FACULTY OF MATHEMATICS AND PHYSICS
Charles University

DOCTORAL THESIS

Jakub Večeřa

Chaotic random variables in applied probability

Department of Probability and Mathematical Statistics

Supervisor of the doctoral thesis:	Prof. RNDr. Viktor Beneš, DrSc.
Study programme:	Mathematics
Study branch:	4M9

Prague 2019

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Title: Chaotic random variables in applied probability

Author: Jakub Večeřa

Department: Department of Probability and Mathematical Statistics

Supervisor: Prof. RNDr. Viktor Beneš, DrSc. , Department of Probability and Mathematical Statistics

Abstract: This thesis deals with modeling of particle processes. In the first part we examine Gibbs facet process on a bounded window with discrete orientation distribution and we derive central limit theorem (CLT) for U-statistics of facet process with increasing intensity. We calculate all asymptotic joint moments for interaction U-statistics and use the method of moments for deriving the CLT. Moreover we present an alternative proof which makes use of the CLT for U-statistics of a Poisson facet process. In the second part we model planar segment processes given by a density with respect to the Poisson process. Parametric models involve reference distributions of directions and/or lengths of segments. Statistical methods are presented which first estimate scalar parameters by known approaches and then the reference distribution is estimated non-parametrically. We also introduce the Takacs-Fiksel estimate and demonstrate the use of estimators in a simulation study and also using data from actin fibres from stem cells images. In the third part we study a stationary Gibbs particle process with deterministically bounded particles on Euclidean space defined in terms of a finite range potential and an activity parameter. For small activity parameters, we prove the CLT for certain statistics of this Gibbs particle process. To this end an exponential decorrelation property is needed.

Keywords: Gibbs facet process; Gibbs particle process; correlation function; central limit theorem; segment process; estimation

Contents

Introduction

This work contains three main topics:

1. asymptotics of non-Poisson facet process functionals with increasing intensity in Chapter 2,
2. applications and estimation of Poisson and non-Poisson models of segment processes in Chapter 3,
3. asymptotics of particle process functionals with increasing window in Chapter 4.

In Chapter 1 we start with defining general locally finite point process. As a special case we introduce finite point process with density with respect to Poisson point process (process with density). The key definition of a correlation function of the point process is given and we continue with the definition and basic properties of functionals of the process. Moreover we introduce partitions in order to express mixed moments of U-statistics. Then we define facet processes and a specific form of the model for facet process with density. From subsection 1.4 we focus on a particle processes, Gibbs particle process on $\mathbb{R}^d$ and we discuss their basic properties, moreover we introduce a notion admissible Gibbs particle process. Besides we define a special class of functionals of Gibbs particle processes - admissible functions. Admissible Gibbs particle process and admissible function together form an admissible pair, which has important properties we later explore.

In Chapter 2 we first discuss existing results for asymptotic properties of functionals of Poisson facet processes and then introduce our results for processes with density in Theorem 5. Our main result are the central limit theorem for a functional of facet processes and evaluation of asymptotic moments, that hold for processes with density with a special size and orientation distribution of facets. The proof of the central limit theorem is based on the method of moments, which utilizes moment formulas and calculation of limit values of the correlation function. Furthermore we show an alternative proof of the Theorem 5. The results are based on Večeřa and Beneš [2016], Večeřa [2016] and Večeřa and Beneš [2017].

In Chapter 3 we start with introducing two models of the segment process (the facet process in $\mathbb{R}^2$) with density. Moreover we introduce the Takacs-Fiksel methodology for parametric estimates and semiparametric estimates, which combine the Takacs-Fiksel parametric estimates with kernel density estimates. We discuss performance of parametric estimates for both models and of semiparametric estimates for the first model. Finally we show an application on some real biological data. We use processed images of actin stress fibres in stem cell samples and assume that the fibres are driven by the second model, we estimate model parameters and validate the selected model. These results are based on Beneš et al. [2017] and Beneš et al. [2019].

In Chapter 4 we first introduce percolation of a particle process, i.e. existence of infinite connected component. We show that, if the process does not percolate almost surely, then it is admissible (under some other technical conditions). Secondly we define an admissible pair as an admissible Gibbs particle process and a score function derived from a U-statistics. We present asymptotic results for admissible pairs, namely the mean value and variance asymptotics in Theorem 15 and the CLT in Theorem 16. The final part contains some proofs, among them of an auxiliary result (Theorem 19) telling that in an admissible pair the weighted mixed moments approximately factorize. The techniques are based on Blaszczyszyn et al. [2019], they are modified from point processes in $\mathbb{R}^d$ to our setting for particle processes. The whole Chapter 4 is based on Beneš et al. [2019+] and only the results where we were employed in the proofs are presented in detail. The assertions with proofs done by foreign co-authors are presented without proofs.

1. Theoretical background

1.1 Point processes

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. Let $\mathbf{B}$ be a locally compact Polish space, i.e. separable completely metrizable topological space with a reference measure Θ and $\mathbf{Y}$ be a measurable space of integer-valued locally finite measures on $\mathbf{B}$, $\mathcal{Y}$ is the smallest σ-algebra which makes the mappings $\mathbf{x} \mapsto \mathbf{x}(\mathbf{C})$, $\mathbf{x} \in \mathbf{Y}$, measurable for all Borel sets $\mathbf{C} \in \mathcal{B}(\mathbf{B})$.

Definition 1. *Let $x \in \mathbf{B}$ and let δ_x be a measure on $\mathbf{B}$, such that*

$$\delta_x(\mathbf{C}) = \mathbf{1}[x \in \mathbf{C}]\,, \mathbf{C} \in \mathcal{B}(\mathbf{B})\,, \tag{1.1}$$

then δ_x is Dirac measure centered in x.

Definition 2. *For k and $\mathbf{x} \in \mathbf{Y}$ we define k-th factorial measure of $\mathbf{x}$*

$$\mathbf{x}^{(k)}(\mathbf{C}_1 \times \cdots \times \mathbf{C}_k) = \int_{\mathbf{C}_1 \times \cdots \times \mathbf{C}_k} \mathbf{1}\left[y_{i_1} \neq y_{i_2}\,, 1 \leq i_1 < i_2 \leq k\right] \mathbf{x}^k(\mathrm{d}(y_1, \ldots, y_k))\,,$$
$$\mathbf{C}_1, \ldots, \mathbf{C}_k \subseteq \mathbf{B}\,.$$

Definition 3. *Let Ξ be a random variable with values in $(\mathbf{Y}, \mathcal{Y})$, then it is called a point process.*

Definition 4. *The intensity measure Λ of a point process Ξ is the measure*

$$\Lambda(\mathbf{C}) := \mathbb{E}[\Xi(\mathbf{C})]\,, \mathbf{C} \in \mathcal{B}(\mathbf{B})\,.$$

For $k \in \mathbb{N}$ and point process Ξ we define k-th factorial moment measure of Ξ

$$\kappa^{(k)}(\mathbf{C}_1 \times \cdots \times \mathbf{C}_k) = \mathbb{E}\,\Xi^{(k)}(\mathbf{C}_1 \times \cdots \times \mathbf{C}_k)\,, \mathbf{C}_1, \ldots, \mathbf{C}_k \subseteq \mathbf{B}\,.$$

Definition 5. *Let Λ be a non-atomic measure on $\mathbf{B}$, such that $\Lambda(\mathbf{B}) > 0$, and let η be a point process, such that*

(i) $$\mathbb{P}(\eta(\mathbf{C}) = n) = \frac{\Lambda(\mathbf{C})^n}{n!} e^{-\Lambda(\mathbf{C})}\,, \mathbf{C} \in \mathcal{B}(\mathbf{B})\,, \Lambda(\mathbf{C}) < \infty\,, n \in \mathbb{N} \cup \{0\}\,,$$

(ii) $\eta(\mathbf{C}_1)\,, \eta(\mathbf{C}_2)$ are independent, whenever $\mathbf{C}_1, \mathbf{C}_2 \in \mathcal{B}(\mathbf{B})\,, \mathbf{C}_1 \cap \mathbf{C}_2 = \emptyset\,,$

then η is a Poisson point process on $\mathbf{B}$ with intensity measure Λ and probability distribution

$$\mathbb{P}_\eta(A) := \mathbb{P}(\eta^{-1}(A))\,, A \in \mathcal{Y}\,.$$

See Last and Penrose [2017] for fundamental properties of general Poisson processes.

Definition 6. *Let $\Psi : \mathbf{B} \times \mathbf{Y} \mapsto [0, \infty)$ be a measurable function. A point process μ is called a Gibbs point process with (Papangelou) conditional intensity Ψ, if*

$$\mathbb{E}\left[\int f(y, \mu - \delta_y)\mu(\mathrm{d}y)\right] = \mathbb{E}\left[\int f(y, \mu)\Psi(y, \mu)\Theta(\mathrm{d}y)\right] \tag{1.2}$$

holds for all measurable $f : \mathbf{B} \times \mathbf{Y} \to [0, \infty)$, where δ_y is the Dirac measure located at y.

Condition for the existence of Gibbs point process can be found in [Ruelle, 1970] and for uniqueness of the Gibbs point process in [Hofer-Temmel and Houdebert].

Definition 7. *For $k \in \mathbb{N}$, $k \geq 2$, define the measurable function $\Psi_k \colon \mathbf{B}^k \times \mathbf{Y} \to [0, \infty)$ by*

$$\Psi_k(y_1, \dots, y_k, \mathbf{x}) := \Psi(y_1, \mathbf{x})\Psi(y_2, \mathbf{x} + \delta_{y_1}) \cdots \Psi(y_k, \mathbf{x} + \delta_{y_1} + \cdots + \delta_{y_{k-1}}) \,. \quad (1.3)$$

This function is called the (Papangelou) conditional intensity of k-th order.

Definition 8. *For $n \in \mathbb{N}$ and $y_1, \dots, y_n \in \mathbf{B}$ distinct, we define the n-th order correlation function of the Gibbs point process μ*

$$\rho_n(y_1, \dots, y_n; \mu) := \mathbb{E}\Psi_n(y_1, \dots, y_n; \mu) \,.$$

Remark 1. Integer-valued measures can be represented by systems of points corresponding to their support, i.e. for $\mathbf{x} \in \mathbf{Y}$ there exist $\{x_i\}_{i \in I}$, $I \subseteq \mathbb{N}$, such that $\mathbf{x} = \sum_{i \in I} \delta_{x_i}$.

Remark 2. If $\Psi \equiv \alpha$ in (1.2), then μ is Poisson point process with intensity measure $\Lambda = \alpha\Theta$. It follows from Slivnyak-Mecke formula (1.2).

Remark 3. Equation (1.2) can be iterated so as to yield

$$\mathbb{E}\left[\int f(y_1, \dots, y_k, \mu - \delta_{y_1} - \cdots - \delta_{y_k}) \mu^{(k)}(\mathrm{d}(y_1, \dots, y_k))\right] \\ = \mathbb{E}\left[\int f(y_1, \dots, y_k, \mu)\Psi_k(y_1, \dots, y_k, \mu)\Theta^k(\mathrm{d}(y_1, \dots, y_k))\right] \,, \quad (1.4)$$

for each measurable $f \colon (\mathcal{C}^{(d)})^k \times \mathbf{Y} \to [0, \infty)$.

Remark 4. Ψ_n and ρ_n are symmetric functions w.r.t. to the permutations of their first n variables.

Remark 5. Let $\mathbf{B}$ be bounded, then $\mathbf{Y}$ are integer-valued finite measures on $\mathbf{B}$, let η be a Poisson point process on $\mathbf{B}$ with intensity measure Θ and $g : \mathbf{Y} \mapsto [0, \infty)$ be a measurable function satisfying

$$\int_{\mathbf{Y}} g(\mathbf{x}) \mathrm{d}\mathbb{P}_\eta(\mathbf{x}) = 1 \,.$$

Let μ be a finite point process on $\mathbf{B}$ with distribution defined by

$$\mathbb{P}(\mu^{-1}(A)) = \int_A g(\mathbf{x}) \mathrm{d}\mathbb{P}_\eta(\mathbf{x}) \,, A \in \mathcal{Y} \,, \quad (1.5)$$

then μ is a (finite volume) Gibbs point process and g is its density w.r.t. η. If g satisfies

$$g(\mathbf{x}_1) > 0 \Rightarrow g(\mathbf{x}_2) > 0$$

for all $\mathbf{x}_1, \mathbf{x}_2 : \mathbf{x}_2(\mathbf{C}) \leq \mathbf{x}_1(\mathbf{C}) \,, \forall \mathbf{C} \subseteq \mathbf{B}$, then we call it hereditary.

The consequence of (1.5) is a formula

$$\mathbb{E}F(\mu) = \mathbb{E}[F(\eta)g(\eta)] \,.$$

In this case holds

$$\Psi_k(y_1, \dots, y_k, \mathbf{x}) = \begin{cases} \frac{g(\mathbf{x} + \delta_{y_1} + \ldots + \delta_{y_k})}{g(\mathbf{x})} \,, & g(\mathbf{x}) > 0 \,, \\ 0 \,, & g(\mathbf{x}) = 0 \,. \end{cases}$$

1.2 Functionals

In this section we denote F a measurable map $F : \mathbf{Y} \mapsto \mathbb{R}$, η a Poisson process with intensity measure Λ and μ is a Gibbs point process. We use symbol $[n] := \{1, \ldots, n\}$, $n \in \mathbb{N}$ and $[0] = \emptyset$.

Definition 9. *For $n \in \mathbb{N}$, $y_1, \ldots, y_n \in \mathbf{B}$, we define the difference operator of the n-th order*

$$D^n_{y_1,\ldots,y_n} F(\mathbf{x}) := \begin{cases} D^1_{y_1} D^{n-1}_{y_2,\ldots,y_n} F(\mathbf{x})\,, & n \in \{2, \ldots\}\,, \\ F(\mathbf{x} + \delta_{y_1}) - F(\mathbf{x})\,, & n = 1\,, \\ F(\mathbf{x})\,, & n = 0\,. \end{cases}$$

Definition 10. *For $n \in \mathbb{N}$, $y_1, \ldots, y_n \in \mathbf{B}$, point process Ξ, we define functions*

$$T^\Xi_n F(y_1, \ldots, y_n) := \mathbb{E} D^n_{y_1,\ldots,y_n} F(\Xi)\,.$$

Theorem 1. *[Last and Penrose, 2011, Theorem 1.1] For functionals $F_1, F_2 : \mathbf{Y} \mapsto \mathbb{R}$, $F_1, F_2 \in L^2(\mathbb{P}_\eta)$, it holds*

$$\mathbb{E}[F_1(\eta) F_2(\eta)] = \mathbb{E} F_1(\eta) \mathbb{E} F_2(\eta) + \sum_{n=1}^{\infty} \frac{1}{n!} \langle T^\eta_n F_1(\cdot), T^\eta_n F_2(\cdot) \rangle\,,$$

where $\langle \cdot, \cdot \rangle$ is the scalar product in $L_2(\Lambda^n)$.

Definition 11 (U-statistic)**.** *A U-statistic of an order $k \in \mathbb{N}$ of a point process Ξ is a functional defined by*

$$F(\Xi) := \int f(y_1, \ldots, y_k) \Xi^{(k)}(\mathrm{d}(y_1, \ldots, y_k))\,,$$

where $f : \mathbf{B}^k \mapsto \mathbb{R}$ is a function symmetric w.r.t. to the permutations of its variables, $f \in L_1(\Lambda^k)$ and $\Xi^{(k)}$ is k-th factorial measure of Ξ. We say that F is driven by f.

Definition 12. *Let $\tilde{\prod}_{\hat{M}}$ be the set of all partitions $\{J_i\}$ of $[\hat{M}]$, where $\bigcup J_i = [\hat{M}]$ and $J_i \cap J_j = \emptyset$, $i \neq j$. For $\hat{M} = k_1 + \cdots + k_m$ and blocks*

$$J_i := \left\{ \sum_{j=1}^{i-1} k_j + 1, \ldots, \sum_{j=1}^{i} k_j \right\}\,, i \in [m]\,,$$

consider the partition

$$\pi_{k_1,\ldots,k_m} := \{J_i \mid i \in [m]\}$$

and let $\Pi_{k_1,\ldots,k_m} \subseteq \tilde{\prod}_{\hat{M}}$ be the set of all partitions $\sigma \in \tilde{\prod}_{\hat{M}}$ such that $|J \cap J'| \in \{0, 1\}$ for all $J \in \pi_{k_1,\ldots,k_m}$ and all $J' \in \sigma$. Here $|J|$ is the cardinality of a block $J \in \sigma$. We will be referring to blocks of $\pi_{k_1,\ldots,k_m}$ as to rows and let

$$S(\sigma) := |\{J \in \pi_{k_1,\ldots,k_m} \mid \forall J' \in \sigma, |J \cap J'| = 1 \Rightarrow |J'| = 1\}|$$

be the number of singleton rows.

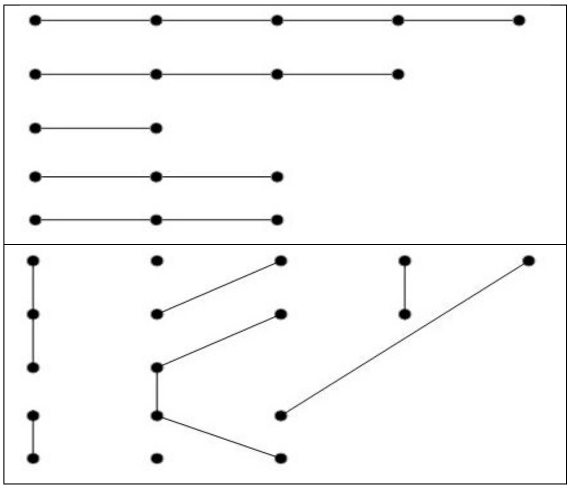

Figure 1.1: Example of partition $\pi_{5,4,2,3,3} \in \tilde{\prod}_{17}$ (upper frame) and $\sigma \in \prod_{5,4,2,3,3}$ (lower frame).
$\pi_{5,4,2,3,3} = \{\{1,2,3,4,5\},\{6,7,8,9\},\{10,11\},\{12,13,14\},\{15,16,17\}\}$
$\sigma = \{\{1,6,10\},\{2\},\{3,7\},\{4,9\},\{5,14\},\{8,11,13,17\},\{12,15\},\{16\}\}$

For a partition $\sigma \in \Pi_{k_1,\dots,k_m}$ and measurable functions $f_i : \mathbf{B}^{k_i} \mapsto \mathbb{R}, i \in [m]$, we define the function $(\otimes_{i=1}^m f_i)_\sigma : \mathbf{B}^{|\sigma|} \mapsto \mathbb{R}$ by replacing all variables of the tensor product $\otimes_{i=1}^m f_i$ that belong to the same block of σ by a new common variable

$$(\otimes_{i=1}^m f_i)_\sigma(y_1,\dots,y_{|\sigma|}) = \prod_{i=1}^m f_i\left(\{y_j \mid |J'_j \cap J_i| \geq 1, J'_j \in \sigma, J_i \in \pi_{k_1,\dots,k_m}\}\right),$$

where $|\sigma|$ is the number of blocks in σ. We define $\Pi_{1,\dots,s}^{m_1,\dots,m_s} := \Pi_{1,\dots,1,\dots,s,\dots,s}$, where i repeats m_i times for $i \in [s]$. Finally we set $\binom{i_1}{i_2} = 0$ for $i_2 > i_1$.

Remark 6. Using the Slivnyak-Mecke Theorem [Schneider and Weil, 2008] for U-statistic $F(\eta)$

$$\mathbb{E}F(\eta) = \int_{\mathbf{B}^k} f(y_1,\dots,y_k)\Lambda^k(\mathrm{d}(y_1,\dots,y_k)).$$

Moreover for a U-statistic $F(\mu) \in L_2(\mathbb{P}_\eta)$ of order k it holds

$$\mathbb{E}F(\mu) = \int_{\mathbf{B}^k} f(y_1,\dots,y_k)\rho_k(y_1,\dots,y_k;\mu)\Lambda^k(\mathrm{d}(y_1,\dots,y_k)). \tag{1.6}$$

In integrals of type similar to (1.6) ρ is defined only for distinct arguments y_j, i.e. only Λ^k-almost everywhere.

In the rest of this section we assume, that $\mathbf{B}$ is bounded and μ is a Gibbs point process with density g with w.r.t. η

Theorem 2. *[Baddeley, 2006] It holds*

$$\mathbb{E}[F(\eta)] = e^{-\Lambda(\mathbf{B})}\sum_{n=0}^{\infty}\frac{1}{n!}\int_{\mathbf{B}}\cdots\int_{\mathbf{B}} F(y_1,\dots,y_n)\Lambda^n(\mathrm{d}(y_1,\dots,y_n)), \tag{1.7}$$

where we write $\Lambda^n(\mathrm{d}(y_1,\dots,y_n))$ instead of $\Lambda(\mathrm{d}y_1)\cdots\Lambda(\mathrm{d}y_n)$.

Theorem 3. *[Beneš and Zikmundová, 2014] Let $m \in \mathbb{N}$, $\prod_{i=1}^m F_i \in L_2(\mathbb{P}_\eta)$, $g \in L_2(\mathbb{P}_\eta)$, where F_i are U-statistics of orders k_i driven by non-negative functions f_i, respectively, $i \in [m]$. Then*

$$\mathbb{E}\left[\prod_{i=1}^m F_i(\mu)\right] = \sum_{\sigma \in \Pi_{k_1,\dots,k_m}} \int_{\mathbf{B}^{|\sigma|}} (\otimes_{i=1}^m f_i)_\sigma(y_1,\dots,y_{|\sigma|}) \times \rho_{|\sigma|}(y_1,\dots,y_{|\sigma|};\mu)\Lambda^{|\sigma|}(\mathrm{d}(y_1,\dots,y_{|\sigma|})) \,. \quad (1.8)$$

Remark 7. For density $g \in L_2(\mathbb{P}_\eta)$, $n \in \mathbb{N}$, it holds

$$T_n^\eta g(y_1,\dots,y_n) = \sum_{J \subseteq [n]} (-1)^{n-|J|} \rho_{|J|}(\{y_i \mid i \in J\};\mu) \,,$$

for Λ^n-almost all $(y_1,\dots,y_n)$, where $|J|$ is the cardinality of J.

1.3 Facet processes

Let $d \in \mathbb{N}$, $\mathcal{B}^d$ be Borel σ-field of $\mathbb{R}^d$, λ^d be the d-dimensional Lebesgue measure, $D \subset \mathbb{R}^d$ be measurable, $\lambda^d(D) \in (0,\infty)$, $\upsilon \in (0,\infty)$, $\mathbb{S}^{d-1}$ be the hemisphere of axial orientations in $\mathbb{R}^d$.

Definition 13. *Let us define a mapping $\iota : D \times (0,\upsilon] \times \mathbb{S}^{d-1} \mapsto \mathcal{B}^d$:*

$$\iota((z,r,\phi)) := \{x \in \mathbb{R}^d \mid \langle \phi, x - z\rangle = 0 \,, \|x - z\|_\infty \in [0,r]\} \,, \quad (1.9)$$

then $\tilde{\mathbf{B}} := \iota((D \times (0,\upsilon] \times \mathbb{S}^{d-1}))$ is a space of facets, $y \in \tilde{\mathbf{B}}$ is a facet, $\|\cdot\|_\infty$ is supremum norm. For $\iota(z,r,\phi) \in \tilde{\mathbf{B}}$, r is a facet size, z is a facet center, ϕ is a normal vector to the hyperplane containing a facet. Further in this section we specially denote

$$\mathbf{B} := D \times (0,\upsilon] \times \mathbb{S}^{d-1} \quad (1.10)$$

a space of facet parameters.

We denote $(\mathbf{Y},\mathcal{Y})$ the measurable space of integer-valued finite measures on $\mathbf{B}$ and the corresponding σ-algebra.

Remark 8. $\mathbf{B}$ is isomorphic with $\tilde{\mathbf{B}}$, where ι is the isomorphism.

When considering facets we assume, that the reference measure Θ has the following form

$$\Theta(\mathrm{d}(z,r,\phi)) := \Theta_\alpha(\mathrm{d}(z,r,\phi)) = \alpha\chi(z)\mathrm{d}zQ(\mathrm{d}r)V(\mathrm{d}\phi) \,, \quad (1.11)$$

where $\alpha \in [1,\infty)$, χ is a intensity measure of facet centres on D, Q is a size distribution, a probability measure on $(0,\upsilon]$, V is an orientation distribution, a probability measure on $\mathbb{S}^{d-1}$.

Definition 14. *Let Ξ be a random variable with values in $(\mathbf{Y},\mathcal{Y})$, then it is called a facet process.*

Definition 15. *The Poisson facet process $\eta_\alpha : (\Omega,\mathcal{A},\mathbb{P}) \mapsto (\mathbf{Y},\mathcal{Y})$ on $\mathbf{B}$ has an intensity measure Λ_α*

$$\Lambda_\alpha(\mathrm{d}(z,r,\phi)) := \Theta_\alpha(\mathrm{d}(z,r,\phi)) = \alpha\chi(z)\mathrm{d}zQ(\mathrm{d}r)V(\mathrm{d}\phi) \,, \quad (1.12)$$

where α, χ, Q, V are taken from (1.11).

Definition 16. *Let $k \in [d]$, then*

$$\mathbb{H}^k(K) = \lim_{t\to 0} \inf_{\substack{\mathcal{S}\subset\mathcal{B}^d \\ \mathcal{S} \text{ countable}}} \left\{ \sum_{L\in\mathcal{S}} (\operatorname{diam} L)^k \mid \bigcup_{L\in\mathcal{S}} L \supseteq K; \operatorname{diam} L < t, L \in \mathcal{S} \right\}, K \in \mathcal{B}^d, \tag{1.13}$$

is Hausdorff measure of order k, where

$$\begin{aligned} \operatorname{diam} L &= \sup_{x_1,x_2\in L} \|x_1 - x_2\|_2\,, & L \neq \emptyset\,, \\ &= 0\,, & L = \emptyset\,, \end{aligned}$$

and $\|\cdot\|_2$ is Euclidean norm.

Definition 17. *Let $k \in [d]$ and Ξ be a facet process, then*

$$G_k(\Xi) := \frac{1}{k!} \int_{\mathbf{B}^k} \mathbb{H}^{d-k} \left(\cap_{i=1}^k \iota(y_i) \right) \Xi^{(k)}(\mathrm{d}(y_1, \ldots, y_k))\,, \tag{1.14}$$

is the facet interaction U-statistic of k-th order.

Definition 18. *Let $\nu := (\nu_1, \ldots, \nu_d) \in \mathbb{R}^d$, $\mathbf{G} := (G_1, \ldots, G_d) : \mathbf{Y}^d \mapsto \mathbb{R}$ and*

$$g(\mathbf{x}) := c \exp(\langle \nu, \mathbf{G}(\mathbf{x}) \rangle)\,, \tag{1.15}$$

then μ_α is defined by density g with respect to η_α. We call μ_α facet process with density.

We assume $\nu_k \in (-\infty, 0]$, $k \in \{2, \ldots, d\}$.

Remark 9. Let $A \subseteq \mathbf{Y}$ be such that

$$A := \{\mathbf{x} = \{(z_1, r_1, \phi_1), \ldots, (z_n, r_n, \phi_n)\} \mid (z_i \neq z_j) \vee (\phi_i \neq \phi_j)\,, 1 \leq i < j \leq n\}\,,$$

then $\mathbb{P}(\eta_\alpha \in A) = 1$ and $\mathbb{P}(\mu_\alpha \in A) = 1$, $\alpha \geq 1$ and for $\mathbf{x} \in A$ it holds

$$G_k(\mathbf{x}) \leq const.\mathbf{x}(\mathbf{B})^k \upsilon^{d-k}\,,$$

which together with the assumption $\nu_k \in (-\infty, 0]$, $k \in \{2, \ldots, d\}$ proves using (1.7) that

$$g \in L_1(\mathbb{P}_{\eta_\alpha}) \cap L_2(\mathbb{P}_{\eta_\alpha})\,.$$

Remark 10. If $\nu_k = 0$, $k \in \{2, \ldots, d\}$, then μ_α is Poisson facet process.

Example. Specially for $d = 3$ the facet process may serve as a model for platelike particles in materials microstructure of metals. Here G_1 yields the total area of all plates, G_2 is the total length of intersection segments of pairs of particles, G_3 is the total number of intersections of triplets of particles. The size of the negative parameter ν_2 or ν_3 gives the measure of neglection of intersections (repulsion) of the corresponding type.

Definition 19. *Let $q \in \{1, \ldots, d\}$ and define $\mu_\alpha^{(q)}$ as facet process with density in form* (1.15) *with $\nu_{q'} = 0$, $\forall q' \neq q$, then we say, that $\mu_\alpha^{(q)}$ is submodel of order q.*

1.4 Particle processes

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and let $\mathbb{R}^d$ be the Euclidean d-dimensional space with Borel σ-algebra $\mathcal{B}^d$ and let $\mathcal{B}_b^d$ denote the system of bounded Borel sets.

