb{P}} 0, \end{aligned}$$

and so $||\boldsymbol{X}_N - \boldsymbol{Y}_N|| \xrightarrow{\mathbb{P}} 0$. Finally a direct application of Slutsky's Theorem (see [20, Theorem 18.8]) implies the assertion. □

The univariate case directly follows by applying the above proposition with $m = 1$:

Corollary 25 (Univariate case)**.** *For every $x \in \mathbb{R}^d$, if $r_N = r_N^{o,AMSE}(x)$, we have*

$$\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x) - \lambda_{\Theta_n}(x)}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x)/w_N}} \xrightarrow{d} Z \quad \text{as } N \to \infty,$$

where $Z \sim \mathrm{N}(\sqrt{(d-n)/2}, 1)$.

Now, we are ready to state the main theorem of this section concerning the confidence region for $\boldsymbol{\lambda}_{\Theta_n}$:

Theorem 26. *Let the assumptions of Theorem 21 be satisfied with $r_N \equiv r_N^{o,c}$ given in (31), and K_α be such that $\mathbb{P}(W \leq K_\alpha) = 1-\alpha$, for $\alpha \in (0,1)$, where $W \sim \chi^2_m$.*

For any $(x_1,\dots,x_m) \in (\mathbb{R}^d)^m$ such that (24) is fulfilled, an asymptotic confidence region for $\boldsymbol{\lambda}_{\Theta_n}$ is the ellipsoid in $\mathbb{R}^m$ so defined

$$\left\{\boldsymbol{y} \in \mathbb{R}^m \,:\, (\boldsymbol{y}-\boldsymbol{b})^t \boldsymbol{Q}(\boldsymbol{y}-\boldsymbol{b}) \leq 1\right\},$$

where the vector $\boldsymbol{b} = (b_1,\dots,b_m)$ and the $m \times m$ matrix $\boldsymbol{Q}$ are defined, respectively, by

$$b_j := \widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j) - \frac{a_j}{A}\sqrt{\frac{\widehat{\Lambda}_N(d-n)}{2w_N}}$$

and

$$\boldsymbol{Q} := \frac{w_N}{K_\alpha}\mathrm{diag}\left(\frac{1}{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_1)},\cdots,\frac{1}{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_m)}\right).$$

Proof. Let us notice that

$$\begin{aligned}\left(\frac{\sqrt{w_N}}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}}(\lambda_j - b_j)\right)^t_{j=1,\dots,m} &\left(\frac{\sqrt{w_N}}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}}(\lambda_j - b_j)\right)_{j=1,\dots,m} \\ &= \sum_{j=1}^m \left(\frac{\sqrt{w_N}}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}}(\lambda_j - b_j)\right)^2 \\ &= \sum_{j=1}^m \left(\frac{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)-\lambda_j}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)/w_N}} - \sqrt{\frac{d-n}{2}}\frac{a_j}{A}\sqrt{\frac{\widehat{\Lambda}_N}{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}}\right)^2\end{aligned}$$

is asymptotically χ^2_m distributed as a consequence of Proposition 24.

Defined K_α as in the statement of the theorem, and denoted by

$$\mathcal{R} := \left\{\boldsymbol{y} \in \mathbb{R}^m \,:\, (\boldsymbol{y}-\boldsymbol{b})^t \boldsymbol{Q}(\boldsymbol{y}-\boldsymbol{b}) \leq 1\right\}$$

the random ellipsoid in $\mathbb{R}^m$, it is easy to observe that

$$\begin{aligned}1-\alpha \simeq \mathbb{P}\Bigg(&\left(\frac{\sqrt{w_N}}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}}(\lambda_j - b_j)\right)^t_{j=1,\dots,m} \\ &\cdot \left(\frac{\sqrt{w_N}}{\sqrt{\widehat{\lambda}^{\mu,N}_{\Theta_n}(x_j)}}(\lambda_j - b_j)\right)_{j=1,\dots,m} \leq K_\alpha\Bigg) = \mathbb{P}(\boldsymbol{\lambda}_{\Theta_n} \in \mathcal{R}),\end{aligned}$$

and so the assertion follows. □

Corollary 27 (Univariate case)**.** *For every $x \in \mathbb{R}^d$, if $r_N = r_N^{o,AMSE}(x)$, an asymptotic confidence interval for $\lambda_{\Theta_n}(x)$ at level α is given by*

$$\left[\widehat{\lambda}_{\Theta_n}^{\mu,N}(x) - \sqrt{\frac{\widehat{\lambda}_{\Theta_n}^{\mu,N}(x)}{w_N}}\Big(\sqrt{\frac{d-n}{2}} + J_\alpha\Big), \widehat{\lambda}_{\Theta_n}^{\mu,N}(x) - \sqrt{\frac{\widehat{\lambda}_{\Theta_n}^{\mu,N}(x)}{w_N}}\Big(\sqrt{\frac{d-n}{2}} - J_\alpha\Big)\right]$$

where J_α is such that $1-\alpha = \mathbb{P}(-J_\alpha \leq Z \leq J_\alpha)$, with $Z \sim \mathrm{N}(0,1)$.