Definition 20. *We define $\mathcal{C}^d$ the space of compact subsets (particles) of $\mathbb{R}^d$ and $\mathcal{C}^{(d)} := \mathcal{C}^d \setminus \{\emptyset\}$, which is equipped with a metric Δ_H*

$$\Delta_H(K, L) = \max\left\{\sup_{x_1 \in K} \inf_{x_2 \in L} \|x_1 - x_2\|_2, \sup_{x_2 \in L} \inf_{x_1 \in K} \|x_1 - x_2\|_2\right\}$$

which is called the Hausdorff metric [Last and Penrose, 2017, Schneider and Weil, 2008] and with Borel σ-algebra $\mathcal{B}(\mathcal{C}^{(d)})$ and we remind that $\|\cdot\|_2$ is Euclidean norm.

Moreover for non-empty sets $\mathcal{R}, \mathcal{S} \in \mathcal{B}(\mathcal{C}^{(d)})$ define

$$\Delta(\mathcal{R}, \mathcal{S}) := \inf_{K \in \mathcal{R}, L \in \mathcal{S}} \Delta_H(K, L)\,. \tag{1.16}$$

and let $\zeta(K) \in \mathbb{R}^d$ denote the centre of the circumscribed ball of $K \in \mathcal{C}^{(d)}$, defined as

$$\zeta(K) := \arg\min_{x \in \mathbb{R}^d} \{t \mid \|x - \tilde{x}\|_2 \leq t, \tilde{x} \in K\}\,.$$

Finally define a reference measure on $\mathcal{C}^{(d)}$:

$$\Theta(\cdot) := \iint \mathbf{1}[K + x \in \cdot]\mathbb{Q}(\mathrm{d}K)\mathrm{d}x\,, \tag{1.17}$$

where $\mathbb{Q}$ is a fixed probability measure on $\mathcal{C}^{(d)}$.

Remark 11. To avoid confusion our notation $\Delta(\mathcal{R}, \mathcal{S})$ does not reflect the underlying metric Δ_H.

Remark 12. Note that $\zeta(K + x) = \zeta(K) + x$ for all $(K, x) \in \mathcal{C}^{(d)} \times \mathbb{R}^d$.

Definition 21. *We define $\mathbf{Y}$ the space of all measures $\mathbf{x}$ on $\mathcal{C}^{(d)}$ with values in $\mathbb{N} \cup \{0, \infty\}$ such that $\mathbf{x}(B(K, t)) < \infty$ for each $K \in \mathcal{C}^{(d)}$ and each $t \geq 0$, where*

$$B(K, t) := \{L \in \mathcal{C}^{(d)} \mid \Delta_H(K, L) \leq t\}$$

is the $\mathcal{C}^{(d)}$-ball with radius t centered at K.

We equip this space with the smallest σ-algebra $\mathcal{Y}$ such that the mappings $\mathbf{x} \mapsto \mathbf{x}(\mathcal{R})$ are measurable for each $\mathcal{R} \in \mathcal{B}(\mathcal{C}^{(d)})$.

Remark 13. We use the same symbol $(\mathbf{Y}, \mathcal{Y})$ for the space of Particle process realizations in this chapter as well as the space of Facet Process realizations in the previous chapter.

Definition 22. *A particle process Ξ on $\mathbb{R}^d$ is a random variable with values in $(\mathbf{Y}, \mathcal{Y})$.*

Such a particle process is said to be stationary if $\mathcal{T}_x \Xi \stackrel{\mathcal{D}}{=} \Xi$, for each $x \in \mathbb{R}^d$, where for each measure $\mathbf{x}$ on $\mathcal{C}^{(d)}$ we set

$$\mathcal{T}_x \mathbf{x} := \int \mathbf{1}[K + x \in \cdot]\mathbf{x}(\mathrm{d}K)\,, \quad K + x := \{x' + x \mid x' \in K\}\,.$$

Definition 23. *We say that a particle process* Ξ *is simple if*

$$\Xi(\zeta^{-1}(x)) \in \{0,1\}\,, \forall x \in \mathbb{R}^d \quad a.s..$$

In the following we consider only simple stationary particle processes. We also assume that $\mathbb{P}(\Xi(\mathcal{C}^{(d)}) \neq 0) = 1$.

Definition 24. *The intensity* α *of a stationary particle process* Ξ *is defined by*

$$\alpha := \mathbb{E}\left[\int \mathbf{1}[\zeta(K) \in [0,1]^d]\Xi(\mathrm{d}K)\right]\,,$$

where $[0,1]^d$ *is d-dimensional unit cube. The intensity measure of a particle process* $\mathbb{E}[\Xi]$ *of* Ξ *is the measure* $\Lambda(\mathcal{R}) := \mathbb{E}[\Xi(\mathcal{R})]$, $\mathcal{R} \in \mathcal{B}(\mathcal{C}^{(d)})$.

Remark 14. Theorem 4.1.1 in Schneider and Weil [2008] implies that there exists a probability measure $\mathbb{Q}$ on $\mathcal{C}^{(d)}$ along with a number $\alpha > 0$ such that

$$\Lambda(\cdot) = \alpha \iint \mathbf{1}[K + x \in \cdot]\mathbb{Q}(\mathrm{d}K)\mathrm{d}x\,, \tag{1.18}$$

where $\mathrm{d}x$ refers to integration w.r.t. the Lebesgue measure λ^d on $\mathbb{R}^d$, $\mathbb{Q}$ is called the particle distribution of Ξ and α is the same as in Definition 24.

Definition 25. *We define a Poisson particle process* $\eta_{\alpha\Theta}$ *as a Poisson point process (Def. 5) on* $\mathcal{C}^{(d)}$ *with intensity measure* $\alpha\Theta$*, where* $\alpha > 0$.

Remark 15. Under some integrability assumptions on $\mathbb{Q}$, the Poisson process $\eta_{\alpha\Theta}$ exists as a stationary particle process; see e.g. [Schneider and Weil, 2008][Theorem 4.1.2.].

It is no restriction of generality to assume that

$$\mathbb{Q}(\mathcal{C}_0^d) = 1\,, \tag{1.19}$$

where $\mathcal{C}_0^d := \{K \in \mathcal{C}^{(d)} \mid \zeta(K) = \mathbf{0}\}$ and $\mathbf{0}$ denotes the origin in $\mathbb{R}^d$. However, we make a crucial assumption that there exists $\upsilon > 0$ such that

$$\mathbb{Q}(\{K \in \mathcal{C}^{(d)} \mid K \subseteq B(\mathbf{0}, \upsilon)\}) = 1\,, \tag{1.20}$$

where $B(x, \upsilon)$ is the closed Euclidean ball with radius υ centered at $x \in \mathbb{R}^d$. This puts deterministic bound on the particle size.

Remark 16. We remind, that for $m \in \mathbb{N}$ and $\mathbf{x} \in \mathbf{Y}$, the m-th factorial measure $\mathbf{x}^{(m)}$ of $\mathbf{x}$ is the measure on $(\mathcal{C}^{(d)})^m$ defined by

$$\mathbf{x}^{(m)}(\cdot) := \int \mathbf{1}[(K_1, \dots, K_m) \in \cdot]\mathbf{1}[K_i \neq K_j \text{ for } i \neq j]\mathbf{x}^m(\mathrm{d}(K_1, \dots, K_m))\,.$$

The m-th factorial moment measure $\kappa^{(m)}$ of a simple particle process Ξ is defined by $\kappa^{(m)} := \mathbb{E}[\Xi^{(m)}]$.

Remark 17. For us this is only of relevance if $\mathbf{x}(\{K\}) \leq 1$, for each $K \in \mathcal{C}^{(d)}$ (simple particle process). In this case, $\mathbf{x}^{(m)}$ coincides with the standard definition of the factorial measure; see [Last and Penrose, 2017].

Definition 26. *Given* $\mathcal{R} \in \mathcal{B}(\mathcal{C}^{(d)})$ *define* $\mathbf{Y}_{\mathcal{R}} := \{\mathbf{x} \in \mathbf{Y} \mid \mathbf{x}(\mathcal{R}^c) = 0\}$ *and let* $\mathcal{Y}_{\mathcal{R}}$ *denote the* σ*-algebra on this set of measures. Given* $\mathbf{x} \in \mathbf{Y}$, $L \in \mathcal{B}^d$ *and* $\mathcal{R} \in \mathcal{B}(\mathcal{C}^{(d)})$*, denote by* $\mathbf{x}_L$ *and* $\mathbf{x}_{\mathcal{R}}$ *the restrictions of* $\mathbf{x}$ *to* $\zeta^{-1}(L)$ *and* $\mathcal{R}$*, respectively. Finally, we set* $\mathcal{B}_b(\mathcal{C}^{(d)}) := \{\zeta^{-1}(L) \mid L \in \mathcal{B}_b^d\}$.

1.5 Gibbs particle processes

In this section we present a few fundamental facts on Gibbs particle process in a general setting.

Remark 18. We remind that for $\Psi : \mathcal{C}^{(d)} \times \mathbf{Y} \mapsto [0,\infty)$ measurable function, a particle process μ is called a Gibbs particle process with (Papangelou) conditional intensity Ψ, if

$$\mathbb{E}\left[\int f(K,\mu-\delta_K)\mu(\mathrm{d}K)\right] = \mathbb{E}\left[\int f(K,\mu)\Psi(K,\mu)\Theta(\mathrm{d}K)\right] \tag{1.21}$$

holds for all measurable $f\colon \mathcal{C}^{(d)} \times \mathbf{Y} \to [0,\infty)$, where δ_K is the Dirac measure located at K and Θ is given by (1.17).

The existence of some Gibbs particle processes can be investigated using the method in [Dereudre et al., 2012].

Remark 19. We remind that the k-th correlation function of a Gibbs particle process μ with Papangelou conditional intensity Ψ is the function $\rho_k\colon (\mathcal{C}^{(d)})^k \to [0,\infty)$ defined by

$$\rho_k(K_1,\dots,K_k;\mu) := \mathbb{E}\left[\Psi_k(K_1,\dots,K_k,\mu)\right], \tag{1.22}$$

where Ψ_k is defined in (1.3).

Remark 20. Putting

$$F(K_1,\dots,K_k,\mathbf{x}) = \mathbf{1}[K_1 \in \mathcal{R}_1,\dots,K_k \in \mathcal{R}_k], \quad \mathcal{R}_1,\dots,\mathcal{R}_k \in \mathcal{B}(\mathcal{C}^{(d)}),$$

in (1.4), we obtain that the k-th factorial moment measure of μ is given by

$$\kappa^{(k)}(\cdot) = \int \mathbf{1}[(K_1,\dots,K_k) \in \cdot]\rho_k(K_1,\dots,K_k;\mu)\Theta^k(\mathrm{d}(K_1,\dots,K_k)), \tag{1.23}$$

justifying our terminology.

Definition 27. *We define a measurable function (Hamiltonian)* $H : \mathbf{Y} \times \mathbf{Y} \mapsto (-\infty,\infty]$ *by*

$$H(\mathbf{x},\Upsilon) := \begin{cases} 0, & \text{if } \mathbf{x}(\mathcal{C}^{(d)}) = 0, \\ -\log\Psi_k(K_1,\dots,K_k,\Upsilon), & \text{if } \mathbf{x} = \delta_{K_1} + \dots + \delta_{K_k} \text{ for } K_1,\dots,K_k \in \mathcal{C}^{(d)}, \\ \infty, & \text{if } \mathbf{x}(\mathcal{C}^{(d)}) = \infty. \end{cases}$$

Remark 21. This definition makes sense, since Ψ_k can be chosen to be symmetric in the first k arguments; see Matthes et al. [1979].

Remark 22. For $\mathcal{R} \in \mathcal{B}(\mathcal{C}^{(d)})$, denote by $\eta_{\mathcal{R},\tau\Theta} = (\eta_{\tau\Theta})_{\mathcal{R}}$ the restriction of the Poisson process $\eta_{\tau\Theta}$ to $\mathcal{R}$. Define

$$c_{\mathcal{R}}(\Upsilon) := \mathbb{E}\left[e^{-H(\eta_{\mathcal{R},\tau\Theta},\Upsilon)}\right], \Upsilon \in \mathbf{Y},$$

as the partition function of μ on $\mathcal{R}$. The following DLR-equations (Dobrushin-Lanford-Ruelle equations, see Matthes et al. [1979], Ruelle [1970], Kallenberg [2017]) hold:

$$\mathbb{E}[f(\mu_{\mathcal{R}}) \mid \mu_{\mathcal{R}^c} = \Upsilon] = c_{\mathcal{R}}(\Upsilon)^{-1}\mathbb{E}\left[f(\eta_{\mathcal{R},\tau\Theta})e^{-H(\eta_{\mathcal{R},\tau\Theta},\Upsilon)}\right], \quad \mathcal{R} \in \mathcal{B}_b(\mathcal{C}^{(d)}), \tag{1.24}$$

for $\mathbb{P}(\mu_{\mathcal{R}^c} \in \cdot)$-a.s. $\Upsilon \in \mathbf{Y}_{\mathcal{R}^c}$ and each measurable $f\colon \mathbf{Y} \to [0,\infty)$.

Definition 28. *The k-th Palm measure $\mathbb{P}_{K_1,\dots,K_k}$, $K_1,\dots,K_k \in \mathcal{C}^{(d)}$, $k \in \mathbb{N}$, of a particle process Ξ is a probability measure on $\mathcal{Y}$ satisfying*

$$\mathbb{E}\left[\int_{(\mathcal{C}^{(d)})^k} F(K_1,\dots,K_k,\Xi)\Xi^{(k)}(\mathrm{d}(K_1,\dots,K_k))\right] = \int_{(\mathcal{C}^{(d)})^k}\int_{\mathcal{Y}} F(K_1,\dots,K_k,\mathbf{x})\mathbb{P}_{K_1,\dots,K_k}(\mathrm{d}\mathbf{x})\kappa^{(k)}(\mathrm{d}(K_1,\dots,K_k)), \quad (1.25)$$

for each non-negative measurable function F on $(\mathcal{C}^{(d)})^k \times \mathbf{Y}$.

Palm distributions are well-defined whenever the k-th factorial moment measure $\kappa^{(k)}$ of Ξ is σ-finite. They can be chosen such that $(K_1,\dots,K_k) \mapsto \mathbb{P}_{K_1,\dots,K_k}(\mathcal{R})$ is a measurable function on $(\mathcal{C}^{(d)})^p$, for each $\mathcal{R} \in \mathcal{Y}$. The reduced Palm distribution $\mathbb{P}^!_{K_1,\dots,K_k}$ of Ξ is defined by means of equality

$$\int_{\mathbf{Y}} F(K_1,\dots,K_k,\mathbf{x})\mathbb{P}^!_{K_1,\dots,K_k}(\mathrm{d}\mathbf{x}) = \int_{\mathbf{Y}} F(K_1,\dots,K_k,\mathbf{x}-\delta_{K_1}-\dots-\delta_{K_k})\mathbb{P}_{K_1,\dots,K_k}(\mathrm{d}\mathbf{x}),$$

valid for every non-negative measurable function F on $(\mathcal{C}^d)^k \times \mathbf{Y}$. We abuse our notation by writing, for each measurable $F : \mathbf{Y} \mapsto \mathbb{R}$,

$$\mathbb{E}_{K_1,\dots,K_k}[F(\Xi)] := \int F(\mathbf{x})\mathbb{P}_{K_1,\dots,K_k}(\mathrm{d}\mathbf{x}),$$
$$\mathbb{E}^!_{K_1,\dots,K_k}[F(\Xi)] := \int F(\mathbf{x})\mathbb{P}^!_{K_1,\dots,K_k}(\mathrm{d}\mathbf{x}).$$

1.6 Admissible Gibbs particle processes

Definition 29. *A family $\mathbf{G} := \{G_l\}_{l=2}^{\infty}$ of higher-order interaction potentials consists of measurable, symmetric and translation-invariant functions $G_l\colon (\mathcal{C}^d)^l \mapsto (-\infty,\infty]$. The potentials have finite interaction range υ_G, if $G_l(K_1,\dots,K_l) = 0$, for every $l \geq 2$ and all $K_1,\dots,K_l \in \mathcal{C}^d$ with $\max_{1\leq i<j\leq l}\Delta_H(K_i,K_j) > \upsilon_G$.*

Define the Papangelou intensity $\Psi : \mathcal{C}^d \times \mathbf{Y} \mapsto [0,\infty)$ by the measurable and translation-invariant mapping with $\Psi(K,\mathbf{x}) := 0$, if $\mathbf{x}(K) > 0$, and otherwise

$$\Psi(K,\mathbf{x}) := \tau\exp\left[-\sum_{l=2}^{\infty}\frac{1}{(l-1)!}\int G_l(K,L_1,\dots,L_{l-1})\mathbf{x}^{(l-1)}(\mathrm{d}(L_1,\dots,L_{l-1}))\right], \quad (1.26)$$

where $\tau \in (0,\infty)$ is called an activity parameter and we assume that the part sum in the exponent of (1.26) over the negative parts of the potentials is finite a.e..

We assume that $\Psi \leq \tau$. While the individual potentials might be attractive (i.e. negative), their cumulative effect must be repulsive (i.e., non-negative).

Remark 23. For all $\mathbf{x},\Upsilon \in \mathbf{Y}$ with disjoint support, the Hamiltonian H takes the form

$$H(\mathbf{x},\Upsilon) := \sum_{l=2}^{\infty}\sum_{i=1}^{l}\frac{1}{i!(l-i)!}\iint G_l(K_1,\dots,K_i,L_1,\dots,L_{l-i}) \times \mathbf{x}^{(i)}(\mathrm{d}(K_1,\dots,K_i))\Upsilon^{(l-i)}(\mathrm{d}(K_1,\dots,K_{l-i})).$$

The assumption $\Psi \le \tau$ implies that $H \ge 0$. If assumptions (1.19) and (1.20) hold, then (1.24) shows that the Gibbs process μ has bounded particles, that is

$$\int \mathbf{1}[K \not\subseteq B(\zeta(K), \upsilon)]\mu(\mathrm{d}K) = 0, \mathbb{P}\text{−a.s.}.$$

Definition 30. *Assume that* $\mathbf{G}$ *is a family of higher-order potentials with finite interaction range. Define* Ψ *by* (1.26) *and assume that* $\Psi \le \tau$*. Assume also that* $\mathbb{Q}$ *is a probability measure on* $\mathcal{C}^{(d)}$ *satisfying* (1.19) *and* (1.20)*. Let* $\tau > 0$ *be given. Assume that* μ *is a Gibbs particle process as in Definition 6, where* Θ *is defined by* (1.17)*. Then, we call* μ *an admissible Gibbs process.*

For an admissible Gibbs particle process it follows from (1.26), (1.3) and (1.4)

$$\rho_k(K_1, \dots, K_k; \mu) \le \tau^k, \quad K_1, \dots, K_k \in \mathcal{C}^{(d)}. \tag{1.27}$$

The classic setup of a repulsive intersection-based pair potential arises from a measurable translation invariant function $F : \mathcal{C}^d \mapsto [0, \infty]$ with $F(\emptyset) = 0$ and setting $G_2(K, L) := F(K \cap L)$ and $G_l := 0$, for $l \ge 3$. Assumption (1.20) implies an interaction range of at most 2υ.

1.7 Admissible functions

A function $f\colon (\mathcal{C}^{(d)})^k \to \mathbb{R}$ is called symmetric if

$$f(K_1, \dots, K_k) = f(K_{\pi(1)}, \dots, K_{\pi(k)}),$$

for all $K_1, \dots, K_k \in \mathcal{C}^{(d)}$ and every permutation π of k elements. It is translation invariant, if

$$f(K_1, \dots, K_k) = f(\mathcal{T}_x K_1, \dots, \mathcal{T}_x K_k),$$

for all $K_1, \dots, K_k \in \mathcal{C}^{(d)}$ and $x \in \mathbb{R}^d$. Given a measurable symmetric and translation invariant function f we can define

$$F_f(\mathbf{x}) := \frac{1}{k!}\int f(K_1, \dots, K_k)\, \mathbf{x}^{(k)}(\mathrm{d}(K_1, \dots, K_k)), \qquad \mathbf{x} \in \mathbf{Y}, \tag{1.28}$$

$$\mathbb{T}(K, \mathbf{x}) := \frac{1}{k!}\int f(K, K_2, \dots, K_k)\mathbf{x}^{(k-1)}(\mathrm{d}(K_2, \dots, K_k)), \quad (K, \mathbf{x}) \in \mathcal{C}^{(d)} \times \mathbf{Y}, \tag{1.29}$$

where the case $k = 1$ has to be read as $\mathbb{T}(K) := f(K)$. Then

$$F_f(\mathbf{x}) := \int \mathbb{T}(K, \mathbf{x})\mathbf{x}(\mathrm{d}K), \mathbf{x} \in \mathbf{Y}. \tag{1.30}$$

The authors of [Blaszczyszyn et al., 2019] call $\mathbb{T}$ a score function.

Definition 31. *Let* $k \in \mathbb{N}$ *and let* $f\colon (\mathcal{C}^{(d)})^k \to \mathbb{R}$ *be measurable, symmetric and translation invariant. Then* F_f *is called an admissible function of order* k *if* $f(K_1, \dots, K_k) = 0$*, whenever either*

$$\max_{2 \le j \le k} \Delta_H(K_j, K_1) > r, \tag{1.31}$$

for some given $r > 0$*, or when* $K_j = K_1$*, for some* $j \in \{2, \dots, k\}$*. Moreover, we assume that*

$$\|f\|_\infty := \sup_{K_1, \dots, K_k \in \mathcal{C}^{(d)}} |f(K_1, \dots, K_k)| < \infty. \tag{1.32}$$

2. Asymptotics of facet processes

In this chapter we use definitions introduced in Chapter 1.1 and further specify them. We base this chapter on results developed in Večeřa and Beneš [2016], Večeřa [2016] and Večeřa and Beneš [2017].

In Večeřa and Beneš [2016] we introduced methods to calculate first and second moment of Gibbsian U-statistics of facets in a bounded window of arbitrary Euclidean dimension. This approach was generalized to an arbitrary moment in Večeřa [2016] and used to derive the central limit theorem for such statistics. In Večeřa and Beneš [2017] we have shown a simplified version of the proof of central limit theorem.

Central limit theorems for U-statistics of Poisson processes were derived based on the Malliavin calculus and the Stein method in Reitzner and Schulte [2013]. The aims to extend developments of this type to functionals of a wider class of spatial processes, e.g. Gibbs processes, were initiated by Schreiber and Yukich [2013].

Our calculations are based on the achievements in Beneš and Zikmundová [2014], where functionals of spatial point processes given by a density with respect to the Poisson process were investigated using the Fock space representation from Last and Penrose [2011]. This formula is applied to the product of a functional and the density and using functionals called U-statistics closed formulas for mixed moments of functionals are obtained. In processes with densities the key characteristic is the correlation function [Georgii and Yoo, 2005] of arbitrary order.

We call facets some compact subsets of hyperplanes with a given shape, size and orientation, cf. Def. 13. Natural geometrical characteristics of the union of the facets, based on Hausdorff measure of the intersections of pairs, triplets, etc., of facets form U-statistics. Building a parametric density from the exponential family, the limitations for the space of parameters have to be given, so called submodels are investigated. In applications of the moment formulas we are interested in the limit behaviour as the intensity of the reference Poisson process tends to infinity.

In this chapter we restrict ourselves to the facet model with finitely many orientations corresponding to canonical vectors. When the order of the submodel is not greater than the order of the observed U-statistic then asymptotically the mean value of the U-statistic vanishes. This leads to a degeneracy in the sense that some orientations are missing. On the other hand when the order of the submodel is greater than the order of the observed U-statistic, then the limit of correlation function is finite and nonzero and under selected standardization we achieve a finite nonzero asymptotic variance. Even if these results are obtained in a special situation with facets of a fixed shape, restricted orientations and size related to the window size, it is important that they allow us to understand the ongoing problems for a possible further investigation of the model.

2.1 The Poisson case

Theorem 4. *[Last et al., 2014] For $m \in \mathbb{N}$ and $i \in [m]$ let $f_i \in L_1(\Lambda_1^{k_i})$ be* *symmetric functions. Consider the following U-statistics*

$$F_i(\eta_\alpha) := \int_{\mathbf{B}^{k_i}} f_i(y_1,\dots,y_{k_i})\eta_\alpha^{(k_i)}(y_1,\dots,y_{k_i})\,,$$

normalized to

$$\tilde{F}_i(\eta_\alpha) := \alpha^{-(k_i-\frac{1}{2})}(F_i(\eta_\alpha) - \mathbb{E}F_i(\eta_\alpha))\,.$$

Then

$$(\tilde{F}_1(\eta_\alpha),\dots,\tilde{F}_m(\eta_\alpha)) \xrightarrow[\alpha\to\infty]{\mathcal{D}} \mathbf{Z}\,, \tag{2.1}$$

where $\mathbf{Z} \sim N(0,\Sigma)$, $\Sigma := \{\sigma_{ij}\}_{i,j=1}^m$,

$$\sigma_{ij} := \lim_{\alpha\to\infty} \operatorname{Cov}(\hat{F}_i(\eta_\alpha), \hat{F}_j(\eta_\alpha)) = \int_{\mathbf{B}} T_1^{\eta_\alpha} F_i(y) T_1^{\eta_\alpha} F_j(y)\Lambda(\mathrm{d}y)\,, i,j \in [m]\,. \tag{2.2}$$

The convergence under the distance between m-dimensional random vectors $\mathbf{U}$, $\mathbf{Z}$

$$d_3(\mathbf{U},\mathbf{Z}) := \sup_{h\in\mathcal{H}} |\mathbb{E}h(\mathbf{U}) - \mathbb{E}h(\mathbf{Z})|\,,$$

where $\mathcal{H}$ is the system of functions $h \in C^3(\mathbf{B})$ with

$$\max_{1\le i_1\le i_2\le m} \sup_{y\in\mathbf{B}} \left|\frac{\partial^2 h(y)}{\partial y_{i_1}\partial y_{i_2}}\right| \in [0,1], \qquad \max_{1\le i_1\le i_2\le i_3\le m} \sup_{y\in\mathbf{B}} \left|\frac{\partial^3 h(y)}{\partial y_{i_1}\partial y_{i_2}\partial y_{i_3}}\right| \in [0,1]\,,$$

implies convergence in distribution. Based on the multi-dimensional Malliavin-Stein inequality for the distance d_3 of a random vector from a centered Gaussian random vector $\mathbf{Z}$ with covariance matrix $\Sigma := (\sigma_{ij})_{i,j=1,\dots,m}$, Last et al. [2014] show that under the assumption

$$\int_{\mathbf{B}} |T_1^{\eta_\alpha} F_i(y)|^3 \Lambda(\mathrm{d}y) < \infty\,, i \in [m]\,, \tag{2.3}$$

there exists a constant such that

$$d_3((\tilde{F}_1(\eta_\alpha),\dots,\tilde{F}_m(\eta_\alpha)),\mathbf{Z}) \le \mathit{const}.\alpha^{-\frac{1}{2}}\,, \alpha \in [1,\infty)\,. \tag{2.4}$$

Example. For the Poisson facet processes η_α, $\alpha \in [1,\infty)$ on $\mathbf{B}$ (1.10) with intensity measure Λ_α (1.12) and the U-statistics $G_k(\eta_\alpha)$, $k \in [d]$, in (1.14) we obtain that

$$T_1^{\eta_\alpha} G_k(x) = \frac{\alpha^{k-1}}{(k-1)!} \int_{\mathbf{B}} \cdots \int_{\mathbf{B}} \mathbb{H}^{d-k}\left(\cap_{i=1}^{k-1} \iota(y_i) \cap \iota(x)\right) \Lambda^{k-1}(\mathrm{d}(y_1,\dots,y_{k-1}))\,. \tag{2.5}$$

The finiteness of the intensity measure Λ in (1.12) and the boundedness of the facets guarantee that integrals in (2.2) and (2.3) are finite. Thus for the random vector $(G_1(\eta_\alpha),\dots,G_d(\eta_\alpha))$ both the CLT when $\alpha\to\infty$ and the Berry-Esseen type inequality (2.4) hold.

2.2 The non-Poisson case

Let

$$\mathbf{B} := [0,\upsilon]^d \times \{2\upsilon\} \times \{e_1,\dots,e_d\}\,,$$
$$\tilde{\mathbf{B}} := \iota^{-1}(\mathbf{B})$$

be a constrained space of facet parameters and space of facets with fixed size 2υ and centres in a cube $[0,\upsilon]^d$ (facets are $d-1$ dimensional cubes), where e_i is the i-th vector of standard basis of $\mathbb{R}^d$. That means facets have the same fixed size and shape and any non-parallel facets intersect. In the case of k facets $y_1,\dots,y_k \in \tilde{\mathbf{B}}$ with different orientations we have bounds for the volume of intersection

$$\mathbb{H}^{d-k}\left(\cap_{i=1}^k \iota(y_i)\right) \in [\upsilon^{d-k},(2\upsilon)^{d-k}]\,.$$

In the intensity measure Λ (1.12) we set $Q := \delta_\upsilon$, the orientation distribution V is set to be uniform on $\{e_1,\dots,e_d\}$. Then for $y=(z,r,\phi)$, it holds

$$\Lambda_\alpha(\mathrm{d}(z,r,\phi)) := \alpha\chi(z)\mathrm{d}z\delta_{2\upsilon}(r)\frac{\sum_{i=1}^d \delta_{e_i}(\phi)}{d}\,,$$

Denote $I = \int_{[0,\upsilon]^d}\chi(z)\mathrm{d}z$.

Remark 24. In the special model it can be shown that for any $q \in \{2,\dots,d\}$ if $\nu_i = 0, i \neq q$ and $\nu_q \in [0,\infty)$, then $g \notin L^1(\mathbb{P}_\eta)$, thus the conditions applied to parameters are not only sufficient, but also necessary conditions for the density existence.

Our main result is that the selected functionals of facet processes are (with increasing intensity) asymptotically normally distributed as are the functionals of the Poisson process. This is expressed in the following theorem.

Theorem 5. *Denote*

$$\tilde{G}_k(\mu_\alpha^{(q)}) := \frac{G_k(\mu_\alpha^{(q)}) - \mathbb{E}G_k(\mu_\alpha^{(q)})}{\alpha^{k-\frac{1}{2}}}\,, k \in [d]\,, q \in \{2,\dots,d\}\,, \tag{2.6}$$

then

$$(\tilde{G}_1(\mu_\alpha^{(q)}),\dots,\tilde{G}_d(\mu_\alpha^{(q)})) \xrightarrow[\alpha\to\infty]{\mathcal{D}} \mathbf{Z}\,, \tag{2.7}$$

where $\mathbf{Z} \sim N(0,\Sigma)$*,* $\Sigma := \{\sigma_{k_1k_2}\}_{k_1,k_2=1}^d$*,*

$$\sigma_{k_1k_2} := \frac{(q-1)}{d^{k_1+k_2-1}}\binom{q-2}{k_1-1}\binom{q-2}{k_2-1}I_{k_1k_2}\,,$$
$$I_{k_1k_2} := \int_{([0,\upsilon]^d)^{k_1+k_2-1}} \mathbb{H}^{d-k_2}\left(\cap_{i=2}^{k_2}\iota(z_{i+k_1-1},2\upsilon,e_i)\cap\iota(z_1,2\upsilon,e_1)\right)$$
$$\times\,\mathbb{H}^{d-k_1}\left(\cap_{i=1}^{k_1}\iota(z_i,2\upsilon,e_i)\right)\chi(z_1)\cdots\chi(z_{k_1+k_2-1})(\mathrm{d}(z_1,\dots,z_{k_1+k_2-1}))\,,$$

moreover

$$G_k(\mu_\alpha^{(q)}) \xrightarrow[\alpha\to\infty]{L^1} 0\,, q \in \{2,\dots,d\}\,, k \in \{q,\dots,d\}\,,$$
$$\frac{G_k(\mu_\alpha^{(q)})}{\alpha^k} \xrightarrow[\alpha\to\infty]{L^2} \frac{I_k}{d^k}\binom{q-1}{k}\,, k \in \{2,\dots,d\}\,, k \in [q-1]\,,$$

where

$$I_k := \int_{([0,\upsilon]^d)^k} \mathbb{H}^{d-k}\left(\cap_{i=1}^k\iota(z_i,2\upsilon,e_i))\right)\chi(z_1)\cdots\chi(z_k)\mathrm{d}z_1\cdots\mathrm{d}z_k\,.$$

Remark 25. Asymptotic moments (when $\alpha \to \infty$) of functionals $G_k(\mathbf{x})$, $k \in \{1, \dots, d\}$, in the submodel $\mu_\alpha^{(q)}$, $q \in \{2, \dots, d\}$, were investigated and see the following Table where crosses mean that the expected value is non-zero.

$\alpha \to \infty$	U-statistics G_k				
Submodel	$\mathbb{E}G_d$	$\mathbb{E}G_{d-1}$	$\dots$	$\mathbb{E}G_2$	$\mathbb{E}G_1$
$\mu_\alpha^{(2)}$	0	0	$\dots$	0	$\times$
$\mu_\alpha^{(3)}$	0	0	$\dots$	$\times$	$\times$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\mu_\alpha^{(d-1)}$	0	0	$\dots$	$\times$	$\times$
$\mu_\alpha^{(d)}$	0	$\times$	$\dots$	$\times$	$\times$

Random variables $G_q(\mu_\alpha^{(q)}), \dots, G_d(\mu_\alpha^{(q)})$, $q \in \{2, \dots, d\}$, are asymptotically degenerate, i.e. their expectations tend to zero and asymptotic variances of these variables are $\sigma_{jj} = 0$, $j \in \{q, \dots, d\}$.

Remark 26. For the random vector $(\tilde{G}_1(\eta_\alpha), \dots, \tilde{G}_d(\eta_\alpha))$ we have similar results [Last et al., 2014, Theorem 4.1] with $\sigma_{k_1 k_2} = \frac{d}{d^{k_1+k_2-1}} \binom{d-1}{k_1-1} \binom{d-1}{k_2-1} I_{k_1 k_2}$.

Corollary. Let $\tilde{G}_k$ be as in (2.6) , then it holds

$$\tilde{G}_k(\mu_\alpha^{(q)}) \xrightarrow[\alpha \to \infty]{\mathcal{D}} Z \,, q \in \{2, \dots, d\} \,, k \in [q-1] \,, \tag{2.8}$$

where $Z \sim N(0, \sigma_{kk})$.

Now we state three auxiliary Lemmas, whose proofs are in Section 2.4.

Lemma 6. *It holds*

$$\rho_k(y_1, \dots, y_k; \mu_\alpha^{(q)}) = \frac{\mathbb{E}\exp(\nu_k G_k(\eta_\alpha \cup \{y_1, \dots, y_k\}))}{\mathbb{E}\exp(\nu_k G_k(\eta_\alpha))} \xrightarrow[\alpha \to \infty]{} \frac{\binom{d-l}{d-k+1}}{\binom{d}{d-k+1}} \,, \tag{2.9}$$

as α tends to infinity, where $y_i \in \mathbf{B}$ and l is number of distinct facet orientations among $\{y_1, \dots, y_k\}$ and $k \in \{2, \dots, d\}$.
Moreover there exist $\beta_0 \in [1, \infty)$, $\beta_1, \beta_2 \in (0, \infty)$, which do not depend on $y_1, \dots, y_k$, such that

$$\left| \rho_k(y_1, \dots, y_k; \mu_\alpha^{(q)}) - \frac{\binom{d-l}{d-k+1}}{\binom{d}{d-k+1}} \right| < \beta_1 e^{-\beta_2 \alpha} \,, \forall \alpha \in [\beta_0, \infty) \,.$$

Remark 27. Consider rotation matrix $A = \{a_{ij}\}_{i,j=1}^d$, where $\forall i \exists j_i : a_{ij_i} = 1; a_{ij} = 0, j \neq j_i$ and $\forall j \exists i_j : a_{i_j j} = 1; a_{ij} = 0, i \neq i_j$. Then it can be shown, that for rotation mapping of facets $\tilde{A} : \mathbf{B} \mapsto \mathbf{B}, \tilde{A}((z, r, \phi)) = (z, r, A\phi)$ around their centers given by A the following relation holds

$$\rho_p(y_1, \dots, y_p; \mu_\alpha^{(q)}) = \rho_p(\tilde{A}(y_1), \dots, \tilde{A}(y_p); \mu_\alpha^{(q)}) \,.$$

Definition 32. *For some* $i_1, i_2, i_3 \in \mathbb{N} \cup \{0\}$*,* $i_4, i_5 \in \mathbb{N}$ *and* $\mathbf{p}^{(i_5)} := (p_1, \dots, p_{i_5})$ *let*

$$\gamma(i_1, i_2, i_3, i_4, \mathbf{p}^{(i_5)}) := \sum_{\substack{F \subseteq [i_4] \\ i_1 - i_2 \le |F| \le i_1 \\ |F \cup [i_3]| + i_2 - i_3 \ge i_1}} \prod_{j \in F} p_j \,.$$

If p_j *is the number of facets among* $y_1, \dots, y_n$ *with orientation* e_j*, then specially* $\gamma(q, 0, 0, d, \mathbf{p}^{(d)})$ *is the total number of intersections of all q-tuples among facets* $y_1, \dots, y_n$ *and* $\gamma(q, l, l, d, \mathbf{p}^{(d)})$ *is the total number of intersections of all q-tuples among facets* $y_1, \dots, y_n, (z_1, 2\upsilon, e_1), \dots, (z_l, 2\upsilon, e_l)$*.*

Lemma 7. *For any* $\nu \in (-\infty, 0)$*,* $q \in \{2, \dots, d\}$*,* $l \in \{0, \dots, q-1\}$*, there exist* $\beta_1, \beta_2 \in (0, \infty)$*, such that for* $\alpha \in [1, \infty)$*,*

$$\left| \sum_{p_1=0}^{\infty} \cdots \sum_{p_d=0}^{\infty} \frac{\alpha^{p_1 + \dots + p_d}}{p_1! \cdots p_d!} \exp\left(\nu \gamma(q, l, l, d, \mathbf{p}^{(d)}) - \alpha(q-1)\right) - \frac{(d-l)!}{(q-1-l)!} \right| < \beta_1 e^{-\beta_2 \alpha} \,.$$

For any $\nu \in (-\infty, 0)$*,* $q \in \{2, \dots, d\}$*,* $l = q$*, there exist* $\beta_1, \beta_2 \in (0, \infty)$*, such that for* $\alpha \in [1, \infty)$*,*

$$\left| \sum_{p_1=0}^{\infty} \cdots \sum_{p_d=0}^{\infty} \frac{\alpha^{p_1 + \dots + p_d}}{p_1! \cdots p_d!} \exp\left(\nu \gamma(q, l, l, d, \mathbf{p}^{(d)}) - \alpha(q-1)\right) \right| < \beta_1 e^{-\beta_2 \alpha} \,.$$

Lemma 8. *For any* $q \in \{2, \dots, d\}$*,* $m_1, \dots, m_d \in \mathbb{N} \cup \{0\}$ *and* $\sigma \in \Pi_{1,\dots,d}^{m_1,\dots,m_d}$*, there exist* $\beta_0 \in [1, \infty)$*,* $\beta_1, \beta_2 \in (0, \infty)$*, such that for* $\alpha \in [\beta_0, \infty)$*,*

$$\begin{aligned} \Bigg| \int_{\mathbf{B}^{|\sigma|}} \left(\otimes_{k=1}^{d} \left(\bar{\mathbb{H}}^{d-k} \right)^{\otimes m_k} \right)_{\sigma} (y_1, \dots, y_{|\sigma|}) \rho_{|\sigma|}(y_1, \dots, y_{|\sigma|}; \mu_{\alpha}^{(q)}) \Lambda^{|\sigma|}(\mathrm{d}(y_1, \dots, y_{|\sigma|})) \\ - \int_{\mathbf{B}_{q-1}^{|\sigma|}} \left(\otimes_{k=1}^{d} \left(\bar{\mathbb{H}}^{d-k} \right)^{\otimes m_k} \right)_{\sigma} (y_1, \dots, y_{|\sigma|}) \Lambda^{|\sigma|}(\mathrm{d}(y_1, \dots, y_{|\sigma|})) \Bigg| < \beta_1 e^{-\beta_2 \alpha} \,, \end{aligned}$$

where $\mathbf{B}_{q-1} := [0, \upsilon]^d \times \{2\upsilon\} \times \{e_1, \dots, e_{q-1}\}$ *is the space of facets with* $d - q + 1$ *orientations missing (which can be selected arbitrarily) and* $\bar{\mathbb{H}}^{d-k}(y_1, \dots, y_k) := \mathbb{H}^{d-k}\left(\cap_{i=1}^{k} \iota(y_i)\right)$*.*

Remark 28. The expression

$$\int_{\mathbf{B}^{|\sigma|}} \left(\otimes_{k=1}^{d} \left(\bar{\mathbb{H}}^{d-k} \right)^{\otimes m_k} \right)_{\sigma} (y_1, \dots, y_{|\sigma|}) \rho_{|\sigma|}(y_1, \dots, y_{|\sigma|}; \mu_{\alpha}^{(q)}) \Lambda^{|\sigma|}(\mathrm{d}(y_1, \dots, y_{|\sigma|})) \,,$$

in the statement of Lemma 8 is used in the calculation of the product of moments and this Lemma shows that for large α the correlation function can be removed from the integral, if we consider only $q - 1$ orientations instead of d.
This will be used in the proof of Theorem 5.

Remark 29. For $\nu = (\nu_1, \dots, \nu_d)$, consider process μ with density in more general form

$$g(\mathbf{x}) := c \exp\left(\langle \nu, \mathbf{G}(\mathbf{x}) \rangle\right) \,.$$

Assume that there is $q \in \{2, \dots, d\}$, $\nu_q \in (0, \infty)$ and select such minimal q. Then

$$\begin{aligned}
&\mathbb{E}\exp\left(\nu \cdot \mathbf{G}(\eta_\alpha)\right) \\
&= \sum_{n=0}^{\infty} \frac{\alpha^n e^{-\alpha I}}{n!} \int_{\mathbf{B}^n} \exp\left(\langle \nu, \mathbf{G}(\{y_1, \dots, y_n\})\rangle\right) \Lambda^n(\mathrm{d}(y_1, \dots, y_n)) \\
&\geq e^{-\alpha I} \sum_{p_1=0}^{\infty} \cdots \sum_{p_d=0}^{\infty} \frac{\left(\frac{\alpha I}{d}\right)^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\left(\sum_{k=1}^{d} \nu'_k \sum_{\{i_1,\dots,i_k\}\subseteq[d]} \prod_{l=1}^{k} p_{i_l} \right) \\
&\overset{(1)}{\geq} e^{-\alpha I} \sum_{p_1=0}^{\infty} \cdots \sum_{p_q=0}^{\infty} \frac{\left(\frac{\alpha I}{d}\right)^{p_1+\cdots+p_q}}{p_1! \cdots p_q!} \exp\left(\sum_{k=1}^{q} \nu'_k \sum_{\{i_1,\dots,i_k\}\subseteq[q]} \prod_{j=1}^{k} p_{i_j} \right) \\
&\overset{(2)}{\geq} e^{-\alpha I} \sum_{n=0}^{\infty} \frac{\left(\frac{\alpha I}{d}\right)^{nq}}{(n!)^q} \exp\left(\sum_{k=1}^{q} \nu'_k \binom{q}{k} n^k \right), \qquad (2.10)
\end{aligned}$$

where

$$\nu'_k := \begin{cases} \nu_k \inf\{\mathbb{H}^{d-k}(\cap_{i=1}^k \iota(y_i)) \mid \mathbb{H}^{d-k}(\cap_{i=1}^k \iota(y_i)) > 0\}, \nu_k \in [0, \infty), \\ \nu_k \sup\{\mathbb{H}^{d-k}(\cap_{i=1}^k \iota(y_i)) \mid \mathbb{H}^{d-k}(\cap_{i=1}^k \iota(y_i)) > 0\}, \nu_k \in (-\infty, 0). \end{cases}$$

In (1) we set the last $d - q$ summing variables to zero and in (2) we kept only summands, where all of the summing variables have the same value. It can be proved (e.g. by using ratio test), that the sum in (2.10) is divergent, because $\nu'_k\binom{q}{k} > 0$ is the highest power of n in the exponential. Therefore $g \notin L^1(\mathbb{P}_{\eta_\alpha})$ in this case.

On the other hand if all parameters ν_i are non-positivitem, then $g \in L^2(\mathbb{P}_{\eta_\alpha})$. Thus it holds $\nu_k \in (-\infty, 0]$, $k \in \{2, \dots, d\}$, if and only if $g \in L^2(\mathbb{P}_{\eta_\alpha})$.

Remark 30. Consider process μ_α with density

$$g(\mathbf{x}) = c \exp\left(\langle \nu, \mathbf{G}(\mathbf{x})\rangle\right),$$

where $\nu_k \in (-\infty, 0], k \in \{2, \dots, d\}$. Assume there is $q \in \{2, \dots, d\}$, $\nu_q \in (-\infty, 0)$ and select minimal such q. Then using similar techniques as in proof of Lemma 6 and Lemma 7 we can show that there exist $\beta_1, \beta_2 \in (0, \infty)$, $\beta_0 \in [1, \infty)$ such that

$$\left| \rho_p(y_1, \dots, y_p; \mu^{(q)}) - \lim_{\alpha\to\infty} \rho_p(y_1, \dots, y_p; \mu_\alpha^{(q)}) \right| < \beta_1 e^{-\beta_2 \alpha}, \alpha \in [\beta_0, \infty),$$

which leads to the same asymptotic distribution of statistics $(\tilde{G}_1(\mu_\alpha), \dots, \tilde{G}_d(\mu_\alpha))$ as $(\tilde{G}_1(\mu_\alpha^{(q)}), \dots, \tilde{G}_d(\mu_\alpha^{(q)}))$.

2.3 Proof of the main theorem

Proof of Theorem 5. It holds for moments of U-statistic [Beneš and Zikmundová, 2014, Theorem 3]

$$\mathbb{E}G_k(\mu_\alpha^{(q)}) = \frac{\alpha^k}{k!}\int_{\mathbf{B}^k} \mathbb{H}^{d-k}\left(\cap_{i=1}^k \iota(y_i)\right)\rho_k(y_1,\dots,y_k;\mu_\alpha^{(q)})\Lambda^k(\mathrm{d}(y_1,\dots,y_k)) \tag{2.11}$$

$$\mathbb{E}\prod_{k=1}^{q-1} G_k^{m_k}(\mu_\alpha^{(q)}) = \sum_{\sigma\in\Pi_{1,\dots,q-1}^{m_1,\dots,m_{q-1}}}\prod_{k=1}^{q-1}\frac{1}{k!^{m_k}}\alpha^{|\sigma|}\int_{\mathbf{B}^{|\sigma|}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_\sigma \tag{2.12}$$
$$\times (y_1,\dots,y_{|\sigma|})\rho_{|\sigma|}(y_1,\dots,y_{|\sigma|};\mu_\alpha^{(q)})\Lambda^{|\sigma|}(\mathrm{d}(y_1,\dots,y_{|\sigma|}))\,.$$

We can also get a relation for joint moments of centered random variables

$$\begin{aligned}\mathbb{E}\prod_{k=1}^{q-1}\tilde{G}_k^{m_k}(\mu_\alpha^{(q)}) &= \frac{1}{\alpha^{\mathbf{M}}}\mathbb{E}\prod_{k=1}^{q-1}\left(G_k(\mu_\alpha^{(q)}) - \mathbb{E}G_k(\mu_\alpha^{(q)})\right)^{m_k} \\ &= \frac{1}{\alpha^{\mathbf{M}}}\sum_{i_1=0}^{m_1}\cdots\sum_{i_{q-1}=0}^{m_{q-1}}\binom{m_1}{i_1}\cdots\binom{m_{q-1}}{i_{q-1}}(-1)^{\sum_{k=1}^{q-1}i_k} \\ &\quad\times\left(\prod_{k=1}^{q-1}\left(\mathbb{E}G_k(\mu_\alpha^{(q)})\right)^{i_k}\right)\mathbb{E}\left(\prod_{k=1}^{q-1}G_k^{m_k-i_k}(\mu_\alpha^{(q)})\right)\,,\end{aligned} \tag{2.13}$$

where $\mathbf{M} := \sum_{k=1}^{q-1}(k-\frac{1}{2})m_k$.

Expectations Firstly we calculate normalized asymptotic expectations of the U-statistics. Using Lemma 8 we obtain

$$\begin{aligned}\frac{\mathbb{E}G_k(\mu_\alpha^{(q)})}{\alpha^k} &\xrightarrow[\alpha\to\infty]{} \frac{1}{k!}\int_{\mathbf{B}_{q-1}^k}\mathbb{H}^{d-k}\left(\cap_{i=1}^k\iota(y_i)\right)\Lambda^k(\mathrm{d}(y_1,\dots,y_k)) \\ &= \frac{1}{d^k}I_k\binom{q-1}{k}\,,\end{aligned}$$

where $\binom{q-1}{k}$ is the number of combinations how to select distinct k orientations from $q-1$ and d^k is number of all k-selections of d orientations. The value I_k is an integral of the Hausdorff measure of intersection of k facets with distinct orientations. It does not depend on the currently selection orientations, they only need to be distinct.

Using Lemma 6 for $k \geq q$, we have that $\rho_k(y_1,\dots,y_k;\mu_\alpha^{(q)})$ tends to zero at exponential rate, therefore $\lim_{\alpha\to\infty}\alpha^k\rho_k(y_1,\dots,y_k;\mu_\alpha^{(q)}) = 0$. Moreover the limit and the integral can be interchanged by using Lebesgue's dominated convergence theorem and we obtain

$$G_k(\mu_\alpha^{(q)}) \xrightarrow[\alpha\to\infty]{L^1} 0\,, q\in\{2,\dots,d\}\,, k\geq q\,.$$

So we only need to investigate the U-statistics of the order lower than q.

Limit of correlation function Secondly we calculate all joint moments. To do this we need first to use formula (2.13) and Lemma 8

$$\begin{aligned}
&\left(\prod_{k=1}^{q-1} k!^{m_k}\right)\left(\prod_{k=1}^{q-1}\left(\mathbb{E}G_k(\mu_\alpha^{(q)})\right)^{i_k}\right)\mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\mu_\alpha^{(q)})\right) \qquad (2.14)\\
&\simeq \prod_{k=1}^{q-1}\left(\int_{\mathbf{B}_{q-1}^k}\mathbb{H}^{d-k}\left(\cap_{i=1}^k \iota(y_i)\right)\Lambda^k(\mathrm{d}(y_1,\dots,y_k))\right)^{i_k}\\
&\times \sum_{\sigma\in\Pi_{1,\dots,q-1}^{m_1-i_1,\dots,m_{q-1}-i_{q-1}}} \alpha^{|\sigma|+\sum_{k=1}^{q-1} k i_k}\int_{\mathbf{B}_{q-1}^{|\sigma|}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes(m_k-i_k)}\right)_\sigma\\
&\times (y_1,\dots,y_{|\sigma|})\Lambda^{|\sigma|}(\mathrm{d}(y_1,\dots,y_{|\sigma|}))\,,
\end{aligned}$$

where we use the limit values of correlation function. To describe the relation between the original formula and the formula with correlation function replaced by its limit value we use $\simeq$. We justify the use of the limit of the correlation function, in general we are considering expression in the form

$$\begin{aligned}
&\sum_{i\in F}\alpha^{|\sigma_i|+\tilde{M}^{(i)}} t_i \int_{\mathbf{B}^{|\sigma_i|+\tilde{M}^{(i)}}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes\tilde{m}_k^{(i)}}\right)(x_1,\dots,x_{\tilde{M}^{(i)}})\\
&\qquad\times\left(\otimes_{k=1}^{q-1}\rho_k\left(\cdot;\mu_\alpha^{(q)}\right)^{\otimes\tilde{m}_k^{(i)}}\right)(x_1,\dots,x_{\tilde{M}^{(i)}})\\
&\qquad\times\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k^{(i)}}\right)_{\sigma_i}(y_1,\dots,y_{|\sigma_i|})\rho_{|\sigma_i|}(y_1,\dots,y_{|\sigma_i|};\mu_\alpha^{(q)})\\
&\qquad\times\Lambda^{|\sigma_i|+\tilde{M}^{(i)}}(\mathrm{d}(x_1,\dots,x_{\tilde{M}^{(i)}},y_1,\dots,y_{|\sigma_i|},))\,,
\end{aligned}$$

where $\tilde{M}^{(i)}=\sum_{k=1}^{q-1}k\tilde{m}_k^{(i)}$, $F\subset\mathbb{N}$, $t_i\in\mathbb{R}$, $m_k^i,\tilde{m}_k^i\in\mathbb{N}\cup\{0\}$, $\sigma_i\in\Pi_{1,\dots,q-1}^{m_1,\dots,m_{q-1}}$ and $\rho_k(\cdot;\mu_\alpha^{(q)})$ has a limit $\rho_k^{lim}(\cdot)$ with exponential rate of convergence to this limit, such that $|\rho_k-\rho_k^{\lim}|<\beta_1^{(k)}e^{-\beta_2^{(k)}\alpha}$ for some $\beta_1^{(k)},\beta_2^{(k)}>0$ and

$$\begin{aligned}
&\sum_{i\in F}\alpha^{|\sigma_i|+\tilde{M}^{(i)}} t_i \int_{\mathbf{B}^{|\sigma_i|+\tilde{M}^{(i)}}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes\tilde{m}_k^{(i)}}\right)\left(\otimes_{k=1}^{q-1}\rho_k^{lim}\left(\cdot;\mu_\alpha^{(q)}\right)^{\otimes\tilde{m}_k^{(i)}}\right)\\
&\qquad\times\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k^{(i)}}\right)_{\sigma_i}\left(\rho_{|\sigma_i|}^{lim}\right)\Lambda^{|\sigma_i|+\tilde{M}^{(i)}}=0\,,
\end{aligned}$$

where for better readability we omit integrating variables from the expression. It holds

$$\begin{aligned}
&\Bigg|\sum_{i\in F}\alpha^{|\sigma_i|+\tilde{M}^{(i)}} t_i \int_{\mathbf{B}^{|\sigma_i|+\tilde{M}^{(i)}}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes\tilde{m}_k^{(i)}}\right)\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k^{(i)}}\right)_{\sigma_i}\\
&\times\left(\rho_{|\sigma_i|}^{lim}\left(\otimes_{k=1}^{q-1}\rho_k^{lim}\left(\cdot;\mu_\alpha^{(q)}\right)^{\otimes\tilde{m}_k^{(i)}}\right)-\rho_{|\sigma_i|}\left(\otimes_{k=1}^{q-1}\rho_k\left(\cdot;\mu_\alpha^{(q)}\right)^{\otimes\tilde{m}_k^{(i)}}\right)\right)\Lambda^{|\sigma_i|+\tilde{M}^{(i)}}\Bigg|\\
&\le\beta_1 e^{-\beta_2\alpha}\sum_{i\in F}\alpha^{|\sigma_i|+\tilde{M}^{(i)}} t_i \int_{\mathbf{B}^{|\sigma_i|+\tilde{M}^{(i)}}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes\tilde{m}_k^{(i)}}\right)\\
&\times\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k^{(i)}}\right)_{\sigma_i}\Lambda^{|\sigma_i|+\tilde{M}^{(i)}}\xrightarrow[\alpha\to\infty]{}0\,,
\end{aligned}$$

because it holds

$$\left(\left(\rho^{lim}_{|\sigma_i|}\right)\left(\otimes_{k=1}^{q-1}\rho_k^{lim}\left(\cdot;\mu_\alpha^{(q)}\right)^{\otimes\tilde{m}_k^{(i)}}\right)-\left(\rho_{|\sigma_i|}\right)\left(\otimes_{k=1}^{q-1}\rho_k\left(\cdot;\mu_\alpha^{(q)}\right)^{\otimes\tilde{m}_k^{(i)}}\right)\right)$$
$$\leq\sum_{k=1}^{q-1}\tilde{m}_k^{(i)}\beta_1^{(k)}\exp\left(-\beta_2^{(k)}\alpha\right)+\beta_1^{(|\sigma_i|)}\exp\left(-\beta_2^{(|\sigma_i|)}\alpha\right)\,,$$

we can choose

$$\beta_1=\sum_{i\in F}\left(\sum_{k=1}^{q-1}\tilde{m}_k^{(i)}\beta_1^{(k)}+\beta_1^{(|\sigma_i|)}\right)\,,$$
$$\beta_2=\min\left\{\min_{k\in\{1,\dots,q-1\}}\beta_1^{(k)}\,,\min_{i\in F}\beta_1^{(|\sigma_i|)}\right\}\,.$$

Higher moments Now we return to (2.14). We are interested only in terms with power of α higher or equal than $\mathbf{M}$, because the other terms will tend to zero with increasing α due to normalization, i.e. for $(i_1,\dots,i_{q-1})$ partitions fulfilling condition $|\sigma|\geq\mathbf{M}-\sum_{k=1}^{q-1}i_kk$.

Moreover we do not have to examine odd moments, i.e. those with $\sum_{k=1}^{q-1}m_k$ odd, because there is not any summand with the power of α matching $\mathbf{M}$ in the denominator, thus asymptotically they can only be zero or infinite and if we prove that there is finite even moment of higher order, then the odd moment tend to zero.

So we fix $(m_1,\dots,m_{q-1})$ even mixed moment ($\sum_{i=1}^{q-1}\frac{m_i}{2}\in\mathbb{N}$) and moreover fix $\mathbf{s}:=(s_1,\dots,s_{q-1})$, so that $m_k\geq s_k\geq 0$, $k\in[q-1]$ and $\exists k'\in[q-1]:m_{k'}>s_{k'}$, fix any partition $\sigma_\mathbf{s}\in\Pi^\mathbf{s}_{1,\dots,q-1}$ fulfilling conditions $|\sigma_\mathbf{s}|\geq\mathbf{M}-\sum_{k=1}^{q-1}i_kk$ and $S(\sigma_\mathbf{s})=0$.