Proof. The assertion directly follows either by applying Theorem 26 with $m = 1$, or as a consequence of Corollary 25. □

Acknowledgements

The authors wish to thank the anonymous referee for the careful reading, and the valuable comments and suggestions which improved significantly the quality of this paper.

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., Lauder I.J. (2004). *Statistical Concepts and Applications in Clinical Medicine*, Chapman and Hall/CRC, London.

[2] Ambrosio, L., Capasso, V., Villa, E. (2009). On the approximation of mean densities of random closed sets. *Bernoulli* **15**, 1222–1242. MR2597590

[3] Ambrosio, L., Fusco, N., Pallara, D. (2000). *Functions of Bounded Variation and Free Discontinuity Problems*, Clarendon Press, Oxford. MR1857292

[4] Baddeley, A., Barany, I., Schneider, R., Weil, W. (2007). *Stochastic Geometry*. Lecture Notes in Mathematics, Vol. 1982. Springer, Berlin. MR2327289

[5] Baddeley, A., 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. MR1654394

[6] Bellettini, G. (2013). *Lecture notes on mean curvature flow, barriers and singular perturbations*. Scuola Normale Superiore di Pisa (Nuova Serie), Vol. 12, Edizioni della Normale, Pisa. MR3155251

[7] Billingsley, P. (1995). *Probability and Measure*, 3rd edition. John Wiley & Sons. MR1324786

[8] Camerlenghi, F., Capasso, V., Villa, E. (2014). On the estimation of the mean density of random closed sets. *J. Multivariate Anal.* **125**, 65–88. MR3163831

[9] Camerlenghi, F., Capasso, V., 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. MR3241508

[10] Camerlenghi, F., 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 Vis.* **125**, 65–88. MR3397099

[11] Capasso, V. (Ed.) (2003). *Mathematical Modelling for Polymer Processing. Polymerization, Crystallization, Manufacturing.* ECMI Series on Mathematics in Industry, Vol. 2, Springer Verlag, Heidelberg. MR1964653

[12] Capasso V., Dejana E., 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. MR2477136

[13] Capasso V., 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. MR2487997

[14] Capasso V., Micheletti A., Morale D. (2008). Stochastic geometric models and related statistical issues in tumour-induced angiogenesis. *Mathematical Biosciences*, **214**, 20–31. MR2446608

[15] Chiu, S.N., Stoyan, D., Kendall, W.S., Mecke, J. (2013). *Stochastic Geometry and its Applications*, 3rd edition. John Wiley & Sons, Chichcester. MR3236788

[16] Daley, D.J., Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*. Springer Series in Statistics, New York. MR0950166

[17] Dembo, A., Zeitouni, O. (1998). *Large Deviations Techniques and Applications*, 2nd edition. Springer. MR1619036

[18] Devroye L., Györfi L., Lugosi G. (1996). *A Probabilistic Theory of Pattern Recognition*, Springer Series in Stochastic Modelling and Applied Probability, New York. MR1383093

[19] Heinrich, L. (2005). Large deviations of the empirical volume fraction for stationary Poisson grain models. *Ann. Appl. Probab.* **15**(1A), 392–420. MR2115047

[20] Jacod, J., Protter, P. (2003). *Probability essentials*, 2nd edition. Springer-Verlag, Berlin. MR1956867

[21] Karr, A.F. (1986). *Point Processes and Their Statistical Inference*. Marcel Dekker, New York. MR0851982

[22] Matheron, G. (1975). *Random Sets and Integral Geometry*. John Wiley & Sons, New York. MR0385969

[23] Sheather, S.J. (2004). Density estimation. *Statistical Science*, **19**, 588–597. MR2185580

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

[25] Villa, E. (2010). Mean densities and spherical contact distribution function of inhomogeneous Boolean models. *Stoch. An. Appl.* **28**, 480–504. MR2739571

[26] Villa, E. (2014). On the local approximation of mean densities of random closed sets. *Bernoulli* **20**, 1–27. MR3160572

[27] Zähle, M. (1982). Random processes of Hausdorff rectifiable closed sets. *Math. Nachr.* **108**, 49–72. MR0695116