Then for any $\mathbf{v}:=(v_1,\dots,v_{q-1})$, such that $m_k\geq v_k\geq s_k$, $k\in[q-1]$ and $\exists k'\in[q-1]$, $v_{k'}>s_{k'}$ select all partitions $\sigma_\mathbf{v}\in\Pi^\mathbf{v}_{1,\dots,q-1}$, such that, $S(\sigma_\mathbf{v})=\sum_{k=1}^{q-1}(v_k-s_k)$, $|\sigma_\mathbf{v}|-|\sigma_\mathbf{s}|=\sum_{k=1}^{q-1}(v_k-s_k)k$ and for whom it holds

$$\alpha^{|\sigma_\mathbf{v}|}\int_{\mathbf{B}_{q-1}^{|\sigma_\mathbf{v}|}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes v_k}\right)_{\sigma_\mathbf{v}}(y_1,\dots,y_{|\sigma_\mathbf{v}|})\Lambda^{|\sigma_\mathbf{v}|}(\mathrm{d}(y_1,\dots,y_{|\sigma_\mathbf{v}|}))$$
$$=\alpha^{|\sigma_\mathbf{s}|}\int_{\mathbf{B}_{q-1}^{|\sigma_\mathbf{s}|}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes s_k}\right)_{\sigma_\mathbf{s}}(y_1,\dots,y_{|\sigma_\mathbf{s}|})\Lambda^{|\sigma_\mathbf{s}|}(\mathrm{d}(y_1,\dots,y_{|\sigma_\mathbf{s}|}))$$
$$\times\prod_{k=1}^{q-1}\left(\alpha^k\int_{\mathbf{B}_{q-1}^k}\mathbb{H}^{d-k}\left(\cap_{i=1}^k\iota(y_i)\right)\Lambda^k(\mathrm{d}(y_1,\dots,y_k))\right)^{v_k-s_k}\,,$$

because we can separate the singleton rows corresponding to the functions $\bar{\mathbb{H}}^{d-k}$ in tensor product, which can be integrated separately, because they do not have any common variables with the other functions in the tensor product and the integral is equal to the expectation of U-statistic. We can see that all summands corresponding to any of the partitions $\sigma_\mathbf{v}$ in the evaluation of (2.13) contain common term

$$\Theta:=\alpha^{|\sigma_\mathbf{s}|}\int_{\mathbf{B}_{q-1}^{|\sigma_\mathbf{s}|}}\left(\otimes_{k=1}^{q-1}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes s_k}\right)_{\sigma_\mathbf{s}}(y_1,\dots,y_{|\sigma_\mathbf{s}|})\Lambda^{|\sigma_\mathbf{s}|}(\mathrm{d}(y_1,\dots,y_{|\sigma_\mathbf{s}|}))$$
$$\times\prod_{k=1}^{q-1}\left(\alpha^k\int_{\mathbf{B}_{q-1}^k}\mathbb{H}^{d-k}\left(\cap_{i=1}^k\iota(y_i)\right)\Lambda^k(\mathrm{d}(y_1,\dots,y_k))\right)^{m_k-s_k}\,,$$

then for fixed $\mathbf{s}$ we sum over all possible $\mathbf{v}$ and all possible partitions $\sigma_{\mathbf{v}}$ in (2.13)

$$\Theta \sum_{v_1=s_1}^{m_1} \cdots \sum_{v_{q-1}=s_{q-1}}^{m_{q-1}} \binom{m_1}{v_1} \cdots \binom{m_{q-1}}{v_{q-1}} \binom{v_1}{s_1} \cdots \binom{v_{q-1}}{s_{q-1}} (-1)^{\sum_{k=1}^{q-1} v_k}$$
$$= \Theta (-1)^{\sum_{k=1}^{q-1} s_k} \binom{m_1}{s_1} \cdots \binom{m_{q-1}}{s_{q-1}}$$
$$\times \sum_{v_1=0}^{m_1-s_1} \cdots \sum_{v_{q-1}=0}^{m_{q-1}-s_{q-1}} \binom{m_1-s_1}{v_1} \cdots \binom{m_{q-1}-s_{q-1}}{v_{q-1}} (-1)^{\sum_{k=1}^{q-1} v_k} = 0\,,$$

where we use the Binomial theorem for summing with necessary condition

$$\sum_{k=1}^{q-1} s_k < \sum_{k=1}^{q-1} m_k$$

and $\binom{m_k}{v_k}$ are original coefficients from formula (2.13) and $\binom{v_k}{s_k}$ is the number of options how to select additional singleton rows.
Therefore all partitions σ, such that $S(\sigma) > 0$ and the ones contained within $\Pi^{\mathbf{s}}_{1,\dots,q-1}, \mathbf{s} < \mathbf{m} := (m_1, \dots, m_{q-1})$ cancel each other out in the summing in the moment formula.

Now we sum over the remaining partitions in (2.13). They all fulfill $\sigma \in \Pi^{\mathbf{m}}_{1,\dots,q-1}$, $S(\sigma) = 0$ and $|\sigma| \geq \sum_{k=1}^{q-1}(k - \frac{1}{2})m_k$. These partitions have each row connected exactly to one another row by one block of two elements in σ ($|\sigma| = M$):

$$\left(\prod_{k=1}^{q-1} k!^{m_k}\right) \mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\mu_\alpha^{(q)})\right) \prod_{k=1}^{q-1} \left(\mathbb{E} G_k(\mu_\alpha^{(q)})\right)^{i_k}$$
$$\simeq \sum_{j_1^{(2)},\dots,j_{m_2}^{(2)}=1}^{2} \cdots \sum_{j_1^{(q-1)},\dots,j_{m_{q-1}}^{(q-1)}=1}^{q-1} \sum_{\substack{\sigma \in \tilde{\prod}_M \\ |J|=2, J\in\sigma}} \prod_{J=\{\varsigma_1,\varsigma_2\}\in\sigma} \alpha^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}$$
$$\times \int_{\mathbf{B}_{q-1}^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}} \mathbb{H}^{d-\mathbf{t}(\varsigma_1)}\left(\cap_{i=1}^{\mathbf{t}(\varsigma_1)} \iota(y_i)\right) \mathbb{H}^{d-\mathbf{t}(\varsigma_2)}\left(\cap_{i=1}^{\mathbf{t}(\varsigma_2)-1} \iota(y_{\mathbf{t}(\varsigma_1)+i}) \cap \iota(y_1)\right)$$
$$\times \Lambda^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}(\mathrm{d}(y_1, \dots, y_{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}))\,,$$
$$\mathbf{t}(\varsigma_i) := \max_{j\in[q-1]} \left\{\sum_{k=1}^{j-1} m_k < i\right\}\,,$$

where we sum first over all possible selections of common elements among the partitions and then over all possible pairings of partition rows, we also divide integral into several parts, where each part consists only of elements which are in the same block of a partition. Function $\mathbf{t}$ connects row of partition to its length. Using definition of Λ, it holds

$$\int_{\mathbf{B}_{q-1}^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}} \mathbb{H}^{d-\mathbf{t}(\varsigma_1)}\left(\cap_{i=1}^{\mathbf{t}(\varsigma_1)} \iota(y_i)\right) \mathbb{H}^{d-\mathbf{t}(\varsigma_2)}\left(\cap_{i=1}^{\mathbf{t}(\varsigma_2)-1} \iota(y_{\mathbf{t}(\varsigma_1)+i}) \cap \iota(y_1)\right)$$
$$\times \Lambda^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}(\mathrm{d}(y_1, \dots, y_{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}))$$
$$= \frac{(q-1)(\mathbf{t}(\varsigma_1)-1)!(\mathbf{t}(\varsigma_2)-1)! I_{\mathbf{t}(\varsigma_1)\mathbf{t}(\varsigma_2)} \binom{q-2}{\mathbf{t}(\varsigma_1)-1}\binom{q-2}{\mathbf{t}(\varsigma_2)-1}}{d^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}}\,, \tag{2.15}$$

$$\sum_{j_1^{(2)},\dots,j_{m_2}^{(2)}=1}^{2} \cdots \sum_{j_1^{(q-1)},\dots,j_{m_{q-1}}^{(q-1)}=1}^{q-1} 1 = \prod_{k=1}^{q-1} k^{m_k}\,, \tag{2.16}$$

where $q-1$ is the number of choices of the one common facet orientation, $\binom{q-2}{\mathbf{t}(\varsigma_1)-1}$, $\binom{q-2}{\mathbf{t}(\varsigma_2)-1}$ is number of combinations how to select the remaining distinct orientations of the rest of the facets orientations in the first and the second function in integrand and $(\mathbf{t}(\varsigma_1)-1)!$, $(\mathbf{t}(\varsigma_2)-1)!$ are the numbers of their allocations into $\mathbf{t}(\varsigma_1)-1$ and $\mathbf{t}(\varsigma_2)-1$ positions, $d^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}$ is $\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1$-selection of d orientations (even non-distinct ones) and $I_{\mathbf{t}(\varsigma_1)\mathbf{t}(\varsigma_2)}$ is integral over facets with fixed orientations over the space of the facet centres. Then using (2.15) and (2.16)

$$\mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\mu_\alpha^{(q)})\right)\prod_{k=1}^{q-1}\left(\mathbb{E}G_k(\mu_\alpha^{(q)})\right)^{i_k}$$

$$\simeq \left(\frac{\alpha}{d}\right)^{\mathbf{M}} \sum_{j_1^{(2)},\ldots,j_{m_2}^{(2)}=1}^{2} \cdots \sum_{j_1^{(q-1)},\ldots,j_{m_{q-1}}^{(q-1)}=1}^{q-1} \sum_{\substack{\sigma\in\tilde{\prod}_M\\ |J|=2, J\in\sigma}} \prod_{J=\{\varsigma_1,\varsigma_2\}\in\sigma}$$

$$\left(\prod_{k=1}^{q-1}\frac{1}{k!^{m_k}}\right)(q-1)(\mathbf{t}(\varsigma_1)-1)!(\mathbf{t}(\varsigma_2)-1)! I_{\mathbf{t}(\varsigma_1)\mathbf{t}(\varsigma_2)}\binom{q-2}{\mathbf{t}(\varsigma_1)-1}\binom{q-2}{\mathbf{t}(\varsigma_2)-1}$$

$$= \left(\frac{\alpha}{d}\right)^{\mathbf{M}} \sum_{\substack{\sigma\in\tilde{\prod}_M\\ |J|=2, J\in\sigma}} \prod_{J=\{\varsigma_1,\varsigma_2\}\in\sigma}(q-1) I_{\mathbf{t}(\varsigma_1)\mathbf{t}(\varsigma_2)}\binom{q-2}{\mathbf{t}(\varsigma_1)-1}\binom{q-2}{\mathbf{t}(\varsigma_2)-1},$$

because the factorial terms $(\mathbf{t}(\varsigma_1)-1)!(\mathbf{t}(\varsigma_2)-1)!$ and the sums cancel out with the term $\prod \frac{1}{k!^{m_k}}$. Therefore it holds

$$\mathbb{E}\prod_{k=1}^{q-1}\tilde{G}_k^{m_k}(\mu_\alpha^{(q)}) \simeq \sum_{\substack{\sigma\in\tilde{\prod}_M\\ |J|=2, J\in\sigma}} \prod_{J=\{\varsigma_1,\varsigma_2\}\in\sigma} \frac{(q-1)I_{\mathbf{t}(\varsigma_1)\mathbf{t}(\varsigma_2)}}{d^{\mathbf{t}(\varsigma_1)+\mathbf{t}(\varsigma_2)-1}}\binom{q-2}{\mathbf{t}(\varsigma_1)-1}\binom{q-2}{\mathbf{t}(\varsigma_2)-1}$$

and as a special case we get

$$\mathbb{E}\tilde{G}_{k_1}(\mu_\alpha^{(q)})\tilde{G}_{k_2}(\mu_\alpha^{(q)}) \simeq \frac{(q-1)I_{k_1k_2}}{d^{k_1+k_2-1}}\binom{q-2}{k_1-1}\binom{q-2}{k_2-1}.$$

Now consider vector of multivariate normal distribution $\mathbf{Z}=(Z_1,\ldots,Z_d)\sim N(0,\Sigma)$, then for any joint moment we have

$$\mathbb{E}\prod_{k=1}^{d} Z_k^{m_k} = \sum_{\substack{\sigma\in\tilde{\prod}_M\\ |J|=2, J\in\sigma}} \prod_{J=\{\varsigma_1,\varsigma_2\}\in\sigma} \mathbb{E}Z_{\mathbf{t}(\varsigma_1)}Z_{\mathbf{t}(\varsigma_2)}.$$

We can see that asymptotically the distribution of statistics has the property of normal distribution, i.e. joint moments of centered variables are equal to sum over all pairs of unordered random variables and this implies the central limit theorem, because normal distribution is defined by its moments [Billingsley, 1995, Theorem 30.2.].

There is only one remaining statement to prove

$$\frac{G_k(\mu_\alpha^{(q)})}{\alpha^k} \xrightarrow[\alpha\to\infty]{L^2} \frac{I_k}{d^k}\binom{q-1}{k}, q\in\{2,\ldots,d\}, k<q,$$

the first moment of the random variable on the left-hand side tends to right-hand side and the variance tends to zero as can be seen from the proof of central limit theorem. □

Alternative proof of Theorem 5. Starting from formula (2.14) for fixed $i_1, \ldots, i_{q-1}$ we have alternative way to prove the central limit theorem. The Poisson process $\eta_\alpha^{(q-1)}$ on $\mathbf{B}_{q-1}$ with intensity measure $\Lambda_\alpha^{(q-1)}$, which is restriction of Λ to $\mathbf{B}_{q-1}$ has higher moments in the following form

$$
\begin{aligned}
&\left(\prod_{k=1}^{q-1} k!^{m_k}\right) \mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\eta_\alpha^{(q-1)})\right) \prod_{k=1}^{q-1} \left(\mathbb{E}G_k(\eta_\alpha^{(q-1)})\right)^{i_k} \\
&= \prod_{k=1}^{q-1} \left(\int_{\mathbf{B}_{q-1}^k} \mathbb{H}^{d-k}\left(\cap_{i=1}^k \iota(y_i)\right) \Lambda^k(\mathrm{d}(y_1, \ldots, y_k))\right)^{i_k} \\
&\times \sum_{\sigma \in \Pi_{1,\ldots,q-1}^{m_1-i_1,\ldots,m_{q-1}-i_{q-1}}} \alpha^{|\sigma|+\sum_{k=1}^{q-1} k i_k} \int_{\mathbf{B}_{q-1}^{|\sigma|}} \left(\otimes_{k=1}^{q-1}\left(\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes(m_k-i_k)}\right)\right)_\sigma \\
&\times (y_1, \ldots, y_{|\sigma|}) \Lambda^{|\sigma|}(\mathrm{d}(y_1, \ldots, y_{|\sigma|}))
\end{aligned}
$$

where we used $\rho \equiv 1$ for Poisson process. Which combined with (2.14) gives us

$$
\begin{aligned}
&\Bigg| \left(\prod_{k=1}^{q-1} k!^{m_k}\right) \mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\eta_\alpha^{(q-1)})\right) \prod_{k=1}^{q-1} \left(\mathbb{E}G_k(\eta_\alpha^{(q-1)})\right)^{i_k} \\
&- \left(\prod_{k=1}^{q-1} k!^{M_k}\right) \mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\mu_\alpha^{(q)})\right) \prod_{k=1}^{q-1} \left(\mathbb{E}G_k(\mu_\alpha^{(q)})\right)^{i_k} \Bigg| < \beta_1 e^{-\beta_2 \alpha},
\end{aligned}
$$

where $\beta_1, \beta_2 > 0$ and also

$$
\beta_1 e^{-\beta_2 \alpha} \left\{ \left(\prod_{k=1}^{q-1} k!^{m_k}\right) \mathbb{E}\left(\prod_{k=1}^{q-1} G_k^{m_k-i_k}(\eta_\alpha^{(q-1)})\right) \prod_{k=1}^{q-1} \left(\mathbb{E}G_k(\eta_\alpha^{(q-1)})\right)^{i_k} \right\} \xrightarrow[\alpha \to \infty]{} 0\,, \tag{2.17}
$$

because moments of Poisson process are at most polynomial in α.
The moments of the Poisson U-statistics behave asymptotically as the moments of a Gaussian distribution, see [Last et al., 2014, Corollary 4.3 and Proposition 5.1]. Therefore using the moment method we have the central limit theorem (2.7) for $(\tilde{G}_1(\mu_\alpha^{(q)}), \ldots, \tilde{G}_{q-1}(\mu_\alpha^{(q)})$. □

2.4 Proofs of the Lemmas

Proof of Lemma 6. First consider submodel $\mu_\alpha^{(q)}$ and facets $y_1, \ldots, y_k$ with $l = k \leq q$ distinct orientations. Moreover without loss of generality we consider orientations $\{e_1, \ldots, e_l\}$, because if we apply rotations from Remark 27, then value of correlation function does not change. It holds

$$
\begin{aligned}
&\rho_k(y_1, \ldots, y_k; \mu_\alpha^{(q)}) \\
&\quad = \frac{\sum_{n=0}^\infty \frac{\alpha^n}{n!} \int_{\mathbf{B}^n} \exp\left(\nu_q G_q(\{x_1, \ldots, x_n, y_1, \ldots, y_k\})\right) \Lambda^n(\mathrm{d}(x_1, \ldots, x_n))}{\sum_{n=0}^\infty \frac{\alpha^n}{n!} \int_{\mathbf{B}^n} \exp\left(\nu_q G_q(\{x_1, \ldots, x_n\})\right) \Lambda^n(\mathrm{d}(x_1, \ldots, x_n))}\,.
\end{aligned}
$$

We can obtain bounds for this expression by using the bounds for the volumes of intersection of facets $\mathbb{H}^{d-q}\left(\cap_{i=1}^{q}\iota(y_i)\right) \in [\upsilon^{d-q}, (2\upsilon)^{d-q}]$ as follows

$$\frac{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q (2\upsilon)^{d-q}\gamma(q,k,k,d,\mathbf{p}^{(d)})\right)}{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q (\upsilon)^{d-q}\gamma(q,0,0,d,\mathbf{p}^{(d)})\right)}$$
$$\leq \rho_k(y_1,\ldots,y_k;\mu_\alpha^{(q)})$$
$$\leq \frac{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q \upsilon^{d-q}\gamma(q,k,k,d,\mathbf{p}^{(d)})\right)}{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q (2\upsilon)^{d-q}\gamma(q,0,0,d,\mathbf{p}^{(d)})\right)}, \quad (2.18)$$

where p_i are the numbers of facets among $x_1,\ldots,x_n$ with orientations e_i and $\mathbf{p}^{(d)} := (p_1,\ldots,p_d)$. Furthermore we will make use of definition $\gamma(q,k,k,d,\mathbf{p}^{(d)})$, because specially $\gamma(q,0,0,d,\mathbf{p}^{(d)})$ is the total number of intersections of all q-tuples of the facets among $x_1,\ldots,x_n$ and $\gamma(q,k,k,d,\mathbf{p}^{(d)})$ is the total number of intersections of all q-tuples of the facets among facets $x_1,\ldots,x_n,y_1,\ldots,y_k$.
Then we substitute $\frac{\alpha I}{d}$ for $\tilde{\alpha}$, extend the both fractions by $e^{-\tilde{\alpha}(q-1)}$ and we get in the case of the lower bound of (2.18)

$$\frac{\sum_{p_1=0}^{\infty}\cdots\sum_{p_d=0}^{\infty} \frac{\tilde{\alpha}^{p_1+\cdots+p_d}}{p_1!\cdots p_d!} \exp\left(\nu_q (2\upsilon)^{d-q}\gamma(q,k,k,d,\mathbf{p}^{(d)}) - \tilde{\alpha}(q-1)\right)}{\sum_{p_1=0}^{\infty}\cdots\sum_{p_d=0}^{\infty} \frac{\tilde{\alpha}^{p_1+\cdots+p_d}}{p_1!\cdots p_d!} \exp\left(\nu_q \upsilon^{d-q}\gamma(q,0,0,d,\mathbf{p}^{(d)}) - \tilde{\alpha}(q-1)\right)}$$

and in the case of the upper bound of (2.18)

$$\frac{\sum_{p_1=0}^{\infty}\cdots\sum_{p_d=0}^{\infty} \frac{\tilde{\alpha}^{p_1+\cdots+p_d}}{p_1!\cdots p_d!} \exp\left(\nu_q \upsilon^{d-q}\gamma(q,k,k,d,\mathbf{p}^{(d)}) - \tilde{\alpha}(q-1)\right)}{\sum_{p_1=0}^{\infty}\cdots\sum_{p_d=0}^{\infty} \frac{\tilde{\alpha}^{p_1+\cdots+p_d}}{p_1!\cdots p_d!} \exp\left(\nu_q (2\upsilon)^{d-q}\gamma(q,0,0,d,\mathbf{p}^{(d)}) - \tilde{\alpha}(q-1)\right)}.$$

Then using Lemma 7 we get the limit of the lower and upper bound, which are both in the same form $\frac{\binom{d-l}{d-q+1}}{\binom{d}{d-q+1}}$. For $d \geq l = k > q$ and $y_1,\ldots,y_k \in \mathbf{B}$ we can get an upper bound

$$0 \leq \rho_k(y_1,\ldots,y_k;\mu_\alpha^{(q)}) \leq \rho_q(y_1,\ldots,y_q;\mu_\alpha^{(q)}),$$
$$\lim_{\alpha\to\infty} \rho_q(y_1,\ldots,y_q;\mu_\alpha^{(q)}) = 0.$$

Now consider more than one facet with the same orientation among $y_1,\ldots y_k$ and with $l < k < q$ distinct orientations, which are without loss of generality set to $e_1,\ldots,e_l$, then we can bound the correlation function in the following way

$$\frac{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q (k-l+1)^d (2\upsilon)^{d-q}\gamma(q,k,k,d,\mathbf{p}^{(d)})\right)}{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q \upsilon^{d-q}\gamma(q,0,0,d,\mathbf{p}^{(d)})\right)}$$
$$\leq \rho_k(y_1,\ldots,y_k;\mu_\alpha^{(q)})$$
$$\leq \frac{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q \upsilon^{d-q}\gamma(q,k,k,d,\mathbf{p}^{(d)})\right)}{\sum_{n=0}^{\infty} \frac{(\frac{\alpha I}{d})^n}{n!} \sum_{p_1+\cdots+p_d=n} \binom{n}{p_1,\ldots,p_d} \exp\left(\nu_q (k-l+1)^d (2\upsilon)^{d-q}\gamma(q,0,0,d,\mathbf{p}^{(d)})\right)}.$$

These bounds lead to expressions in the same form as in the case with unique orientations and therefore we proceed in the same way and get the value of the limit $\frac{\binom{d-l}{d-q+1}}{\binom{d}{d-q+1}}$.
For $d \geq l \geq q$ we need only lower bound for the number of intersections in form $\gamma(q, l, l, d, \mathbf{p}^{(d)})$, which forms an upper bound for the correlation function. This upper bound tends to zero.
Bounds for the numerator and denominator of the correlation function converge to their limits with at least exponential rate and we can also see that the upper bounds can be selected to depend only on the ν and q, therefore they do not depend on currently selected facets $y_1, \ldots, y_k$ in the argument of the correlation function.
Rate of convergence can be extended to the whole fraction, if we denote by $\rho^{num}(\alpha)$ the value of numerator and by $\rho^{den}(\alpha)$ the value of denominator on the left side in (2.9), respectively, and ρ^{num} and ρ^{den} the limit of the numerator and denominator on the right-hand side in (2.9), respectively, then there exist $\beta_1^{(1)}, \beta_2^{(1)}, \beta_1^{(2)}, \beta_2^{(2)} > 0$, such that for every $\alpha \in [1, \infty)$ we have

$$|\rho^{num}(\alpha) - \rho^{num}| < \beta_1^{(1)} \rho^{num} e^{-\beta_2^{(1)}\alpha}, |\rho^{den}(\alpha) - \rho^{den}| < \beta_1^{(2)} \rho^{den} e^{-\beta_2^{(2)}\alpha}.$$

If we choose $\beta_0 := \max\{-\frac{1}{\beta_2^{(2)}} \log \frac{1}{2\beta_1^{(2)}}, 1\}$, $\beta_1 := 4\max\left\{\beta_1^{(1)}, \beta_1^{(2)}\right\}$ and $\beta_2 := \min\{\beta_2^{(1)}, \beta_2^{(2)}\}$, then for $\alpha \geq \beta_0$ we get bounds

$$\frac{\rho^{num}(\alpha)}{\rho^{den}(\alpha)} - \frac{\rho^{num}}{\rho^{den}} \leq \frac{\rho^{num}}{\rho^{den}} \left(\frac{\beta_1^{(1)} e^{-\beta_2^{(1)}\alpha} + \beta_1^{(2)} e^{-\beta_2^{(2)}\alpha}}{1 - \beta_1^{(2)} e^{-\beta_2^{(2)}\alpha}} \right) \leq \frac{\rho^{num}}{\rho^{den}} \beta_1 e^{-\beta_2 \alpha},$$
$$\frac{\rho^{num}(\alpha)}{\rho^{den}(\alpha)} - \frac{\rho^{num}}{\rho^{den}} \geq -\frac{\rho^{num}}{\rho^{den}} \left(\frac{\beta_1^{(1)} e^{-\beta_2^{(1)}\alpha} + \beta_1^{(2)} e^{-\beta_2^{(2)}\alpha}}{1 + \beta_1^{(2)} e^{-\beta_2^{(2)}\alpha}} \right) \geq -\frac{\rho^{num}}{\rho^{den}} \beta_1 e^{-\beta_2 \alpha}.$$

□

Proof of Lemma 7. For $i_1, \ldots, i_4 \in \mathbb{N}$ and $t \in (0, \infty)$, let

$$\Gamma(i_1, i_2, i_3, i_4, t) := \sum_{p_1=0}^{\infty} \cdots \sum_{p_{i_4}=0}^{\infty} \frac{t^{p_1 + \cdots + p_{i_4}}}{p_1! \cdots p_{i_4}!} \exp\left(\nu\gamma(i_1, i_2, i_3, i_4, \mathbf{p}^{(i_4)}) - t(i_1 - 1)\right).$$

and we calculate $\lim_{\alpha\to\infty} \Gamma(q, l, l, d, \alpha)$. To do this firstly we calculate the values of the limit by calculating the sum over

$$(p_1 > 0 \wedge \cdots \wedge p_d > 0)$$

to show that this value tends to zero as α tends to infinity. We show this for $l = 0$ because for $l > 0$, we get upper bound using $l = 0$, because $\gamma(q, l, \cdot, \cdot, \cdot) \geq \gamma(q, 0, \cdot, \cdot, \cdot)$ and the sum is non-negative. In the following we use Chernoff bound for tail probabilities of Poisson distribution

$$\sum_{i=0}^{n} \frac{\alpha^i}{i!} \leq \frac{(e\alpha)^n}{n^n}, n < \alpha.$$

1. First we consider that all the summing variables are between 0 and $\alpha^{2/3}$:

$$\Big((\alpha^{2/3} > p_1 > 0) \wedge \cdots \wedge (\alpha^{2/3} > p_d > 0)\Big)$$

$$\sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_d>0}^{\alpha^{2/3}} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\Big(\nu\gamma(q,0,0,d,\mathbf{p}^{(d)}) - \alpha(q-1)\Big) \leq$$
$$\sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_d>0}^{\alpha^{2/3}} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\left(-\alpha(q-1)\right) \leq \left(\frac{(e\alpha)^{d\alpha^{2/3}}}{(\alpha^{2/3})^{d\alpha^{2/3}}}\right) e^{-\alpha(q-1)} \xrightarrow[\alpha\to\infty]{} 0\,,$$

where we used d times the Chernoff bound.

2. Now consider that one of the summing variables is greater than $\alpha^{2/3}$, without loss of generality we select p_d such that

$$\Big((\alpha^{2/3} > p_1 > 0) \wedge \cdots \wedge (\alpha^{2/3} > p_{d-1} > 0) \wedge (p_d \geq \alpha^{2/3})\Big)$$

and we get

$$\sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_{d-1}>0}^{\alpha^{2/3}} \sum_{p_d\geq\alpha^{2/3}}^{\infty} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\Big(\nu\gamma(q,0,0,d,\mathbf{p}^{(d)}) - \alpha(q-1)\Big)$$
$$\leq \sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_{d-1}>0}^{\alpha^{2/3}} \sum_{p_d=0}^{\infty} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\left(\nu p_d - \alpha(q-1)\right)$$
$$= \sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_{d-1}>0}^{\alpha^{2/3}} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_{d-1}!} \exp\left(\alpha e^{\nu} - \alpha(q-1)\right)$$
$$\leq \left(\frac{(e\alpha)^{(d-1)\alpha^{2/3}}}{(\alpha^{2/3})^{(d-1)\alpha^{2/3}}}\right) \exp\left(\alpha e^{\nu} - \alpha(q-1)\right) \xrightarrow[\alpha\to\infty]{} 0$$

because $e^{\nu} - (q-1) < 0$.

3. When at least two of the summing variables are greater than $\alpha^{2/3}$, without loss of generality select p_{d-1} and p_d, then we have

$$\sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_{d-2}>0}^{\alpha^{2/3}} \sum_{p_{d-1}\geq\alpha^{2/3}}^{\infty} \sum_{p_d\geq\alpha^{2/3}}^{\infty} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\Big(\nu\gamma(q,0,0,d,\mathbf{p}^{(d)}) - \alpha(q-1)\Big)$$
$$\leq \sum_{p_1>0}^{\alpha^{2/3}} \cdots \sum_{p_{d-2}>0}^{\alpha^{2/3}} \sum_{p_{d-1}\geq\alpha^{2/3}}^{\infty} \sum_{p_d\geq\alpha^{2/3}}^{\infty} \frac{\alpha^{p_1+\cdots+p_d}}{p_1! \cdots p_d!} \exp\Big(\nu\alpha^{4/3} - \alpha(q-1)\Big)$$
$$\leq \exp\Big(\nu\alpha^{4/3} + \alpha(d+1-q)\Big) \xrightarrow[\alpha\to\infty]{} 0\,.$$

4. The same applies to the case, where more than two variables are greater than $\alpha^{2/3}$, because we are able to find terms with higher power of α in the exponential.

Therefore we need only to examine the remaining terms, where at least one of the variables is equal to zero, thus we replace $\Gamma(q,l,l,d,\alpha)$ by d sums, where one variable is set to zero

$$\Gamma(q,l,l,d,\alpha) \approx l\Gamma(q,l,l-1,d-1,\alpha) + (d-l)\Gamma(q,l,l,d-1,\alpha)\,, \tag{2.19}$$

where $\approx$ is the equality after omitting the summands, which tend to zero on the left-hand side, $\Gamma(q,l,l-1,d-1,\alpha)$ is the sum after setting to zero one of the variables $p_1,\dots,p_l$, $\Gamma(q,l,l,d-1,\alpha)$ is the sum after setting to zero one of the variables $p_{l+1},\dots,p_d$ and the multiplying numbers are the counts of possible selections of these variables. It can be shown that

$$\gamma(q,i_1,i_2,q-1,\mathbf{p}^{(q-1)}) = \sum_{\substack{F\subseteq[q-1]\\ q-i_1\le|F|\le q\\ |F\cup[i_2]|+i_1-i_2\ge q}} \prod_{j\in F} p_j$$
$$\begin{cases} = \sum_{F\in\emptyset}\prod_{j\in F} p_j = 0\,, & i_1 = i_2\,,\\ \ge \sum_{\substack{F\subseteq[q-1]\\ q-1\in F}}\prod_{j\in F} p_j \ge p_{q-1}\,, & i_1 > i_2\,. \end{cases}$$

Using this we see, that it holds

$$\begin{aligned} &\lim_{\alpha\to\infty}\Gamma(q,i_1,i_2,q-1,\alpha)\\ &= \lim_{\alpha\to\infty}\sum_{p_1=0}^{\infty}\cdots\sum_{p_{q-1}=0}^{\infty}\frac{\alpha^{p_1+\dots+p_{q-1}}}{p_1!\cdots p_{q-1}!}\exp\left(\nu\gamma(q,i_1,i_2,q-1,\mathbf{p}^{(q-1)}) - \alpha(q-1)\right)\\ &= \begin{cases} 0\,, & i_1 < i_2\,,\\ 1\,, & i_1 = i_2\,. \end{cases} \end{aligned}$$

Because the series on the right-hand side of (2.19) are in the same form as the original one and we can again sum only over the indices, where at least one is equal to zero, thus we repeat the $(d-q+1)$ times step in (2.19) and we get

$$\Gamma(q,l,l,d,\alpha) \approx \sum_{j=0}^{d-q+1} t_j\Gamma(q,l,l-j,q-1,\alpha)\,, \tag{2.20}$$

where $t_j \in \mathbb{N}$. All summands tend to zero with one exception of $t_0\Gamma(q,l,l,q-1,\alpha)$ with

$$t_0 = \begin{cases} \frac{(d-l)!}{(q-1-l)!}\,, & q > l\,,\\ 0\,, & q = l\,, \end{cases} \tag{2.21}$$

which is the number of all selections of variables set to zero from $p_{l+1},\dots,p_d$ in $d-q+1$ steps. The overall speed of convergence is implied by the convergence speed of every part of the sum, which converges to its limit at least at exponential rate. □

Proof of Lemma 8. The limit of correlation function depends only on the number l of the distinct orientations among the facets $(y_1,\dots,y_{|\sigma|})$, then correlation function tends to $\frac{\binom{d-l}{d-q+1}}{\binom{d}{d-q+1}}$ and thus we can write

$$
\begin{aligned}
\int_{\mathbf{B}^{|\sigma|}} & \left(\otimes_{k=1}^{d}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_{\sigma}(y_1,\ldots,y_{|\sigma|})\rho_{|\sigma|}(y_1,\ldots,y_{|\sigma|};\mu_\alpha^{(q)})\Lambda^{|\sigma|} \\
& \qquad \times (\mathrm{d}(y_1,\ldots,y_{|\sigma|})) \\
& = \sum_{l=1}^{d}\binom{d}{l}\int_{(\mathbf{B}^{|\sigma|})_{[l]}}\left(\otimes_{k=1}^{d}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_{\sigma}(y_1,\ldots,y_{|\sigma|}) \\
& \qquad \times \rho_{|\sigma|}(y_1,\ldots,y_{|\sigma|};\mu_\alpha^{(q)})\Lambda^{|\sigma|}(\mathrm{d}(y_1,\ldots,y_{|\sigma|})) \\
\leq & \sum_{l=1}^{d}\binom{d}{l}\int_{(\mathbf{B}^{|\sigma|})_{[l]}}\left(\otimes_{k=1}^{d}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_{\sigma}(y_1,\ldots,y_{\sigma})\frac{\binom{d-l}{d-q+1}}{\binom{d}{d-q+1}}\Lambda^{|\sigma|}(\mathrm{d}(y_1,\ldots,y_{|\sigma|})) \\
& + \sum_{l=1}^{d}\binom{d}{l}\int_{(\mathbf{B}^{|\sigma|})_{[l]}}\left(\otimes_{k=1}^{d}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_{|\sigma|}(y_1,\ldots,y_{|\sigma|}) \\
& \times \left|\frac{\binom{d-l}{d-q+1}}{\binom{d}{d-q+1}} - \rho_{|\sigma|}(y_1,\ldots,y_{|\sigma|};\mu_\alpha^{(q)})\right|\Lambda^{|\sigma|}(\mathrm{d}(y_1,\ldots,y_{|\sigma|})) \\
\leq & \sum_{l=1}^{q-1}\binom{q-1}{l}\int_{(\mathbf{B}^{|\sigma|})_{[l]}}\left(\otimes_{k=1}^{d}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_{\sigma}(y_1,\ldots,y_{|\sigma|})\Lambda^{|\sigma|}(\mathrm{d}(y_1,\ldots,y_{|\sigma|})) \\
& + \beta_1 e^{-\beta_2\alpha}\sum_{l=1}^{d}\binom{d}{l}\int_{(\mathbf{B}^{|\sigma|})_{[l]}}\left(\otimes_{k=1}^{d}\left(\bar{\mathbb{H}}^{d-k}\right)^{\otimes m_k}\right)_{\sigma}(y_1,\ldots,y_{\sigma}) \\
& \qquad \times \Lambda^{|\sigma|}(\mathrm{d}(y_1,\ldots,y_{|\sigma|}))\,, \qquad (2.22)
\end{aligned}
$$

where $(\mathbf{B}^{|\sigma|})_{[k]}$ is subspace of $\mathbf{B}^{|\sigma|}$, where facets $(y_1,\ldots,y_{|\sigma|})$ use orientations $e_1,\ldots,e_k$ (each orientation is used at least by one of the facets), $\binom{d}{k}$ is the number of possible selections of orientations used. We have an upper bound for the expression in the absolute value and we can get a lower bound in the same way. □

3. Modelling and estimates of facet process

In this chapter we use definitions introduced in Chapter 1.1 and further specify them. We base this chapter on results developed in Beneš et al. [2017] and Beneš et al. [2019].

The facet process can be represented as a marked point process, where the mark is the orientation and size of the facet. Our research addresses an important problem in the statistics of spatial marked point processes given by a density with respect to the Poisson process. Observing a realization of spatial data which shall be fitted to such a model we first estimate the parameters by a method of point estimation. However, among the quantities to be estimated there may appear also the reference mark distribution which need not coincide with the observed mark distribution of the process. Both the scalar parameters and the reference mark distribution are needed e.g. when we try to simulate the model. Our aim is to estimate the reference mark distribution non-parametrically, i.e. in total to use a semiparametric approach instead of a fully parametric one.

An early paper Baddeley and Turner [2000] mentions parameter estimation of a marked point process by means of the maximum pseudolikelihood method but the authors do not identify our problem. Much more attention is paid to it in Møller and Helisová [2010] where the marks form radii of circles centered at the points of a point process given by a density with respect to the Poisson process. The random set corresponding to the union of circles in a compact window is investigated. Since an exact method is not available the authors use an approximation what means that estimation of the distribution of radii is done by methods for a Boolean model. Then an MCMC maximum likelihood method (Møller and Waagepetersen [2004]) is used for the estimation of parameters of the point process. A recent paper by Dereudre et al. [2014] deals with the same model as Møller and Helisová [2010], their goal is to use the Takacs-Fiksel estimator instead of the computationally demanding maximum likelihood method.

We present two models of a distribution of segment process (facet process in $\mathbb{R}^2$). The first model with reference directional distribution has fixed size segments and repulsive interaction among segments. We propose two methods for the estimation of the directional distribution - the parametric Takacs-Fiksel method (see Coeurjolly et al. [2011]) and a semiparametric method combining the Takacs-Fiksel method for parameters and the kernel estimation of the density. We study performance of these two approaches using simulation based on birth-death Metropolis-Hastings algorithm of Markov chain Monte Carlo from Geyer and Møller [1994].

The second model with a reference length distribution deals with segments in an ellipsoidal window and uniform directional distribution. We show how to estimate model parameters using the Takacs-Fiksel method and we examine performance of this estimate using simulated data. Moreover we apply the suggested method to the real data from fluorescence imaging of stress fibres in adult human mesenchymal stem cells. The stem cells have been cultured on gels for a time span of 24 hours. We evaluate especially the cells on a low stiffness gel since they present more randomness and less inhomogeneity than the others. Using the Filament Sensor [Eltzner et al., 2016] algorithm it is possible to transform the raw data onto a system of segments. Here, the true window corresponds to a single cell which is roughly approximated by an ellipse. The parameters of this model are estimated and the degree of fit of real data with the model is then tested using various statistics by means of Monte-Carlo testing.

3.1 Models

We will consider two models - both present facet processes in $\mathbb{R}^2$, i.e. segment processes.

3.1.1 Model I

The segment process with a reference directional distribution. Let $D \subseteq \mathbb{R}^2$ be bounded and measurable planar set such that $\lambda^2(D) > 0$, let $\mathbb{S}^1 = [0, \pi)$ be the semicircle of axial directions, $\upsilon > 0$ is fixed segment size, and put

$$\begin{aligned} \mathbf{B} &:= D \times \{\upsilon\} \times [0, \pi)\,, \\ \tilde{\mathbf{B}} &:= \iota(\mathbf{B})\,, \end{aligned}$$

as space of segment parameters and segments, where ι is defined in (1.9). A segment $\iota((z, \upsilon, \phi)) \in \tilde{\mathbf{B}}$ has centre z, υ fixed segment size and direction ϕ.

We will use the Poisson segment process η with the intensity measure Λ on $\mathbf{B}$, where

$$\Lambda(\mathrm{d}(z, r, \phi)) := \mathrm{d}z \delta_\upsilon(r) \mathrm{d}r \frac{1}{\pi} \mathrm{d}\phi\,.$$

Let the segment process μ have a density g w.r.t. η, in form:

$$g(\mathbf{x}) := c\alpha^{\mathbf{x}(\mathbf{B})} \exp(\gamma_0 G_2(\mathbf{x})) \prod_{(z,\upsilon,\phi)\in\mathbf{x}} w(\phi)\,, \tag{3.1}$$

where w is probability density on $[0, \pi)$; $\alpha > 0$, $\gamma_0 \leq 0$ are parameters; c is the normalizing constant,

$$G_2(\mathbf{x}) := \int_{\mathbf{B}^2} \mathbf{1}[\iota(y_1) \cap \iota(y_2) \neq \emptyset] \mathbf{x}^{(2)}(\mathrm{d}(y_1, y_2))\,,$$

is the total number of intersections among segments in $\mathbf{x}$.

The conditional intensity corresponding to the density g in (3.1) is defined by

$$\Psi((z, \upsilon, \phi), \mathbf{x}) = \alpha \exp\left(\gamma_0 \int_{\mathbf{B}} \mathbf{1}[\iota((z, \upsilon, \phi)) \cap \iota(y) \neq \emptyset] \mathbf{x}(\mathrm{d}y)\right) w(\phi)\,, \qquad (z, \upsilon, \phi) \notin \mathbf{x}\,. \tag{3.2}$$

Furthermore let the directional density w be that of von Mises distribution on $[0, \pi)$ with parameters $\gamma_1 \geq 0$, $\gamma_2 \in [0, \pi)$, which is suitable for unimodal distribution (γ_2 is the mode and γ_1 reflects the concentration around γ_2) :

$$w(\phi) := \tilde{c}_{\gamma_1} \exp(\gamma_1 \cos(2(\phi - \gamma_2)))\,, \phi \in [0, \pi)\,.$$

3.1.2 Model II

The segment process with reference length distribution. The next model is working with a random segment length. The shape of the window and the location of segments take into account the intended application on the real data. Consider an ellipse $D \subseteq \mathbb{R}^2$ centred in the origin, with axes lengths $e_1 \geq e_2 > 0$ and area $\lambda^2(D) := \frac{\pi e_1 e_2}{4}$. Let be

$$\mathbf{B} := D \times (0, e_1] \times [0, \pi)\,,$$
$$\tilde{\mathbf{B}} := \iota(\mathbf{B})\,,$$

as space of segment parameters and segments.

Let the Poisson segment process η have the intensity measure Λ on $\mathbf{B}$ of the form

$$\Lambda(\mathrm{d}(z, r, \phi)) := \frac{1}{\pi e_1} \mathrm{d}z\mathrm{d}r\mathrm{d}\phi\,.$$

Let the segment process μ have a density g with respect to η, we consider model II:

$$g(\mathbf{x}) := c\alpha^{\mathbf{x}(\mathbf{B})} \exp(\gamma_0 F(\mathbf{x})) \prod_{(z,r,\phi)\in\mathbf{x}} w\left(\frac{r}{e_1}\right) \mathbf{1}[\iota((z, r, \phi)) \subseteq D]\,,$$

with parameters $\gamma_0 \in \mathbb{R}$, $\alpha > 0$, c is the normalizing constant, w is a reference probability density of normalized lengths on $(0, 1]$ and

$$F(\mathbf{x}) := \int_{\mathbf{B}} f(y)\mathbf{x}(\mathrm{d}y)\,, f(y) := \max_{x\in\iota(y)} \frac{||x||}{e_1}\,.$$

For $y \in \mathbf{B}$ the $f(y) \in (0, \frac{1}{2})$ is a distance of the most distant point of the segment y from the centre of the cell. The corresponding conditional intensity is

$$\Psi((z, r, \phi), \mathbf{x}) = \alpha \exp(\gamma_0 f((z, r, \phi)))w\left(\frac{r}{e_1}\right) \prod_{y\in\mathbf{x}\cup(z,r,\phi)} \mathbf{1}[\iota(y) \subseteq D]\,.$$

Furthermore suppose that the reference length density w of the beta distribution with parameters $\gamma_1, \gamma_2 > 0$

$$w(x) := \frac{x^{\gamma_1-1}(1 - x)^{\gamma_2-1}}{B(\gamma_1, \gamma_2)}\,, x \in (0, 1)\,,$$

where $B(\cdot, \cdot)$ is the beta function in the denominator.

3.2 Estimates

3.2.1 Takacs-Fiksel

We will demonstrate the general use of Takacs-Fiksel method of parameter estimation in Gibbs point process. From the Georgii-Nguyen-Zessin formula

$$\mathbb{E}\left[\int_{\mathbf{B}} h(y, \mu \setminus \{y\})\mu(\mathrm{d}y)\right] = \int_{\mathbf{B}} \mathbb{E}(\Psi(y, \mu)h(y, \mu))\mathrm{d}y\,, \tag{3.3}$$

we obtain the so called innovation

$$\int_{\mathbf{B}} h(y,\mu \setminus \{y\})\mu(\mathrm{d}y) - \int_{\mathbf{B}} \Psi(y,\mu)h(y,\mu)\mathrm{d}y\,,$$

which is a centered random variable. Using suitable test functions h and setting innovations equal to zero leads to a system of equations for the unknown parameters and the integral can be for the estimation replaced by its Monte Carlo approximation. Then in general the Takacs-Fiksel estimate is solution of system of equations

$$\int_{\mathbf{B}} h_k(y,\mu \setminus \{y\})\mu(\mathrm{d}y) = \int_{\mathbf{B}} \Psi(y,\mu)h_k(y,\mu)\tilde{\mathbf{x}}(\mathrm{d}y)\,, k \in [m]\,,$$

where $\tilde{\mathbf{x}}$ is random sample from $\mathbf{B}$ represented by a random measure and ν is vector of optimized parameters, which affect the innovation via Ψ. If the conditional intensity is in form $\Psi(y,\mathbf{x}) = \exp(\sum_{k=1}^{m} \nu_k F_k(y,\mathbf{x}))$, then the optimal choice of h_k is

$$h_k(y,\mathbf{x}) := \frac{\partial}{\partial \nu_k} \log(\Psi(y,\mathbf{x})) = F_k(y,\mathbf{x})\,.$$

3.2.2 Semiparametric estimate

We use this estimate for a directional distribution density of a Model I, that has von Mises distribution.

Definition 33. *We define measure $\overline{\Lambda}_\mu^{cen}$ on space $(D,\mathcal{B}(D))$ and measure $\overline{\Lambda}_\mu$ on product space $(D\times[0,\pi),\mathcal{B}(D\times[0,\pi)))$*

$$\overline{\Lambda}_\mu^{cen}(K) := \mathbb{E}\int_{\mathbf{B}} \delta_z(K)\mu(\mathrm{d}(z,r,\phi))\,, \qquad K\in\mathcal{B}(D)\,,$$
$$\overline{\Lambda}_\mu(K\times L) := \mathbb{E}\int_{\mathbf{B}} \delta_z(K)\delta_\phi(L)\mu(\mathrm{d}(z,r,\phi))\,, \qquad K\in\mathcal{B}(D), L\in\mathcal{B}([0,\pi))\,,$$

where $\mathcal{B}(\cdot)$ denote Borel sets.

Theorem 9. *For fixed $L\in\mathcal{B}([0,\pi))$, the $\overline{\Lambda}_\mu(\cdot\times L)$ is absolutely continuous with respect to the $\overline{\Lambda}_\mu^{cen}(\cdot)$ and with Radon-Nikodym density $P(\cdot,L)$, i.e.*

$$\overline{\Lambda}_\mu(K\times L) = \int_K P(z,L)\overline{\Lambda}_\mu^{cen}(\mathrm{d}(z))\,.$$

For given $z\in D$, $P(z,\cdot)$ is absolutely continuous with respect to the λ measure with Radon-Nikodym density $w_\mu^{(z)}(\phi)$

$$w_\mu^{(z)}(\phi) = \frac{\rho_1((z,\upsilon,\phi);\mu)}{\int_0^\pi \rho_1((z,\upsilon,\overline{\phi});\mu)\mathrm{d}\overline{\phi}}\,.$$

Proof. Absolute continuity in both cases is trivial and can be seen right from the definition. For Borel sets $K\subseteq D$, $L\subseteq[0,\pi)$ it holds using the Campbell theorem

$$\overline{\Lambda}_\mu(K\times L) = \frac{1}{\pi}\int_K\int_L \rho_1((z,\upsilon,\phi);\mu)\mathrm{d}\phi\mathrm{d}z\,.$$

Specially for $L = [0, \pi)$ we have

$$\overline{\Lambda}_\mu(K \times [0,\pi)) = \overline{\Lambda}^{cen}_\mu(K) = \frac{1}{\pi}\int_K \int_0^\pi \rho_1((z,\upsilon,\phi);\mu)\mathrm{d}\phi\mathrm{d}z\,.$$

We see that

$$P(z,L) = \frac{\int_L \rho_1((z,\upsilon,\phi);\mu)\mathrm{d}\phi}{\int_0^\pi \rho_1((z,\upsilon,\phi);\mu)\mathrm{d}\phi}$$

and the results follows. □

For $x = (z,\upsilon,\phi)$ we get from (3.2) that

$$\rho_1((z,\upsilon,\phi);\mu) = \alpha w(\phi)\mathbb{E}\exp\left(\gamma_0\int_{\mathbf{B}} \mathbf{1}[\iota((z,\upsilon,\phi))\cap\iota(y)\neq\emptyset]\mu(\mathrm{d}y)\right) \tag{3.4}$$

is the correlation function of μ. The expectation is not analytically tractable, therefore Baddeley and Nair [2012] suggest an approximation

$$\mathbb{E}\exp\left(\gamma_0\int_{\mathbf{B}} \mathbf{1}[\iota(x)\cap\iota(y)\neq\emptyset]\mu(\mathrm{d}y)\right) \approx \mathbb{E}\exp\left(\gamma_0\int_{\mathbf{B}} \mathbf{1}[\iota(x)\cap\iota(y)\neq\emptyset]\eta_\rho(\mathrm{d}y)\right)\,, \tag{3.5}$$

where $\eta(\rho)$ is a Poisson process with intensity function $\rho_1((z,r,\phi);\mu)\Lambda(\mathrm{d}(z,r,\phi))$.

Lemma 10. *For $y = (z,r,\phi) \in \mathbf{B}$ we have*

$$\mathbb{E}\exp\left(\gamma_0\int_{\mathbf{B}} \mathbf{1}[\iota(x)\cap\iota(y)\neq\emptyset]\eta_\rho(\mathrm{d}y)\right) = \exp\left((e^{\gamma_0}-1)\int_{M_y}\rho_1(x;\mu)\mathrm{d}x\right)\,,$$

where $M_y := \{x \in \mathbf{B} \mid \iota(y)\cap\iota(x)\neq\emptyset\}$.

Proof. Using Theorem 2 we have

$$\begin{aligned}
&\mathbb{E}\exp\left(\gamma_0\int_{\mathbf{B}} \mathbf{1}[\iota(x)\cap\iota(y)\neq\emptyset]\eta_\rho(\mathrm{d}y)\right)\\
&= e^{-\int_{\mathbf{B}}\rho_1(y;\mu)\Lambda(\mathrm{d}y)}\sum_{n=0}^{\infty}\frac{1}{n!}\prod_{i=1}^{n}\int_{\mathbf{B}}\exp(\gamma_0\mathbf{1}[\iota(x)\cap\iota(y_i)\neq\emptyset])\rho_1(y_i;\mu)\Lambda(\mathrm{d}y_i)\\
&= \exp\left(\int_{\mathbf{B}}\left(\exp(\gamma_0\mathbf{1}[\iota(x)\cap\iota(y)\neq\emptyset])-1\right)\rho_1(y;\mu)\Lambda(\mathrm{d}y)\right)\\
&= \exp\left((e^{\gamma_0}-1)\int_{M_x}\rho_1(y;\mu)\mathrm{d}y\right)\,.
\end{aligned}$$

□

From Proposition 1 we have

$$\rho_1((z,\upsilon,\phi);\mu) = \overline{c}_z w^{(z)}_\mu(\phi)\,, \phi\in[0,\pi)\,,$$

where $\overline{c}_z$ is normalizing constant, $z \in D$. Assuming that there exists a stationary segment process in $\mathbb{R}^2$ with given conditional intensity (this time the conditional intensity cannot be defined by means of densities, but from the energy function), we have that $\overline{c} := \overline{c}_z$, $w^{(z)}_\mu(\phi) = w_\mu(\phi)$ do not depend on z in the extension of D onto the whole $\mathbb{R}^2$. Under this assumption we estimate

$$\begin{aligned}
\int_{M_{(z,\upsilon,\phi)}} w^{(z)}_\mu(\phi')\mathrm{d}z'\mathrm{d}\phi' &\approx \int_{\mathbb{R}^2\times[0,\pi)}\mathbf{1}[\iota((z,\upsilon,\phi))\cap\iota((z',\upsilon,\phi'))\neq\emptyset]w_\mu(\phi')\mathrm{d}z'\mathrm{d}\phi'\\
&= \upsilon^2\int_0^\pi|\sin(\phi'-\phi)|w_\mu(\phi')\mathrm{d}\phi'\,.
\end{aligned}$$

We can then express the desired density w approximately as

$$w(\phi) \approx \frac{C w_\mu(\phi)}{\alpha \exp((e^{\gamma_0} - 1)C\upsilon^2 I(\phi))}\,, \tag{3.6}$$

where C is normalizing constant and

$$I(\phi) = \int_0^\pi |\sin(\phi' - \phi)| w_\mu(\phi') \mathrm{d}\phi'\,.$$

3.3 Results

3.3.1 Model I - Parametric estimate

We first suppose that the parameters γ_0, γ_1, α are unknown and υ, γ_2 are known, where the former is directly observable and the latter we assume to know. The asymptotic properties of this estimator are studied in Coeurjolly et al. [2011]. We use test functions

$$\begin{aligned} h_1(y,\mu) &= \int_{\mathbf{B}} \mathbf{1}[\iota(y) \cap \iota(x) \neq \emptyset]\mu(\mathrm{d}x)\,, \\ h_2(y,\mu) &= 1\,, \\ h_3((z,\upsilon,\phi),\mu) &= \int_{\mathbf{B}} \cos(2(\phi - \gamma_2))\mu(\mathrm{d}(z',\upsilon,\phi'))\,. \end{aligned}$$

Dividing the first and third equations by the second one we obtain a system of two equations for unknown γ_0, γ_1:

$$\begin{aligned} &\frac{G_2(\mu)}{\mu(\mathbf{B})} \\ &= \frac{\int_{\mathbf{B}}\int_{\mathbf{B}} c(\gamma_0,\gamma_1,\gamma_2,x)\mathbf{1}[\iota(y)\cap\iota(x)\neq\emptyset]\mu(\mathrm{d}y)\tilde{\mathbf{x}}(x)}{\int_{\mathbf{B}} c(\gamma_0,\gamma_1,\gamma_2,x)\tilde{\mathbf{x}}(\mathrm{d}x)}\,, \\ &\frac{\int_{\mathbf{B}} \cos(2(\phi-\gamma_2))\mu(\mathrm{d}(z,\upsilon,\phi))}{\mu(\mathbf{B})} \\ &= \frac{\int_{\mathbf{B}} c(\gamma_0,\gamma_1,\gamma_2,(z,\upsilon,\phi))\cos(2(\phi-\gamma_2))\tilde{\mathbf{x}}(\mathrm{d}(z,\upsilon,\phi)}{\int_{\mathbf{B}} c(\gamma_0,\gamma_1,\gamma_2,x)\tilde{\mathbf{x}}(x)}\,, \end{aligned}$$

where

$$\begin{aligned} c(\gamma_0,\gamma_1,\gamma_2,(z,\upsilon,\phi)) = \exp(\gamma_0 \int_{\mathbf{B}} &\mathbf{1}[\iota((z',\upsilon,\phi')) \cap \iota((z,\upsilon,\phi) \neq \emptyset]\mu(\mathrm{d}(z',\upsilon,\phi')) \\ &+ \gamma_1 \cos(2(\phi - \gamma_2)))\,. \end{aligned}$$

On the left hand side of the equations we have statistics of the data μ. Having solved (numerically) the above system of two equations we estimate the third parameter α as

$$\alpha = \left(\frac{\pi\lambda(D)}{\mu(\mathbf{B})I_0(\kappa)\tilde{\mathbf{x}}(\mathbf{B})}\int_{\mathbf{B}} c(\gamma_0,\gamma_1,\gamma_2,x)\tilde{\mathbf{x}}(x)\right)^{-1}\,. \tag{3.7}$$

For a numerical demonstration of the Takacz-Fiksel method we simulated 100 realizations of the segment process on $[0,1]^2$ of model I with parameters $\gamma_0 = -3$, $\gamma_1 = 1$, $\gamma_2 = \pi/2$, $\upsilon = 0.06$, $\alpha = 1000$.

The Takacz-Fiksel estimators of γ_0, γ_1, α were evaluated with random sample of 1000 segments for each realization. Empirical means and standard deviations (sd) of the estimators are printed in Table 3.1.

Model I	true	mean	sd
γ_0	-3	-2.998	0.299
γ_1	1	0.999	0.078
α	1000	1008.9	64.2

Table 3.1: Means and standard deviations (sd) of Takacz-Fiksel estimates from 100 simulations on $[0,1]^2$ of model I (with observable parameters $\gamma_2 = \pi/2$, $\upsilon = 0.06$). The true values of estimated parameters are in the table.

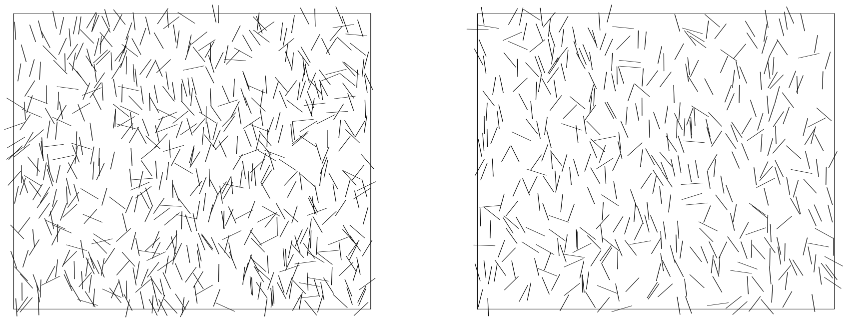

Figure 3.1: Simulated realizations of model I segment processes on $[0,1]^2$ with parameters $\gamma_1 = 1$, $\gamma_2 = \pi/2$, $\upsilon = 0.06$, $\alpha = 1000$ in the left pattern we have $\gamma_0 = -0.5$, and statistics $\mu(\mathbf{B}) = 624$, $G_2(\mu) = 204$. In the right pattern we have $\gamma_0 = -3$ (more repulsion) and statistics $\mu(\mathbf{B}) = 433$, $G_2(\mu) = 5$.

3.3.2 Model II

The parameters e_1, e_2 are known as described in the subsection on cell shape.

Here we use four test functions

$$\begin{aligned}
h_1((z,r,\phi),\mathbf{x}) &= f((z,r,\phi))\,,\\
h_1((z,r,\phi),\mathbf{x}) &= \log\left(\frac{r}{e_1}\right)\,,\\
h_1((z,r,\phi),\mathbf{x}) &= \log\left(1-\frac{r}{e_1}\right)\,,\\
h_1((z,r,\phi),\mathbf{x}) &= 1\,,
\end{aligned}$$

the second and third is not derived using the derivative of conditional intensity. We divide the first three equations by the fourth one and we obtain system for

Model II	true	mean	sd
γ_0	0.5	0.525	0.463
γ_1	3	3.089	0.296
γ_2	3	3.147	0.458
α	100	104.32	40.25

Table 3.2: Means and standard deviations (sd) of Takacz-Fiksel estimates from 100 simulations of model II with parameters $e_1 = e_2 = 1$, $\alpha = 100$, $\gamma_0 = 0.5$, $\gamma_1 = 3$, $\gamma_2 = 3$. The true values of estimated parameters are in the table.

unknown γ_0, γ_1, γ_2:

$$\frac{\int_{\mathbf{B}} F((z,r,\phi))\mu(\mathrm{d}(z,r,\phi))}{\mu(\mathbf{B})} = \frac{\int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})F((z,r,\phi))\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))}{\int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))},$$

$$\frac{\int_{\mathbf{B}} \log(r)\mu(\mathrm{d}(z,r,\phi))}{\mu(\mathbf{B})} = \frac{\int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})\log(r)\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))}{\int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))},$$

$$\frac{\int_{\mathbf{B}} \log(e_1 - r)\mu(\mathrm{d}(z,r,\phi))}{\mu(\mathbf{B})} = \frac{\int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})\log(e_1 - r)\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))}{\int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))},$$

and finally we put the estimator of γ_0 in the equation for α

$$\alpha = \frac{\tilde{\mathbf{x}}(\mathbf{B})\mu(\mathbf{B})}{\Lambda^3(\mathbf{B}) \int_{\mathbf{B}} \mathbf{1}[\iota((z,r,\phi)) \subseteq D] \exp(\gamma_0 F((z,r,\phi)))w(\frac{r}{e_1})\tilde{\mathbf{x}}(\mathrm{d}(z,r,\phi))}. \tag{3.8}$$

For a numerical demonstration of the Takacz-Fiksel method we simulated $n = 100$ realizations of the segment process of model II with parameters $e_1 = e_2 = 1$, $\alpha = 100$, $\gamma_0 = 0.5$, $\gamma_1 = 3$, $\gamma_2 = 3$.
The Takacz-Fiksel estimators of α, γ_0, γ_1, γ_2 were evaluated with $\tilde{\mathbf{x}}(B) = 1000$ for each realization. Empirical means and standard deviations (sd) of the estimators are printed in Table 3.2.

3.3.3 Model I - Semiparametric estimate

In this section we suggest a method of estimation of parameters C, γ_0, α and the density w from the previous section using the Takacs-Fiksel method. From formula (3.3) we obtain innovation equations

$$\int_{\mathbf{B}} h(y, \mu \setminus \{y\})\mu(\mathrm{d}y) - \int_{\mathbf{B}} \Psi(y,\mu)h(y,\mu)\mathrm{d}y = 0 \tag{3.9}$$

and solve them for various test functions h. We take Ψ from (3.2) where we insert approximation (3.6) for unknown w. First the density w_μ is estimated using a kernel estimator for directional data Mardia and Jupp [1999]. Then put

$$\beta(\gamma_0, v, C, \phi) = \exp((e^{\gamma_0} - 1)v^2 CI(\phi)),$$

	true	mean	sd	CV
γ_0	-0.5	-0.496	0.071	0.14
α	1000	1011	154.7	0.15
	true	mean	sd	CV
γ_0	-3	-3.03	0.356	0.12
α	1000	976	141.0	0.14

Table 3.3: Empirical mean, standard deviation (sd) and coefficient of variation (CV) of Takacs-Fiksel estimates of scalar parameters in the model having density (3.1) with reference directional distribution. It is based on 100 simulations, the two cases correspond to $\gamma_0 = -0.5$, $\gamma_0 = -3$.

and we estimate C, γ_0 from the system of Takacs-Fiksel equations:

$$\begin{aligned} G_2(\mu) &= \frac{\pi^2 \Lambda(D) C}{\tilde{\mathbf{x}}(\mathbf{B})} \int_{\mathbf{B}} \frac{w_\mu(\phi) G_2^\mu((z,r,\phi)) e^{\gamma_0 G_2^\mu((z,r,\phi))}}{\exp((e^{\gamma_0}-1) v^2 C I(\phi))} \tilde{\mathbf{x}}(\mathrm{d}((z,r,\phi))) \\ \mu(\mathbf{B}) &= \frac{\pi^2 \Lambda(D) C}{\tilde{\mathbf{x}}(\mathbf{B})} \int_{\mathbf{B}} \frac{w_\mu(\phi) e^{\gamma_0 G_2^\mu((z,r,\phi))}}{\exp((e^{\gamma_0}-1) v^2 C I(\phi))} \tilde{\mathbf{x}}(\mathrm{d}((z,r,\phi))) \\ G_2^\mu(x) &= \int_{\mathbf{B}} \mathbf{1}[\iota(y) \cap \iota(x) \neq \emptyset] \mu(\mathrm{d}(y)) \,. \end{aligned}$$

Here in the innovations equations we take score functions $h(y,\mu) = G_2^\mu(y)$ and $h(y,\mu) = 1$, respectively, the integrals in the second term of (3.9) are evaluated by Monte Carlo method using $\tilde{\mathbf{x}}(\mathbf{B})$ independent simulations of segments uniformly distributed in $\mathbf{B}$. Then we plug the estimators of C, γ_0 in a formula obtained by integrating (3.6)

$$\alpha = \frac{\pi C}{\tilde{\mathbf{x}}(\mathbf{B})} \int_{\mathbf{B}} \frac{w_\mu(\phi)}{\exp((e^{\gamma_0}-1) v^2 C I(\phi))} \tilde{\mathbf{x}}(\mathrm{d}((z,r,\phi))) \,,$$

and finally estimate w from (3.6).

A numerical study is based on twice 100 simulated realizations of segment process with parameters $\gamma_1 = 1$, $\gamma_2 = 0$, $\alpha = 1000$, $v = 0.12$ on $[0,1]^2 \times \{v\} \times [-\frac{\pi}{2}, \frac{\pi}{2}]$. The two cases I, II investigated are $\gamma_0 = -0.5$, $\gamma_0 = -3$, respectively. The results are in Table 3.3 and in Figures 3.2 and 3.3.

In Table 3.2 we observe a small difference between the true and mean values for both estimates of γ_0 and α. The coefficient of variation

$$CV = \frac{\mathrm{sd}}{|\mathrm{mean}|}$$

is also comparable. In Figure 3.2 we can observe how the kernel estimator of the observed directional distribution differs from the true reference directional distribution (von Mises). The results in Figure 3.3 suggest that the estimate of the reference density is slightly better (smaller bias and variability) for the case I than for the case II. We conclude that the approximation (3.5) works well in the Takacs-Fiksel method here.

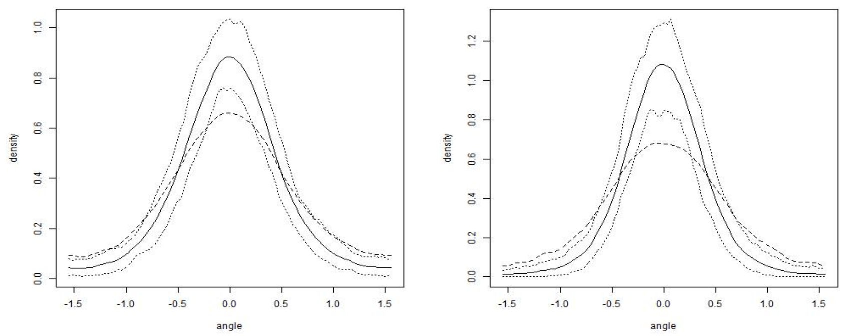

Figure 3.2: Kernel estimation of the observed directional density based on 100 simulations of the segment process μ, $\gamma_0 = -0.5$ (left), $\gamma_0 = -3$ (right). The average kernel estimator of the observed directional density (full line) compared to the true reference density (dashed line) of von Mises distribution with parameters $\gamma_1 = 0$, $\gamma_2 = 1$. The envelopes (dotted lines) correspond to empirical 90% confidence interval for the kernel estimator, pointwise in 100 points on horizontal axis.

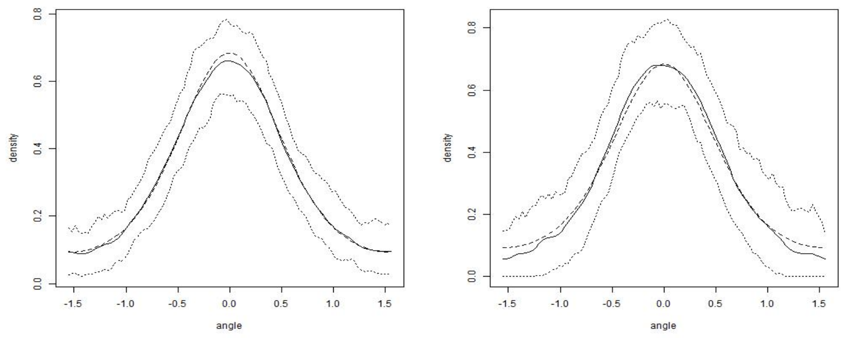

Figure 3.3: Semiparametric estimation based on 100 simulations of the segment process μ, $\gamma_0 = -0.5$ (left), $\gamma_0 = -3$ (right). The average estimator of the reference density (full line) compared to the true reference density (dashed line) of von Mises distribution with parameters $\gamma_1 = 0$, $\gamma_2 = 1$. The envelopes (dotted lines) correspond to empirical 90% confidence interval for the estimated reference density, pointwise in 100 points on horizontal axis.

3.4 Real data

3.4.1 Description

In many applications, systems of randomly dispersed segments in the plane or space are investigated. In biology, such systems occur e.g. when using fluorescence imaging to observe stress fibres in stem cells. Real data from an ongoing research consists of actin stress fibres in human mesenchymal stem cells (hMSCs) taken from the bone marrow. In the experiment, stem cells have been cultured on gels of different stiffness for 24 hours. This stiffness is given in terms of the Young's modulus, the ratio of stress by strain, i.e. the force per area needed to deform the material.

Earlier experiments have found that hMSCs can be mechanically guided to differentiate towards various cell types depending on the substrate elasticity they are grown on, namely neuron precursor cells for 1 kPa, muscle precursor cells for 10 kPa and bone precursor cells for 30 kPa Engler et al. [2006]. Especially the differentiation into neuron precursor cells is remarkable, since hMSC stem from the mesodermal tissue layer, while neurons are ectodermal cells. It has also been found that these three populations of cells on different gels express significantly disparate fibre patterns after 24 hours on the gel, Zemel et al. [2010]. It is therefore interesting to closely examine the stress fibre patterns especially for cells on a gel with 1 kPa stiffness. Here we investigate group $G1$ of $n_1 = 138$ cells which corresponds to a Young's modulus of 1 kPa and it is mostly suitable for a simple stochastic modeling.

Using the Filament Sensor algorithm Eltzner et al. [2016] it is possible to transform the raw data into a system of segments. Fig. 3.4 shows an example cell and the automatic line detection result. The corresponding segment systems of each cell are characterized by the following geometrical parameters:

- cell shape,
- spatial distribution of segments,
- length distribution of segments,
- directional distribution of segments.

In the following we suggest a methodology of quantitative description of these attributes using pixilized input from automatic image analysis.

3.4.2 Description of the cell shape

The cell shape is determined automatically by the Filament Sensor. From the raw shape, the program derives the center point, the area and the aspect ratio of the cell. This yields a simple approximation of the shape by an ellipse. An elliptical window with axes $e_1 \geq e_2$ can be realized in our simulations, however this suggests restricting attention to cells whose real shape closely fits the elliptical approximation. For the example cell from Fig. 3.4 we illustrate the corresponding elliptical approximation in Fig. 3.5. If A is the cell and B is the ellipse, the number of pixels of ellipsis outside of the cell is $card\,(B \setminus A)$ and the number of

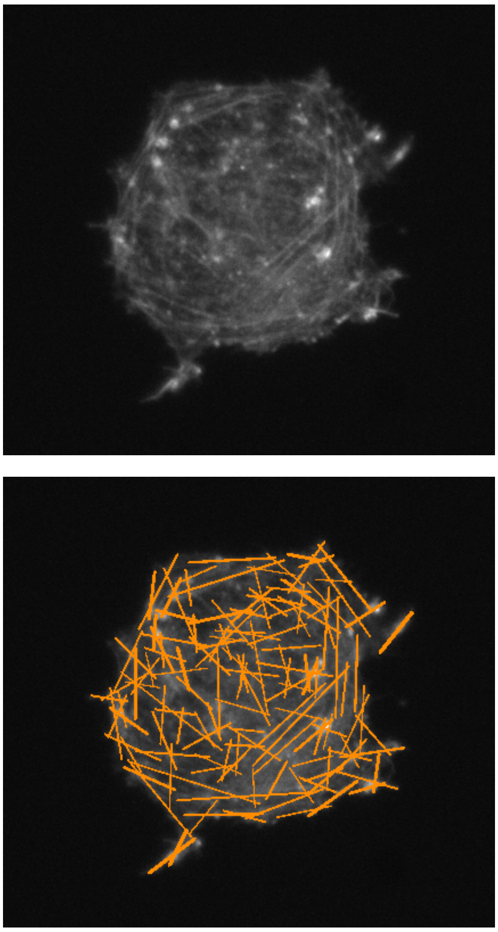

Figure 3.4: Cell 12 on 1 kPa gel, original microscopy image (upper image) and microscopy image overlayed with fibers (lower image) detected by the Filament Sensor Eltzner et al. [2016].

pixels of cell outside of ellipsis is $card\,(A \setminus B)$ and to find ideal ellipsis we minimize the sum of these two numbers. For the cells considered in our study, less than 10% of the cells' pixels lie outside the ellipse. Of the segment pixels even less than 3% lie outside the ellipse for every image. For these images we can therefore consider the elliptical approximation of the cell shape as sufficient and to accept an assumption that the segments are completely included in the ellipse.

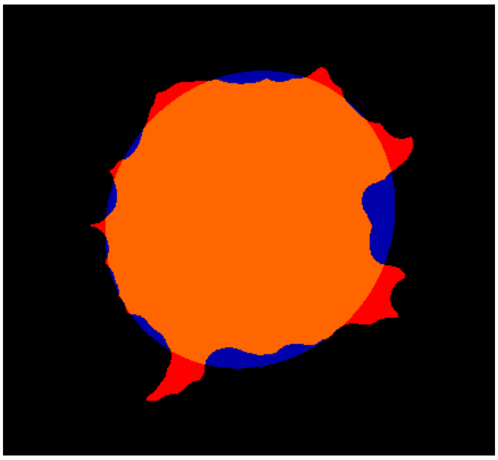

Figure 3.5: The shape of cell 12 on 1 kPa gel as detected by the Filament Sensor Eltzner et al. [2016] compared to the elliptical approximation.

3.4.3 Real data estimates

In Fig. 3.4.3 there are twenty cells from group $G1$ for further analysis. Those cells were selected which fit the ellipsoidal shape best. The cell numbers and the relative amount of pixels outside the ellipse are given in Tab. 3.4.

No.	AOE [%]	No.	AOE [%]
001	3.03	043	5.87
002	4.98	049	4.63
005	7.96	055	3.61
006	5.76	059	2.78
018	6.91	060	5.00
019	8.39	064	6.99
020	7.20	092	6.98
030	9.75	093	7.01
031	7.75	127	5.59
034	7.71	131	5.50

Table 3.4: Shape description: the numbers of cells from group $G1$ investigated and their area outside ellipse (AOE) in percent of pixels.

The selected cells were then fitted to the model II with reference length distribution. The estimated parameter values are in the Table 3.5.

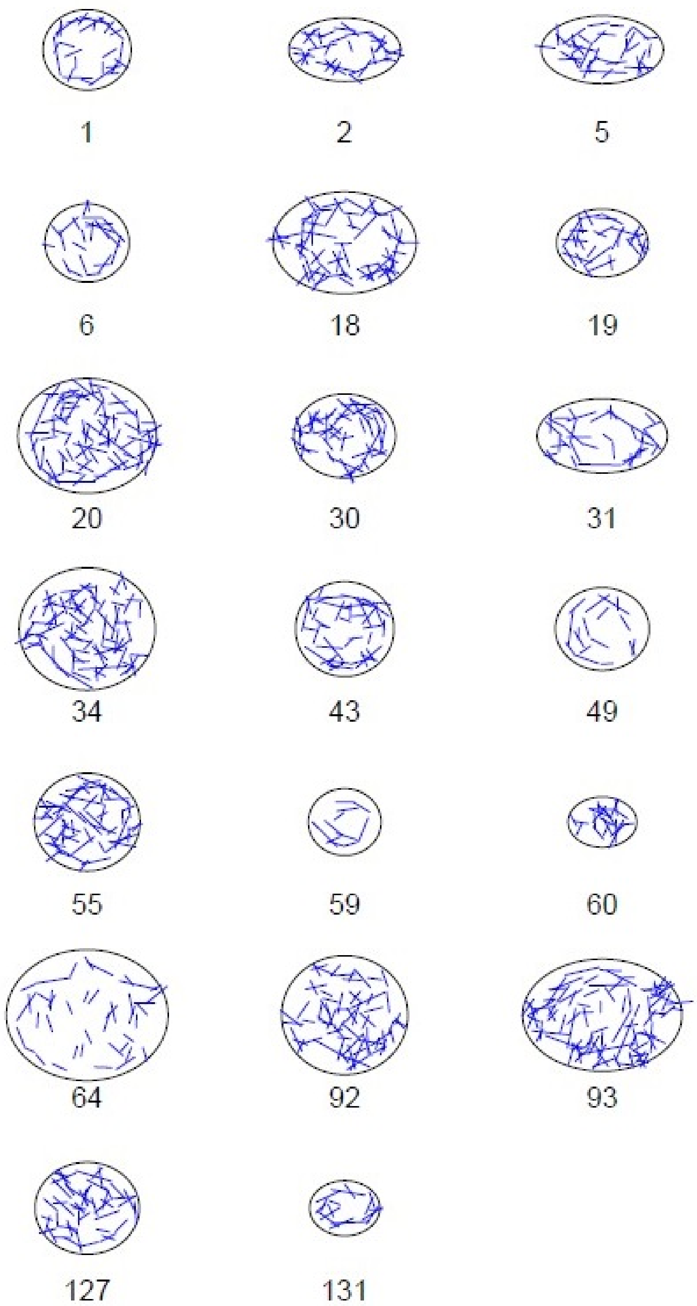

Figure 3.6: Analysed segment systems corresponding to stress fibres in cells from group $G1$ and their numbers. The shape of the cell is approximated by ellipse with axes lengths in Table 3.5.

No	e_1	e_2	$\alpha.10^3$	$\gamma_0.10^2$	γ_1	γ_2	$I.10^2$
001	59	54	0.46	4.99	1.76	4.27	0.83
002	76	43	1.67	2.88	1.58	4.35	1.13
005	83	46	4.49	0.45	1.59	4.50	0.97
006	57	53	1.98	1.70	2.00	4.47	0.76
018	97	69	1.77	1.65	1.34	4.95	1.03
019	62	46	1.12	3.97	1.87	4.05	1.05
020	94	78	4.07	0.41	1.41	4.67	1.12
030	69	57	3.54	1.13	1.75	4.38	1.27
031	88	50	1.18	2.40	1.44	4.46	0.81
034	92	83	9.44	-1.05	1.43	4.71	1.07
043	67	64	2.43	1.38	1.81	4.52	0.95
049	63	56	2.54	-0.02	1.86	4.34	0.48
055	72	66	2.25	1.85	1.91	5.22	1.20
059	49	45	19.03	-5.41	2.74	4.42	0.49
060	46	34	19.10	-2.57	2.84	4.92	1.31
064	109	89	1.48	0.28	1.29	5.75	0.42
092	86	80	7.89	-1.00	1.69	5.91	0.98
093	108	76	1.98	1.42	1.25	4.63	1.13
127	70	65	7.63	-0.60	1.65	4.28	1.13
131	48	34	0.14	10.26	5.73	12.3	0.90

Table 3.5: The results of the Takacz-Fiksel estimator with reference beta length distribution, model II. The columns involve subsequently: the cell number, the axes lengths e_1, e_2 (in pixels, where 1 pixel=0.32 μm), estimated parameters α, γ_0, γ_1 ,γ_2 and the ratio $I = \frac{n(\mathbf{x})}{e_1 e_2}$ which is proportional to number density of segments.

3.4.4 Testing of the fit of the model

Once we have estimated parameters of the model from the data, it is necessary to test whether the model fits the data well. Monte Carlo tests are common in spatial statistics, which are based on some scalar or functional test statistics of the data pattern Møller and Waagepetersen [2004]. Then we simulate realizations of the model based on estimated parameters, and find upper and lower limits of values of estimated test statistics. In the case of test functions the envelopes formed by pointwise minima and maxima are plotted and it is evaluated how well the test function estimated from the data falls between the envelopes. Various designs of these methods are developed in Myllymäki et al. [2015].

Here we restrict ourselves to scalar test statistics. Let $h : \mathbf{Y} \mapsto \mathbb{R}$, be a test statistic and $\mu_1, \ldots, \mu_m$ be a random sample from model with estimated parameters of a statistic h,

$$h_{lower} = \min(h(\mu_1), \ldots, h(\mu_m)) \,, \tag{3.10}$$

$$h_{upper} = \max(h(\mu_1), \ldots, h(\mu_m)) \tag{3.11}$$

and $h(\mu)$ is the observed value from the data. If $h(\mu) > h_{upper}$ or $h(\mu) < h_{lower}$ we reject the hypothesis that the data come from the model. The significance level is unknown since the testing procedure is not independent of the estimation procedure and also because we are doing multiple tests.

The test of the fit of the model II is based on scalar test statistics $n(\mathbf{x})$, $D(\mathbf{x})$ from the model and moreover on $N(\mathbf{x})$, $L(\mathbf{x})$, the total number of intersections, total length of segments, respectively. In Figures 3.7 the results from $n = 20$ simulations are presented for each cell. The bounds (3.10) are plotted by dashed lines and the values of the test statistic from real data lie between the bounds in all cases. For the test statistics $n(\mathbf{x})$, $D(\mathbf{x})$ from the model naturally the fit is better and the line corresponding to data lies almost in the middle of the bounds. We summarize that based on the selected statistics we cannot reject the hypothesis of model compatibility for any cell. To obtain the level of our test is computationally demanding as explained above.

3.5 Discussion

The segment process having a density with respect to the Poisson process and reference directional and/or length distributions presents a new model for segment systems on a bounded set which may posses interactions. A more complex models can be built by using joint direction-length distributions, but in fact model II is of this kind where the directional distribution is uniform. It should be mentioned that generally the reference distribution need not coincide with the observed distribution. The model I is a Gibbs type homogeneous process while model II is an inhomogeneous Poisson process. We suggested the parameter estimator based on Takacz-Fiksel method for both models, the estimating equations were solved numerically using the Nelder-Mead method. First we showed the capabilities of the estimation procedure in simulated segment systems.

Model I was introduced for simulation and demonstration purposes, because of its homogeneity it arised not to be useful for the modeling of real data from hMSCs. Theferore we tried to apply the model II to real data of stress fibres observed by fluorescence imaging and transformed into segment systems. In model II we involve a special statistics $D(\mathbf{x})$ which performs quite well. The negative values of the corresponding parameter γ_0 for cells 34, 49, 59, 60, 92, 127 (cf. Table 3.5) correspond to a uniform distribution of short filaments across the cell or even tendency to cluster around the centre. Positive values in other cases correspond to a typical accumulation more close to the boundary than to the centre of the cell. The beta distribution of the length is also stable in parameters γ_1, γ_2 (with the exception of cell 131). Positive results of the degree-of-fit test in Figure 3.7 do not yet mean that the data completely correspond to the model since the test is conservative. Moreover functional characteristics (like the contact distribution function) could be implemented as descriptors of spatial distribution. Nevertheless all the presented arguments together make the model II interesting and valuable for the underlying biological problem.

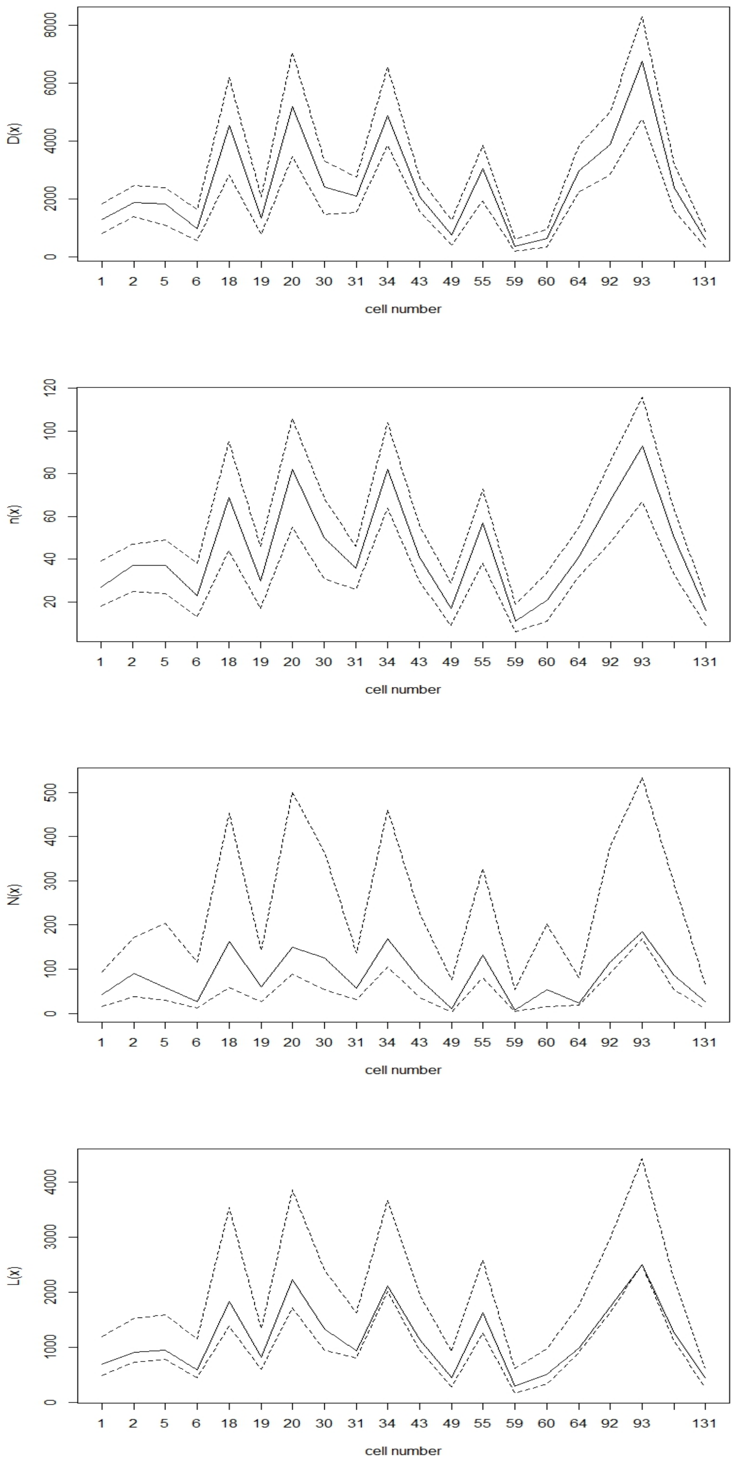

Figure 3.7: The result of model II testing for the statistics $D(\mathbf{x})$, $n(\mathbf{x})$, $N(\mathbf{x})$, $L(\mathbf{x})$ for analysed segment systems corresponding to selected cells from group $G1$. On the horizontal axis there are numbers of cells. The bounds (3.10) are plotted and joined by dashed lines, the values of the test statistics from real data lie between them (joined by a full line). This does not lead to the rejection of model II fit.

4. Particle process asymptotics

In this chapter we use definitions introduced in Chapter 1.4 and further specify them. This chapter is based on Beneš et al. [2019+].

Starting with the seminal paper Schreiber and Yukich [2013], the limit theory for functionals of Gibbs point processes on Euclidean space has recently attracted a lot of attention; see Blaszczyszyn et al. [2019], Torrisi [2017], Xia and Yukich [2015]. In this chapter we derive variance asymptotics and central limit theorems for certain statistics of Gibbs processes of geometrical objects (particles). One possible approach is to extend asymptotic results to Gibbs marked point processes, see e.g. Mase [2000]. In applications marks describe the geometric properties of particles. The marks may be scalar, vectorial or particles themselves. In the literature we find remarks [Dereudre et al., 2012, Remark 3.7] telling that results for (Gibbs) point processes on $\mathbb{R}^d$ generalize easily to marked point processes. However, depending on the circumstances the details of such an extension requires additional effort. Another approach is to parametrize some particle attributes and to deal with the point processes on the parametric space, see Chapter 2 of this thesis, Večeřa and Beneš [2016], Večeřa [2016], Večeřa and Beneš [2017] for an application of the method of moments to a specific Gibbs model of this type. In this chapter we work directly with particle processes, defined as point processes on the space of compact sets equipped with the Hausdorff distance as in Schneider and Weil [2008].

We study a stationary Gibbs particle process μ on $\mathbb{R}^d$ defined in terms of a non-negative potential and an activity parameter, assuming that the size of the particles is deterministically bounded. The background was presented in Sections 1.4-1.6.

Let W_n denote a centered cube of volume $n \in \mathbb{N}$. We are interested in the asymptotic behavior of statistics of the form

$$F_n := \frac{1}{k!} \int f(K_1, \dots, K_k) \mu_n^{(k)}(\mathrm{d}(K_1, \dots, K_k)), \quad n \in \mathbb{N},$$

where f is a symmetric and measurable function of $k \in \mathbb{N}$ particles and $\mu_n^{(k)}$ is the restriction of the k-th factorial measure of μ to $(W_n)^k$. For small activity parameters (and under some additional technical assumptions) we prove a central limit theorem (CLT) for the standardized sequence $\{F_n\}_{n \in \mathbb{N}}$. Our main technical tools are some methods from Blaszczyszyn et al. [2019] combined with a new decorrelation property.

In Section 4.1 we mention the fast decay of correlations provided that the activity is below the percolation threshold of the associated Boolean model. This result is of some independent interest. It not only strengthens and generalizes the results in Schreiber and Yukich [2013], but also holds for a wider range of the activity parameter. The main technical tool is a disagreement coupling [Hofer-Temmel and Houdebert] of two Gibbs processes with a dominating Poisson particle process.

In Section 4.2 we derive the CLT through the factorization of weighted mixed moments analogously to the procedure used for point processes on $\mathbb{R}^d$ in Blaszczyszyn et al. [2019]. Our results are complemented by mean and variance asymptotics of F_n.

4.1 Percolation

Define a symmetric geometric relation on $\mathcal{C}^{(d)}$ by setting $K \sim L$, if and only if, $K \cap L \neq \emptyset$. For $\mathbf{x} \in \mathbf{Y}$, this defines a combinatorial graph $(\mathrm{supp}(\mathbf{x}), \sim)$. For $\{K, L\} \subseteq \mathcal{C}^{(d)}$, we say that $\mathbf{x}$ connects K and L, if there exists a finite path between K and L in the graph on $\mathbf{x} + \delta_K + \delta_L$. For $\mathcal{R}, \mathcal{S} \in \mathcal{B}(\mathcal{C}^{(d)})$, we say that $\mathbf{x}$ connects $\mathcal{R}$ and $\mathcal{S}$, if there exist $K \in \mathcal{R}$ and $L \in \mathcal{S}$ such that $\mathbf{x}$ connects K and L. We write $\mathcal{R} \overset{\mathbf{x}}{\leftrightarrow} \mathcal{S}$ for this.

We say that $\mathbf{x}$ percolates, if its graph contains an infinite connected component. There is a critical percolation intensity $\alpha_c(d) := \alpha_c(d, \mathbb{Q}) \in [0, \infty]$ for percolation of Poisson particle process $\eta_{\alpha\Theta}$.

Lemma 11. *For $0 \leq \alpha < \alpha_c(d)$ and $\mathcal{R}, \mathcal{S} \in \mathcal{B}_b(\mathcal{C}^{(d)})$ with $\mathcal{R} \subseteq \mathcal{S}$, there exist a monotone decreasing $\varpi_1 : [0, \infty) \mapsto [0, \infty)$ and $\varpi_2 \in (0, \infty)$ such that*

$$\mathbb{P}(\mathcal{R} \overset{\eta_{\alpha\Theta}}{\longleftrightarrow} \mathcal{S}^c) \leq \varpi_1(\mathrm{diam}(\mathcal{R})) \exp\left(-\varpi_2 \Delta(\mathcal{R}, \mathcal{S}^c)\right) . \tag{4.1}$$

Monotonicity in the particle shapes allows to control the percolation threshold. In the special case of $\mathbb{Q} = \delta_{B(\mathbf{0},\upsilon)}$, the measure Θ becomes

$$\Theta_\upsilon := \int \mathbf{1}[B(x, \upsilon) \in \cdot] \mathrm{d}x \,.$$

Assumption (1.20) implies that Θ-a.e. $\mathbf{x}$ fulfils

$$\bigcup_{K \in \mathbf{x}} K \subseteq \bigcup_{K \in \mathbf{x}} B(\zeta(K), \upsilon) \,.$$

Hence, we can couple $\eta_{\alpha\Theta}$ and $\eta_{\alpha\Theta_\upsilon}$ such that

$$\mathbb{P}\left(\bigcup_{K \in \eta_{\alpha\Theta}} K \subseteq \bigcup_{L \in \eta_{\alpha\Theta_\upsilon}} L\right) = 1 \,.$$

A well known lower bound [Meester and Roy, 1996, Section 3.10] is

$$\alpha_c(d, \mathbb{Q}) \geq \alpha_c(d, \delta_{B(\mathbf{0},\upsilon)}) \geq \frac{1}{B_d^{vol} \upsilon^d} \,, \tag{4.2}$$

where B_d^{vol} is the volume of the d-dimensional unit sphere.

In the subcritical percolation regime of $\eta_{\tau\Theta}$, the finiteness of the percolation clusters guarantees uniqueness of the Gibbs process μ.

Theorem 12. *If $\tau < \alpha_c(d)$, then the distribution of μ is uniquely determined.*

Proof. The proof generalises straightforward from the proof of [Hofer-Temmel and Houdebert, Theorem 3.2], with the only change being that the interaction range and particle size are here two separate parameters. Because of the deterministic bound υ from (1.20) on the particle size and the finiteness of the interaction range, the arguments remain the same. □

With the following theorem we establish the particle counterpart of fast decay of correlations in [Blaszczyszyn et al., 2019, Definition 2.1] in the subcritical regime.

Theorem 13. *Assume that $\tau < \alpha_c$. Then the Gibbs process μ satisfies*

$$\begin{aligned}&|\rho_{k_1+k_2}(K_1,\dots,K_{k_1+k_2};\mu) - \rho_{k_1}(K_1,\dots,K_{k_1};\mu)\rho_{k_2}(K_{k_1+1},\dots,K_{k_1+k_2};\mu)| \\ &\le \tau^{k_1+k_2}\min\{k_1,k_2\}\beta_1\exp(-\beta_2\Delta(\{K_1,\dots,K_{k_1}\},\{K_{k_1+1},\dots,K_{k_1+k_2}\}))\,,\end{aligned} \tag{4.3}$$

for all $k_1, k_2 \in \mathbb{N}$ and $\kappa^{(k_1+k_2)}$-a.e. $(K_1,\dots,K_{k_1+k_2})$, where β_1 and β_2 are defined in (4.1)*.*

The proof of Theorem 13 is based on several lemmas, see Beneš et al. [2019+].

Combining Theorem 13 with known bounds on the percolation threshold (4.2) implies the following constraint on the activity as sufficient condition for exponential mixing:

$$\tau \le \frac{1}{B_d^{vol}\upsilon^d}\,. \tag{4.4}$$

The equivalent statement in [Schreiber and Yukich, 2013] needs to be translated into our language. Because our potentials have finite range υ_G and since $\Psi \le \tau$, [Schreiber and Yukich, 2013, Proposition 2.1] gives the following constraint on the activity as sufficient condition for exponential concentration:

$$\tau \le \frac{1}{B_d^{vol}(1+\upsilon_G)^d}. \tag{4.5}$$

Our improvement (4.4) comes from a perfect usage of the information about the percolation threshold and the proof of Theorem 13 does not weaken this information.

4.2 Asymptotic properties of admissible functions

In this section we fix an admissible Gibbs process μ as in Definition 30.

For $n \in \mathbb{N}$, let $W_n := \left[-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}\right]^d$ be the centered cube of volume n, $\mathcal{C}_n^d := \zeta^{-1}(W_n)$ and $\mathbf{x}_n := \mathbf{x}_{W_n}$, $\mathbf{x} \in \mathbf{Y}$. Let $\mu_n := \mu_{W_n}$, μ_n^c be the restriction of the Gibbs process to $\mathcal{C}_n^d$, $(\mathcal{C}_n^d)^c$, respectively.

Theorem 14. *Let F_f be an admissible function of order k and $m \in \mathbb{N}$. Then*

$$\sup_{n\in\mathbb{N}}\sup_{1\le j\le m}\sup_{K_1,\dots,K_j\in\mathcal{C}_n^d}\mathbb{E}_{K_1,\dots,K_j}\left[\max\{|\mathbb{T}(K_1,\mu_n)|,1\}^m\right] < \infty\,, \tag{4.6}$$

where $\mathbb{T}$ is given by (1.29) *and the inner supremum is an essential supremum with respect to the j-th reduced factorial moment measure of μ.*

Proof. Complete proof can be found in [Beneš et al., 2019+]. □

Definition 34. *We call $(\mu,\mathbb{T})$ an admissible pair if μ is an admissible Gibbs process with $\tau < \alpha_c(d)$ and the score function $\mathbb{T}$ corresponds to an admissible function F_f, cf. Definition 30 and Definition 31.*

For a given mapping F_f in (1.30), we are interested in the asymptotic properties of $F(\mu_n)$ as $n \to \infty$, using notation from Subsection 1.7.

We state the mean and variance asymptotics of $F_n = F_f(\mu_n)$, $n \in \mathbb{N}$, as well as a central limit theorem. The proofs are presented in Subsection 4.2.1.

Definition 35. *Let $n, k, m_1, \dots, m_k \in \mathbb{N}$ we define the weighted mixed moment*

$$\mathbb{M}^{(m_1,\dots,m_k)}(K_1,\dots,K_k;n) := \rho_k(K_1,\dots,K_k;\mu) \times \int_{\mathbf{Y}} (\mathbb{T}(K_1,\mathbf{x}_n)^{m_1}\cdots\mathbb{T}(K_k,\mathbf{x}_n)^{m_k})\,\mathbb{P}_{K_1,\dots,K_k}(\mathrm{d}\mathbf{x})\,, K_1,\dots,K_k \in \mathcal{C}^d\,. \quad (4.7)$$

For the weighted mixed moments in (4.7) we denote some special cases in the following definition.

Definition 36. *For $K, L \in \mathcal{C}^{(d)}$ we set*

$$\mathcal{M}_{(1)}(K) := \mathbb{E}_K[\mathbb{T}(K,\mu)]\rho_1(K;\mu)\,,$$
$$\mathcal{M}_{(2)}(K,L) := \mathbb{E}_{K,L}[\mathbb{T}(K,\mu)\mathbb{T}(L,\mu)]\rho_2(K,L;\mu)\,,$$
$$\mathcal{M}_{(1,2)}(K) := \mathbb{E}_K[\mathbb{T}^2(K,\mu)]\rho_1(K;\mu)\,.$$

Theorem 15. *Let $(\mu, \mathbb{T})$ be an admissible pair. Then it holds that*

$$\lim_{n\to\infty}\frac{1}{n}\mathbb{E}F_n = \int_{\mathcal{C}_0^d}\mathcal{M}_{(1)}(K)\mathbb{Q}(\mathrm{d}K)\,, \quad (4.8)$$

$$\lim_{n\to\infty}\frac{1}{n}\mathrm{var}F_n = \int_{\mathcal{C}_0^d}\mathcal{M}_{(1,2)}(K)\mathbb{Q}(\mathrm{d}K) + \int_{(\mathcal{C}_0^d)^2}\int_{\mathbb{R}^d}(\mathcal{M}_{(2)}(K,L+x)-\mathcal{M}_{(1)}(K)\mathcal{M}_{(1)}(L))\mathrm{d}x\mathbb{Q}(\mathrm{d}K)\mathbb{Q}(\mathrm{d}L) < \infty\,. \quad (4.9)$$

Next, we write $f(n) = \Omega(g(n))$ when $g(n) = O(f(n))$ as $n \to \infty$.

Theorem 16. *Let $(\mathbb{T}, \mu)$ be an admissible pair. If, for some $c \in (0, \infty)$,*

$$\mathrm{var}F_n = \Omega(n^c)\,, \quad (4.10)$$

then we have the CLT

$$\frac{F_n - \mathbb{E}F_n}{(\mathrm{var}F_n)^{1/2}} \xrightarrow[n\to\infty]{\mathcal{D}} N(0,1)\,. \quad (4.11)$$

Remark 31. Condition (4.10), which does not necessarily follow from (4.9), says that there exists $c > 0$ such that

$$\liminf_{n\to\infty}\frac{\mathrm{var}F_n}{n^c} > 0\,.$$

In [Xia and Yukich, 2015, Theorem 1.1] it is shown that for Gibbs point processes in $\mathbb{R}^d$ the condition (4.10) holds with $c = 1$. This assertion is based on a sufficient non-degeneracy condition which can be rewritten in our context as follows: There exists $q \in (0, \infty)$ and $b_0 = b_0(q) > 0$ such that

$$\inf_{n\in[q,\infty)}\mathbb{E}[\mathrm{var}(F_n|\zeta^{-1}(W_n^c)\cap\mu)] \geq b_0\,. \quad (4.12)$$

The result on CLT can be extended to the so called stabilizing functionals, as defined and investigated in [Blaszczyszyn et al., 2019].

4.2.1 Factorization of weighted mixed moments

In the following all equations and inequalities involving Palm distributions and correlation functions are to be understood in the a.e.-sense with respect to the appropriate factorial moment measures of μ.

Definition 37. *We say that the weighted mixed moments approximately factorize, if there exist constants $\beta_{1,M}, \beta_{2,M} > 0$, such that*

$$\begin{aligned}|\mathbb{M}^{(m_1,\ldots,m_{k_1})}(K_1,\ldots,K_{k_1};n)\mathbb{M}^{(m_{k_1+1},\ldots,m_{k_1+k_2})}(K_{k_1+1},\ldots,K_{k_1+k_2};n) \\ - \mathbb{M}^{(m_1,\ldots,m_{k_1+k_2})}(K_1,\ldots,K_{k_1+k_2};n)| \\ \leq \beta_{1,M}\exp(-\beta_{2,M}\Delta(\{K_1,\ldots,K_{k_1}\},\{K_{k_1+1},\ldots,K_{k_1+k_2}\}))\,, \end{aligned} \tag{4.13}$$

for all $n, k_1, k_2, m_1, \ldots, m_{k_1+k_2} \in \mathbb{N}$ and for all $K_1, \ldots, K_{k_1+k_2} \in \zeta^{-1}(W_n)$, where $M = \sum_{i=1}^{k_1+k_2} m_i$.

In the following we use Theorem 13 to show that (4.13) holds in our context. The method from [Blaszczyszyn et al., 2019] is used and transformed from point processes on $\mathbb{R}^d$ to particle processes.

The space $(\mathcal{C}^{(d)}, \Delta_H)$ is a complete and separable metric space. By [Dudley, 1989, Theorem 13.1.1], the spaces $(\mathcal{C}^{(d)}, \mathcal{B}(\mathcal{C}^{(d)}))$ and $\mathbb{R}$ equipped with the Borel σ-algebra are Borel isomorph. That is, there exists a measurable bijection from $\mathcal{C}^{(d)}$ to $\mathbb{R}$ with measurable inverse. We use this bijection to pull back the total order from $\mathbb{R}$ to $\mathcal{C}^{(d)}$ and denote it by $\prec$. Hence, intervals with respect to $\prec$ are in $\mathcal{B}(\mathcal{C}^{(d)})$.

For $\mathbf{x} \in \mathbf{Y}$ and $K \in \mathcal{C}^{(d)}$, define the measure

$$\mathbf{x}_{|K}(\cdot) := \mathbf{x}(\cdot \cap \{L \mid L \prec K\})\,.$$

Let o be the zero-measure, i.e., $o(\mathcal{R}) = 0$, for all $\mathcal{R} \in \mathcal{B}(\mathcal{C}^{(d)})$. We define a difference operator for a measurable function $F : \mathbf{Y} \to \mathbb{R}$, $l \in \mathbb{N} \cup \{0\}$ and $K_1, \ldots, K_l \in \mathcal{C}^{(d)}$ by

$$D^l_{K_1,\ldots,K_l}F(\mathbf{x}) := \begin{cases} \sum_{J\subseteq[l]}(-1)^{l-|J|}F(\mathbf{x}_{|K_*} + \sum_{j\in J}\delta_{K_j}) & \text{if } l > 0, \\ F(o) & \text{if } l = 0, \end{cases} \tag{4.14}$$

where the minimum $K_* := \min\{K_1, \ldots, K_l\}$ is taken with respect to the order $\prec$. We say that F is $\prec$-continuous at ∞ if, for all $\mathbf{x} \in \mathbf{Y}$, we have $\lim_{K\uparrow\mathbb{R}^d} F(\mathbf{x}_{|K}) = F(\mathbf{x})$.

We use the following factorial moment expansion (FME) proved in [Błaszczyszyn et al., 1997, Theorem 3.1] on a general Polish space. For stronger results in the special case of a Poisson particle process we refer to [Last, 2014] and [Last and Penrose, 2017, Chapter 19].

Theorem 17. *Let $F : \mathbf{Y} \to \mathbb{R}$ be $\prec$-continuous at ∞. Assume that, for all $l \in \mathbb{N}$,*

$$\int_{(\mathcal{C}^{(d)})^l} \mathbb{E}^!_{K_1,\ldots,K_l}[D^l_{K_1,\ldots,K_l}F(\mu)]\rho_l(K_1,\ldots,K_l;\mu)\Theta^l(\mathrm{d}(K_1,\ldots,K_l)) < \infty \tag{4.15}$$

and

$$\lim_{l\to\infty}\frac{1}{l!}\int_{(\mathcal{C}^{(d)})^l}\mathbb{E}^!_{K_1,\dots,K_l}[D^l_{K_1,\dots,K_l}F(\mu)]\rho_l(K_1,\dots,K_l;\mu)\Theta^l(\mathrm{d}(K_1,\dots,K_l))=0. \tag{4.16}$$

Then, $\mathbb{E}[F(\mu)]$ *has the FME*

$$\mathbb{E}[F(\mu)]=F(o)+\sum_{l=1}^{\infty}\frac{1}{l!}\int_{(\mathcal{C}^{(d)})^l}[D^l_{K_1,\dots,K_l}F(o)]\rho_l(K_1,\dots,K_l;\mu)\Theta^l(\mathrm{d}(K_1,\dots,K_l))\,. \tag{4.17}$$

Definition 38. *For an admissible pair* $(\mathbb{T},\mu)$*,* $K_1,\dots,K_l\in\mathcal{C}^{(d)}$ *and* $\mathbf{x}\in\mathbf{Y}$*, set*

$$F_{m_1,\dots,m_l}(K_1,\dots,K_l;\mathbf{x}):=\prod_{i=1}^{l}\mathbb{T}(K_i,\mathbf{x})^{m_i}\,, \tag{4.18}$$

$$F^!_{m_1,\dots,m_l}(K_1,\dots,K_l;\mathbf{x}):=\prod_{i=1}^{l}\mathbb{T}(K_i,\mathbf{x}+\sum_{j=1}^{l}\delta_{K_j})^{m_i}\,, \tag{4.19}$$

with $m_1,\dots,m_l\ge 1$.

Remark 32. It holds that $\mathbb{E}_{K_1,\dots,K_l}[F(\mu_n)]=\mathbb{E}^!_{K_1,\dots,K_l}[F^!(\mu_n)]$.

Definition 39. *For* $l,k\in\mathbb{N}$*,* $K_1,\dots,K_l,L_1,\dots,L_k\in\mathcal{C}^{(d)}$ *we denote by*

$$\rho_l^{(K_1,\dots,K_l)}(L_1,\dots,L_k;\mu)$$

the l*-th correlation function of* $\mathbb{P}^!_{K_1,\dots,K_l}$.

Further for $l,k\in\mathbb{N}$*,* $K_1,\dots,K_l\in\mathcal{C}^{(d)}$ *we denote*

$$(\mathbb{P}^!_{K_1,\dots,K_l})^!_{L_1,\dots,L_k}\,,L_1,\dots,L_k\in\mathcal{C}^{(d)}\,,$$

the reduced Palm distributions of $\mathbb{P}^!_{K_1,\dots,K_l}$.

Remark 33. It is easy to show that

$$\rho_l(K_1,\dots,K_l;\mu)\rho_k^{(K_1,\dots,K_l)}(L_1,\dots,L_k;\mu)=\rho_{l+k}(K_1,\dots,K_l,L_1,\dots,L_k;\mu)\,, \tag{4.20}$$

and

$$(\mathbb{P}^!_{K_1,\dots,K_l})^!_{L_1,\dots,L_k}=\mathbb{P}^!_{K_1,\dots,K_l,L_1,\dots,L_k}\,. \tag{4.21}$$

Lemma 18. *For distinct* $K_1,\dots,K_p\in\mathcal{C}^{(d)}$*, non-negative integers* $m_1,\dots,m_p,n$ *and* $M_p:=\sum_{i=1}^{p}m_i$*, the functional* $F^!$ *admits the FME*

$$\begin{aligned}\mathbb{E}^!_{K_1,\dots,K_p}[F^!_{m_1,\dots,m_p}(K_1,\dots,K_p;\mu_n)]&=F^!_{m_1,\dots,m_p}(K_1,\dots,K_p;o)\\&+\sum_{l=1}^{M_p(k-1)}\frac{1}{l!}\int_{(\mathcal{C}^{(d)})^l}D^l_{L_1,\dots,L_l}F^!_{m_1,\dots,m_p}(K_1,\dots,K_p,o)\\&\qquad\times\rho_l^{(K_1,\dots,K_p)}(L_1,\dots,L_l;\mu)\Theta^l(\mathrm{d}(L_1,\dots,L_l))\,.\end{aligned}$$

Proof. We abbreviate $F_{m_1,\dots,m_p}(K_1,\dots,K_p;\mu)$ by $F(K_1,\dots,K_p;\mu)$. The radius bound r from (1.31) for the function f implies that $F^!$ is $\prec$-continuous at ∞. In [Blaszczyszyn et al., 2019, Lemma 5.1] it is shown that $F^!$ is the sum of admissible statistics of orders not larger than $M_p(k-1)$, where k is the order of admissible statistics. Thus, for $l \in (M_p(k-1),\infty)$ and all $L_1,\dots,L_l \in \mathcal{C}^{(d)}$, we have

$$D^l_{L_1,\dots,L_l}F^!(K_1,\dots,K_p;\mathbf{x}) = 0\,. \tag{4.22}$$

This implies that (4.15), for $l \in (M_p(k-1),\infty)$, and (4.16) are satisfied for $F^!$ from (4.19). We need to verify (4.15), for $l \in [1, M_p(k-1)]$. For $L_1,\dots,L_l \in \mathcal{C}^{(d)}$, $\mathbf{x} \in \mathbf{Y}$ and $J \subseteq [l]$, set

$$\mathbf{x}_J := \mathbf{x}\,|_{L_*} + \sum_{j\in J}\delta_{L_j}\,,$$

where the minimum $L_* := \min\{L_1,\dots,L_l\}$ is taken with respect to the order $\prec$. The difference operator $D^l_{L_1,\dots,L_l}$ vanishes like in (4.22) as soon as $L_q \notin \bigcup_{i=1}^p B(K_i,2r)$ for some $q \in [l]$. To prove this, expand (4.14) to obtain

$$\begin{aligned} D^l_{L_1,\dots,L_l}F^!(K_1,\dots,K_p;\mathbf{x}) = &\sum_{J\subseteq[l],q\notin J}(-1)^{l-|J|}F^!(K_1,\dots,K_p;\mathbf{x}_J)\\ &+\sum_{J\subseteq[l],q\notin J}(-1)^{l-|J|-1}F^!(K_1,\dots,K_p;\mathbf{x}_{J\cup\{q\}}) = 0\,, \end{aligned}$$

since, for fixed $J \subseteq [l]$ and $q \notin J$, $F^!(K_1,\dots,K_p;\mathbf{x}_J) = F^!(K_1,\dots,K_p;\mathbf{x}_{J\cup\{q\}})$. Then, we have

$$\begin{aligned} F^!(K_1,\dots,K_p;\mathbf{x}_J) &\le \prod_{i=1}^p \|f\|_\infty^{m_i}\left(\mathbf{x}\left(\bigcup_{i=1}^p B(K_i,2r)\right)+|J|+p\right)^{m_i(k-1)}\\ &\le \|f\|_\infty^{M_p}\left(\mathbf{x}\left(\bigcup_{i=1}^p B(K_i,2r)\right)+|J|+p\right)^{M_p(k-1)}. \end{aligned}$$

Using this for difference operator we get

$$\begin{aligned} |D^l_{L_1,\dots,L_l}F^!(K_1,\dots,K_p;\mathbf{x})| &\le \|f\|_\infty^{M_p}\sum_{J\subseteq[l]}\left(\mathbf{x}\left(\bigcup_{i=1}^p B(K_i,2r)\right)+|J|+p\right)^{M_p(k-1)}\\ &\le \|f\|_\infty^{M_p}2^l\left(\mathbf{x}\left(\bigcup_{i=1}^p B(K_i,2r)\right)+l+p\right)^{M_p(k-1)}. \end{aligned} \tag{4.23}$$

Using (4.21), the defining equation (1.25) and (4.23) results in

$$\begin{aligned} &\frac{1}{l!}\int_{(\mathcal{C}^{(d)})^l}(\mathbb{E}^!_{K_1,\dots,K_p})^!_{L_1,\dots,L_l}[|D^l_{L_1,\dots,L_l}F^!(K_1,\dots,K_p;\mu_n)|]\\ &\times\rho_l^{(K_1,\dots,K_p)}(L_1,\dots,L_l;\mu)\Theta^l(\mathrm{d}(L_1,\dots,L_l))\\ &=\frac{1}{l!}\int_{(\mathcal{C}^{(d)})^l}\mathbb{E}^!_{K_1,\dots,K_p,L_1,\dots,L_l}[|D^l_{L_1,\dots,L_l}F^!(K_1,\dots,K_p;\mu_n)|]\\ &\times\rho_l^{(K_1,\dots,K_p)}(L_1,\dots,L_l;\mu)\Theta^l(\mathrm{d}(L_1,\dots,L_l))\\ \le&\|f\|_\infty^{M_p}2^l\mathbb{E}_{K_1,\dots,K_p}\left[\mu\left(\bigcup_{i=1}^p B(K_i,r)\right)^l\left(\mu\left(\bigcup_{i=1}^p B(K_i,r)\right)+l+p\right)^{M_p(k-1)}\right]. \end{aligned}$$

Since μ has all moments under the Palm measure the finiteness of the last term and hence the validity of the condition for $l \in [1, M_p(k-1)]$ follows. This justifies the FME expansion. $\square$

Theorem 19. *Let $(\mathbb{T}, \mu)$ be an admissible pair. Then the weighted moments approximately factorize.*

Proof. Let $k_1, k_2, m_1, \ldots, m_{k_1+k_2} \in \mathbb{N}$ be fixed. Let $t := \max(4\upsilon + \upsilon_G, r)$, with υ_G being the finite interaction range of the admissible particle process as outlined in Definition 30 and taking into account the particle size from (1.20), as well as (1.31).

Given $n \in \mathbb{N}$, $K_1, \ldots, K_{k_1+k_2} \in \zeta^{-1}(W_n)$, we set

$$R := \Delta(\{K_1, \ldots, K_{k_1}\}, \{K_{k_1+1}, \ldots, K_{k_1+k_2}\}) .$$

Without loss of generality we assume that $R \in (8t, \infty)$. Put $M := \sum_{i=1}^{k_1+k_2} m_i$, $M_{k_1} := \sum_{i=1}^{k_1} m_i$ and $M_{k_2} := \sum_{i=k_1+1}^{k_1+k_2} m_i$. Then, using Lemma 18, (4.22) and (4.20) we obtain

$$\begin{aligned}
&\mathbb{M}^{(m_1,\ldots,m_{k_1+k_2})}(K_1, \ldots, K_{k_1+k_2}; n) \\
&= \mathbb{E}^!_{K_1,\ldots,K_{k_1+k_2}}[F^!(K_1, \ldots, K_{k_1+k_2}; \mu_n)]\rho_{k_1+k_2}(K_1, \ldots, K_{k_1+k_2}; \mu) \\
&= \sum_{l=0}^{M(k-1)} \frac{1}{l!} \int_{(\mathcal{C}^d_n)^l} D^l_{L_1,\ldots,L_l} F^!(o) \\
&\times \rho_{l+k_1+k_2}(K_1, \ldots, K_{k_1+k_2}, L_1, \ldots, L_l; \mu)\Theta^l(\mathrm{d}(L_1, \ldots, L_l)) \\
&= \sum_{l=0}^{M(k-1)} \frac{1}{l!} \int_{(\bigcup_{i=1}^{k_1+k_2} B(K_i,2t))^l} D^l_{L_1,\ldots,L_l} F^!(o) \\
&\times \rho_{l+k_1+k_2}(K_1, \ldots, K_{k_1+k_2}, L_1, \ldots, L_l; \mu)\Theta^l(\mathrm{d}(L_1, \ldots, L_l)) .
\end{aligned}$$

Let $\varpi_{p_1,p_2} := (\bigcup_{i=1}^{k_1} B(K_i, 2t))^{p_1} \times (\bigcup_{i=1}^{k_2} B(K_{k_1+i}, 2t))^{p_2}$. Then, using the FME from Lemma 18,

$$\begin{aligned}
&\mathbb{M}^{(m_1,\ldots,m_{k_1+k_2})}(K_1, \ldots, K_{k_1+k_2}; n) \\
&\quad= \sum_{l=0}^{M(k-1)} \frac{1}{l!} \sum_{j=0}^{l} \frac{l!}{j!(l-j)!} \int_{\varpi_{j,l-j}} D^l_{L_1,\ldots,L_l} F^!(K_1, \ldots, K_{k_1+k_2}; o) \\
&\quad\times \rho_{l+k_1+k_2}(K_1, \ldots, K_{k_1+k_2}, L_1, \ldots, L_l)\Theta^l(\mathrm{d}(L_1, \ldots, L_l)) \\
&\quad= \sum_{l=0}^{M(k-1)} \sum_{j=0}^{l} \frac{1}{j!(l-j)!} \int_{\varpi_{j,l-j}} \sum_{J\subseteq[l]} (-1)^{l-|J|} F^!(K_1, \ldots, K_{k_1+k_2}; \sum_{j\in J} \delta_{L_j}) \\
&\quad\times \rho_{l+k_1+k_2}(K_1, \ldots, K_{k_1+k_2}, L_1, \ldots, L_l)\Theta^l(\mathrm{d}(L_1, \ldots, L_l)) .
\end{aligned}$$

To compare the (k_1+k_2)-th mixed moment with the product of the k_1-th and k_2-th mixed moments we use factorization that holds for $L_1, \ldots, L_j \in \bigcup_{i=1}^{k_1} B(K_i, 2t)$ and $L_{j+1}, \ldots, L_l \in \bigcup_{i=1}^{k_2} B(K_{k_1+i}, 2t)$. If

$$N_1 \in \bigcup_{i=1}^{k_1} B(K_i, 2t) \, , N_2 \in \bigcup_{i=1}^{k_2} B(K_{k_1+i}, 2t) \, ,$$

then $N_1 \cap N_2 = \emptyset$. Hence

$$F^!(K_1, \ldots, K_{k_1+k_2}; \sum_{i=1}^{l} \delta_{L_i}) = F^!(K_1, \ldots, K_{k_1}; \sum_{i=1}^{j} \delta_{L_i}) F^!(K_{k_1+1}, \ldots, K_{k_1+k_2}; \sum_{i=j+1}^{l} \delta_{L_i}) . \quad (4.24)$$

Using (4.24) and the similar steps as in the case of the (k_1+k_2)-th mixed moment we work with the product of k_1-th and k_2-th mixed moments.

$$\begin{aligned}
&\mathbb{M}^{(m_1,\ldots,m_{k_1})}(K_1, \ldots, K_{k_1}; n)\mathbb{M}^{(m_{k_1+1},\ldots,m_{k_1+k_2})}(K_{k_1+1}, \ldots, K_{k_1+k_2}; n) \\
&= \mathbb{E}^!_{K_1,\ldots,K_{k_1}}[F^!(K_1, \ldots, K_{k_1}, \mu_n)] \\
&\quad \times \mathbb{E}^!_{K_{k_1+1},\ldots,K_{k_1+k_2}}[F^!(K_{k_1+1}, \ldots, K_{k_1+k_2}, \mu_n)] \\
&\quad \times \rho_{k_1}(K_1, \ldots, K_{k_1}; \mu)\rho_{k_2}(K_{k_1+1}, \ldots, K_{k_1+k_2}; \mu) \\
&= \sum_{l_1,l_2=0}^{\infty} \int_{\varpi_{l_1,l_2}} D^{l_1}_{L_1,\ldots,L_{l_1}} F^!(K_1, \ldots, K_{k_1}; o) \\
&\quad \times D^{l_2}_{N_1,\ldots,N_{l_2}} F^!(K_{k_1+1}, \ldots, K_{k_1+k_2}; o) \\
&\quad \times \rho_{l_1+k_1}(K_1, \ldots, K_{k_1}, L_1, \ldots, L_{l_1}; \mu) \\
&\quad \times \rho_{l_2+k_2}(K_{k_1+1}, \ldots, K_{k_1+k_2}, N_1, \ldots, N_{l_2}; \mu) \\
&\quad \times \Theta^{l_1}(\mathrm{d}(L_1, \ldots, L_{l_1}))\Theta^{l_2}(\mathrm{d}(N_1, \ldots, N_{l_2})) \\
&= \sum_{l_1,l_2=0}^{\infty} \frac{1}{l_1! l_2!} \int_{\varpi_{l_1,l_2}} \sum_{J_1 \subseteq [l_1], J_2 \subseteq [l_2]} (-1)^{l_1+l_2-|J_1|-|J_2|} \\
&\quad \times F^!(K_1, \ldots, K_{k_1}; \sum_{i \in J_1} \delta_{L_i}) F^!(K_{k_1+1}, \ldots, K_{k_1+k_2}; \sum_{i \in J_2} \delta_{N_i}) \\
&\quad \times \rho_{l_1+k_1}(K_1, \ldots, K_{k_1}, L_1, \ldots, L_{l_1}; \mu) \\
&\quad \times \rho_{l_2+k_2}(K_{k_1+1}, \ldots, K_{k_1+k_2}, N_1, \ldots, N_{l_2}; \mu) \\
&\quad \times \Theta^{l_1}(\mathrm{d}(L_1, \ldots, L_{l_1}))\Theta^{l_2}(\mathrm{d}(N_1, \ldots, N_{l_2})) \\
&= \sum_{l=0}^{M(k-1)} \sum_{j=0}^{l} \frac{1}{j!(l-j)!} \int_{\varpi_{j,l-j}} \sum_{J_1 \subseteq [j], J_2 \subseteq [l] \setminus [j]} (-1)^{l-|J_1|-|J_2|} \\
&\quad \times F^!(K_1, \ldots, K_{k_1}; \sum_{i \in J_1} \delta_{L_i}) F^!(K_{k_1+1}, \ldots, K_{k_1+k_2}; \sum_{i \in J_2} \delta_{L_i}) \\
&\quad \times \rho_{j+k_1}(K_1, \ldots, K_{k_1}, L_1, \ldots, L_j; \mu) \\
&\quad \times \rho_{l-j+k_2}(K_{k_1+1}, \ldots, K_{k_1+k_2}, L_{j+1}, \ldots, L_l; \mu)\Theta^l(\mathrm{d}(L_1, \ldots, L_l)) \\
&= \sum_{l=0}^{M(k-1)} \sum_{j=0}^{l} \frac{1}{j!(l-j)!} \int_{\varpi_{j,l-j}} \sum_{J \subseteq [l]} (-1)^{l-|J|} \\
&\quad \times F^!(K_1, \ldots, K_{k_1+k_2}; \sum_{i \in J} \delta_{L_i}) \rho_{j+k_1}(K_1, \ldots, K_{k_1}, L_1, \ldots, L_j) \\
&\quad \times \rho_{l-j+k_2}(K_{k_1+1}, \ldots, K_{k_1+k_2}, L_{j+1}, \ldots, L_l; \mu)\Theta^l(\mathrm{d}(L_1, \ldots, L_l)) .
\end{aligned}$$

Altogether we have, using β_1, β_2 from (4.3), that

$$
\begin{aligned}
&|\mathbb{M}^{(m_1,\dots,m_{k_1+k_2})}(K_1,\dots,K_{k_1+k_2};n)\\
&-\mathbb{M}^{(m_1,\dots,m_{k_1})}(K_1,\dots,K_{k_1};n)\mathbb{M}^{(m_{k_1+1},\dots,m_{k_1+k_2})}(K_{k_1+1},\dots,K_{k_1+k_2};n)|\\
&\le \sum_{l=0}^{M(k-1)}\sum_{j=0}^{l}\sum_{J\subseteq[l]}\frac{(-1)^{l-|J|}}{j!(l-j)!}\int_{\varpi_{j,l-j}} F^{!}(K_1,\dots,K_{k_1+k_2};\sum_{i\in J}\delta_{L_i})\\
&\times|\rho_{j+k_1}(K_1,\dots,K_{k_1},L_1,\dots,L_j)\rho_{l-j+k_2}(K_{k_1+1},\dots,K_{k_1+k_2},L_{j+1},\dots,L_l)\\
&-\rho_{l+k_1+k_2}(K_1,\dots,K_{k_1+k_2},L_1,\dots,L_l)|\Theta^l(\mathrm{d}(L_1,\dots,L_l))\\
\le&\sum_{l=0}^{M(k-1)}\sum_{j=0}^{l}\sum_{J\subseteq[l]}\frac{(-1)^{l-|J|}}{j!(l-j)!}\int_{\varpi_{j,l-j}} F^{!}(K_1,\dots,K_{k_1+k_2};\sum_{i\in J}\delta_{L_i})\\
&\times\tau^{l+k_1+k_2}\min\{j+k_1,l-j+k_2\}\beta_1\\
&\times\exp(-\beta_2\Delta(\{K_1,\dots,K_{k_1},L_1,\dots,L_j\},\{K_{k_1+1},\dots,K_{k_1+k_2},L_{j+1},\dots,L_l\})\\
&\times\Theta^l(\mathrm{d}(L_1,\dots,L_l))\,.
\end{aligned}
$$

Using

$$\mathbb{T}(K,\mathbf{x})\mathbf{1}[\mathbf{x}(\mathcal{C}^{(d)})=p]\le\frac{p^{k-1}}{k}\|f\|_\infty\,,$$

we have

$$\sum_{J\subseteq[l]}|F^{!}(K_1,\dots,K_{k_1+k_2};\sum_{i\in J}\delta_{L_i})|\le 2^l\left(\|f\|_\infty\frac{|k_1+k_2+l|^{k-1}}{k}\right)^M\,.$$

The difference of weighted mixed moments is finally bounded by

$$
\begin{aligned}
&|\mathbb{M}^{(m_1,\dots,m_{k_1+k_2})}(K_1,\dots,K_{k_1+k_2};n)\\
&-\mathbb{M}^{(m_1,\dots,m_{k_1})}(K_1,\dots,K_{k_1};n)\mathbb{M}^{(m_{k_1+1},\dots,m_{k_1+k_2})}(K_{k_1+1},\dots,K_{k_1+k_2};n)|\\
&\le\sum_{l=0}^{M(k-1)}\sum_{j=0}^{l}\left(\|f\|_\infty\frac{|k_1+k_2+l|^{k-1}}{k}\right)^M\frac{\exp(-8t)2^l(-1)^{l-|J|}}{j!(l-j)!}\Theta(B(\{\mathbf{0}\},3t))^l\\
&\times\tau^{l+k_1+k_2}\min\{j+k_1,l-j+k_2\}\\
&\times\beta_1\exp(-\beta_2\Delta(\{K_1,\dots,K_{k_1}\},\{K_{k_1+1},\dots,K_{k_1+k_2}\})\,.
\end{aligned}
$$

As $\min\{j+k_1,l-j+k_2\}\le l+k_1+k_2\le kM$, we obtain the desired constants for approximate factorization depending only on M and the attributes of the admissible pair. □

4.2.2 Proofs of limit theorems

In this subsection we prove mean and variance asymptotics of $F_n=F_f(\mu_n)$, $n\in\mathbb{N}$, as well as a central limit theorem. For $K,L\in\mathcal{C}^d$, $n\in\mathbb{N}$, $x\in\mathbb{R}^d$ we abbreviate

$$\mathcal{M}_{(1)}(K;n,x)=\mathbb{E}_K[\mathbb{T}(K,\mu_n^x)]\rho_1(K;\mu)\,,\tag{4.25}$$

$$\mathcal{M}_{(2)}(K,L;n,x)=\mathbb{E}_{K,L}[\mathbb{T}(K,\mu_n^x)\mathbb{T}(L,\mu_n^x)]\rho_2(K,L;\mu)\,,\tag{4.26}$$

where $\mu_n^x:=\mu\cap(W_n-n^{\frac{1}{d}}x)$.

Proof of Theorem 15. From (1.25), (1.23) and (1.17), we have

$$\mathbb{E}F_n = \int_{\mathcal{C}_n^d} \mathbb{E}_K \mathbb{T}(K,\mu_n)\rho_1(K;\mu)\Theta(\mathrm{d}K)$$
$$= \int_{\mathcal{C}_0^d}\int_{W_n} \mathbb{E}_{K+x}[\mathbb{T}(K+x,\mu_n)]\rho_1(K+x;\mu)\mathrm{d}x\mathbb{Q}(\mathrm{d}K)\,.$$

The stationarity of μ and translation invariance of $\mathbb{T}$ imply that

$$\rho_1(K+x;\mu) = \rho_1(K;\mu) \text{ and } \mathbb{T}(K+x,\mu) = \mathbb{T}(K,\mu)\,, K\in\mathcal{C}^{(d)}\,, x\in\mathbb{R}^d\,.$$

Thus,

$$\int_{\mathcal{C}_0^d}\int_{W_n} \mathbb{E}_{K+x}[\mathbb{T}(K+x,\mu)]\rho_1(K+x;\mu)\mathrm{d}x\mathbb{Q}(\mathrm{d}K) = n\int_{\mathcal{C}^{(d)}} \mathcal{M}_{(1)}(K)\mathbb{Q}(\mathrm{d}K)\,.$$

To prove (4.8) it remains to show that

$$\frac{1}{n}\int_{\mathcal{C}_0^d}\int_{W_n} |\mathbb{E}_{K+x}[\mathbb{T}(K+x,\mu_n)] - \mathbb{E}_{K+x}[\mathbb{T}(K+x,\mu)]|\mathrm{d}x\rho_1(K;\mu)\mathbb{Q}(\mathrm{d}K)$$

tends to zero as $n\to\infty$. The function ρ_1 is bounded. For $K\in\mathcal{C}_0^d$, $K\subseteq B(\mathbf{0},\upsilon)$ fixed and $x\in W_n$, we use (1.31) to obtain $\mathbb{T}(K+x,\mu_n) = \mathbb{T}(K+x,\mu)$ whenever $\Delta(x,\partial W_n)\le 2\upsilon$ for the distance from x to the boundary of W_n holds. The 1-moment condition (4.6) implies the existence of some $0<t<\infty$ such that $|\mathbb{T}(K+x,\mu_n)|\le t$ and $|\mathbb{T}(K+x,\mu)|\le t$, uniformly in $n\in\mathbb{N}$ and $K+x\in\mathcal{C}_n^d$. Since

$$\lim_{n\to\infty}\frac{1}{n}\lambda^d\{x\in W_n \mid \Delta(x,\partial W_n)\le 2\upsilon\} = 0\,,$$

the first assertion of the theorem is proved.
For the second moment we obtain as above

$$\mathbb{E}F_n^2 = J_1 + J_2 = \int_{\mathcal{C}_0^d}\int_{W_n} \mathbb{E}_{K+x_1}\mathbb{T}^2(K+x_1,\mu_n)\rho_1(K+x_1;\mu)\mathrm{d}x_1\mathbb{Q}(\mathrm{d}K)$$
$$+\int_{(\mathcal{C}_0^d)^2}\int_{(W_n)^2} \mathbb{E}_{K+x_1,L+x_2}[\mathbb{T}(K+x_1,\mu_n)\mathbb{T}(L+x_2,\mu_n)]$$
$$\times\rho_2(K+x_1,L+x_2;\mu)\mathrm{d}x_1\mathrm{d}x_2\mathbb{Q}(\mathrm{d}K)\mathbb{Q}(\mathrm{d}L)\,.$$

Then

$$\lim_{n\to\infty}\frac{J_1}{n} = \int_{\mathcal{C}^{(d)}} \mathcal{M}_{(1,2)}(K)\mathbb{Q}(\mathrm{d}K)$$

is obtained analogously to the mean value asymptotics above using the 2-moment condition (4.6).
In the second term J_2 we use the substitutions $\tilde{x}_1 = n^{-\frac{1}{d}}x_1$, $\tilde{x}_2 = x_2 - x_1$, obtaining

$$\frac{J_2}{n} = \int_{(\mathcal{C}_0^d)^2}\int_{W_1}\int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} \rho_2(K+n^{\frac{1}{d}}\tilde{x}_1, L+\tilde{x}_2+n^{\frac{1}{d}}\tilde{x}_1;\mu)$$
$$\times\mathbb{E}_{K+n^{\frac{1}{d}}\tilde{x}_1, L+\tilde{x}_2+n^{\frac{1}{d}}\tilde{x}_1}[\mathbb{T}(K+n^{\frac{1}{d}}\tilde{x}_1,\mu_n)\mathbb{T}(L+\tilde{x}_2+n^{\frac{1}{d}}\tilde{x}_1,\mu_n)]$$
$$\times\mathrm{d}\tilde{x}_2\mathrm{d}\tilde{x}_1\mathbb{Q}(\mathrm{d}K)\mathbb{Q}(\mathrm{d}L)$$
$$= \int_{(\mathcal{C}_0^d)^2}\int_{W_1}\int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} \rho_2(K, L+\tilde{x}_2;\mu)$$
$$\times\mathbb{E}_{K,L+\tilde{x}_2}[\mathbb{T}(K,\mu_n^{\tilde{x}_1})\mathbb{T}(L+\tilde{x}_2,\mu_n^{\tilde{x}_1})]\mathrm{d}\tilde{x}_2\mathrm{d}\tilde{x}_1\mathbb{Q}(\mathrm{d}K)\mathbb{Q}(\mathrm{d}L)\,,$$

Since $\mathrm{var}(F_n) = \mathbb{E}F_n^2 - (\mathbb{E}F_n)^2$, we investigate the expression $\frac{1}{n}(J_2 - (\mathbb{E}F_n)^2)$. It takes the form

$$\int_{(\mathcal{C}_0^d)^2} \int_{W_1} \int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} (\mathcal{M}_{(2)}(K, L+\tilde{x}_2; n, \tilde{x}_1) - \mathcal{M}_{(1)}(K; n, \tilde{x}_1)\mathcal{M}_{(1)}(L+\tilde{x}_2; n, \tilde{x}_1)) \times \mathrm{d}\tilde{x}_2 \mathrm{d}\tilde{x}_1 \mathrm{d}\mathbb{Q}(K)\mathrm{d}\mathbb{Q}(L)\,. \quad (4.27)$$

Splitting the innermost integral in (4.27) into the two terms

$$\int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} (\dots)\mathrm{d}\tilde{x}_2 = \int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} (\dots)\mathbf{1}[|\tilde{x}_2| \le t]\mathrm{d}\tilde{x}_2 + \int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} (\dots)\mathbf{1}[|\tilde{x}_2| > t]\mathrm{d}\tilde{x}_2\,, \quad (4.28)$$

for an arbitrary $t > 0$, we observe that the part of (4.27) corresponding to the first term of (4.28), i.e.

$$\int_{(\mathcal{C}_0^d)^2} \int_{W_1} \int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} (\mathcal{M}_{(2)}(K, L+\tilde{x}_2; n, \tilde{x}_1) - \mathcal{M}_{(1)}(K; n, \tilde{x}_1)\mathcal{M}_{(1)}(L+\tilde{x}_2; n, \tilde{x}_1)) \times \mathbf{1}[|\tilde{x}_2| \le t]\mathrm{d}\tilde{x}_2 \mathrm{d}\tilde{x}_1 \mathrm{d}\mathbb{Q}(K)\mathrm{d}\mathbb{Q}(L)\,, \quad (4.29)$$

converges to

$$\int_{(\mathcal{C}_0^d)^2} \int_{\mathbb{R}^d} (\mathcal{M}_{(2)}(K, L+x) - \mathcal{M}_{(1)}(K)\mathcal{M}_{(1)}(L))\mathrm{d}x\mathbb{Q}(\mathrm{d}K)\mathbb{Q}(\mathrm{d}L),$$

when first $n \to \infty$ and then $t \to \infty$. Using (4.13) absolute value of the second term in (4.28) can be bounded uniformly in n by

$$\beta_{1,2} \int_{|x|>t} \exp\left(-\beta_{2,2}\Delta_H(K, L+x)\right)\mathrm{d}x\,,$$

which tends to zero when $t \to \infty$. Thus the part of (4.27) corresponding to the second term in (4.28), i.e.

$$\int_{(\mathcal{C}_0^d)^2} \int_{W_1} \int_{W_n - n^{\frac{1}{d}}\tilde{x}_1} (\mathcal{M}_{(2)}(K, L+\tilde{x}_2; n, \tilde{x}_1) - \mathcal{M}_{(1)}(K; n, \tilde{x}_1)\mathcal{M}_{(1)}(L+\tilde{x}_2; n, \tilde{x}_1)) \times \mathbf{1}[|\tilde{x}_2| > t]\mathrm{d}\tilde{x}_2 \mathrm{d}\tilde{x}_1 \mathrm{d}\mathbb{Q}(K)\mathrm{d}\mathbb{Q}(L)\,, \quad (4.30)$$

converges to zero. We can justify these statements in detail similarly to [Blaszczyszyn et al., 2019, p.39-41]. The boundedness in (4.9) follows from the 2-moment condition (4.6) for the first term and from (4.13) for the second term. □

Sketch of the proof of Theorem 16. Denote $\bar{F}_n = F_n - \mathbb{E}F_n$. The idea is to prove that the k-th order cumulants of $(\mathrm{var}F_n)^{-1/2}(\bar{F}_n)$ vanish as $n \to \infty$ and k large. This follows by showing that (4.6) and (4.13) imply volume order growth (i.e., of order $O(n)$) for the k-th order cumulant of $\bar{F}_n$, $k \ge 2$, and using the assumption (4.10). Then (4.11) holds. The details are analogous to [Blaszczyszyn et al., 2019, p.43-49], with the difference that we deal with measures defined on $\mathcal{C}_n^d$. □

Conclusion

We have presented our results divided into three parts based on topics.

- The first part in Chapter 2 presented central limit theorem for U-statistics of Gibbs facet process based on method of moments and moment formulas for moments of such statistics.
- The second part shows parametric and semiparametric methods for estimation in facet (segment) processes. We illustrate usage of these methods in simulation study and on real data retrieved from stem cells imaging.
- The third part deals with Gibbs particle process modeling. We derive conditions under which the mean and variance asymptotics and the central limit theorem for some statistics of Gibbs particle processes hold.

Bibliography

A. Baddeley. *Spatial Point Processes and their Applications.* Springer, 01 2006.

A. Baddeley and G. Nair. Fast approximation of the intensity of Gibbs point processes. *Electron. J. Statist.*, 6, 2012.

A. Baddeley and R. Turner. Practical Maximum Pseudolikelihood for Spatial Point Patterns. *Australian & New Zealand Journal of Statistics*, 42(3), 2000.

B. Błaszczyszyn, E. Merzbach, and V. Schmidt. A note on expansion for functionals of spatial marked point processes. *Statistics & Probability Letters*, 36 (3), 1997.

V. Beneš and M. Zikmundová. Functionals of spatial point processes having a density with respect to the Poisson process. *Kybernetika*, 50, 2014.

V. Beneš, J. Večeřa, B. Eltzner, C. Wollnik, F. Rehfeldt, V. Kralova, and S. Huckemann. Estimation of parameters in a planar segment process with a biological application. *Image Analysis & Stereology*, 36(1), 2017.

V. Beneš, J. Večeřa, and M. Pultar. Planar Segment Processes with Reference Mark Distributions, Modeling and Estimation. *Methodology and Computing in Applied Probability*, 2019. doi: 10.1007/s11009-017-9608-x. Preprint ArXiv 1701.01893.

V. Beneš, Ch. Hofer-Temmel, G. Last, and J. Večeřa. Decorrelation of a class of Gibbs particle processes and asymptotic properties of U-statistics. 2019+. Preprint ArXiv 1903.06553.

P. Billingsley. *Probability and Measure.* Wiley, 1995.

B. Blaszczyszyn, D. Yogeshwaran, and J.E. Yukich. Limit theory for geometric statistics of point processes having fast decay of correlations. *Annals of Probability*, 47(2), 2019.

J.F. Coeurjolly, D. Dereudre, R. Drouilhet, and F. Lavancier. Takacs-Fiksel Method for Stationary Marked Gibbs Point Processes. *Scandinavian Journal of Statistics*, 39, 2011.

D. Dereudre, R. Drouilhet, and H.-O. Georgii. Existence of Gibbsian point processes with geometry-dependent interactions. *Probability Theory and Related Fields*, 153(3), 2012.

D. Dereudre, F. Lavancier, and K. Helisova. Estimation of the intensity parameter of the germ-grain Quermass-interaction model when the number of germs is not observed. *Scandinavian Journal of Statistics*, 41(3), 2014.

R.M. Dudley. *Real Analysis and Probability.* The Wadsworth & Brooks/Cole Mathematics Series. Wadsworth & Brooks/Cole Advanced Books & Software, 1989.

B. Eltzner, C. Wollnik, C. Gottschlich, S. Huckemann, and F. Rehfeldt. The filament sensor for near real-time detection of cytoskeletal fiber structures. *PLOS ONE*, 10(5), 2016.

A.J. Engler, S. Sen, H.L. Sweeney, and D.E. Discher. Matrix elasticity directs stem cell lineage specification. *Cell*, 126(4), 2006.

H.O. Georgii and H.J. Yoo. Conditional intensity and Gibbsianness of determinantal point processes. *Journal of Statistical Physics*, 118, 2005.

C.J. Geyer and J. Møller. Simulation and likelihood inference for spatial point processes. *Scandinavian Journal of Statistics*, 21(4), 1994.

C. Hofer-Temmel and P. Houdebert. Disagreement percolation for marked Gibbs point processes. Preprint ArXiv 1507.02521.

O. Kallenberg. *Random Measures, Theory and Applications.* Springer International Publishing, 2017.

G. Last. Perturbation analysis of Poisson processes. *Bernoulli*, 20(2), 05 2014.

G. Last and M.D. Penrose. Poisson process Fock space representation, chaos expansion and covariance inequalities. *Probability Theory and Related Fields*, 150(3-4), 2011.

G. Last and M.D. Penrose. *Lectures on the Poisson Process.* Cambridge University Press, UK United Kingdom, 12 2017.

G. Last, M.D. Penrose, M. Schulte, and Ch. Thäle. Moments and central limit theorems for some multivariate Poisson functionals. *Advances in Applied Probability*, 46(2), 2014.

K.V. Mardia and P.E. Jupp. *Directional Statistics.* Wiley, 1999.

S. Mase. Marked Gibbs Processes and Asymptotic Normality of Maximum Pseudo-Likelihood Estimators. *Mathematische Nachrichten*, 209, 2000.

K. Matthes, W. Warmuth, and J. Mecke. Bemerkungen zu einer Arbeit von Nguyen Xuan Xanh und Hans Zessin. *Mathematische Nachrichten*, 88(1), 1979.

R. Meester and R. Roy. *Continuum Percolation.* Cambridge University Press, 1996.

J. Møller and K. Helisová. Likelihood inference for unions of interacting discs. *Scandinavian Journal of Statistics*, 37, 2010.

J. Møller and R. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes.* Chapman and Hall/CRC, 2004.

M. Myllymäki, P. Grabarnik, H. Seijo, and D. Stoyan. Deviation test construction and power comparison for marked spatial point patterns. *Spatial Statistics*, 11, 2015.

M. Reitzner and M. Schulte. Central limit theorems for U-statistics of Poisson point processes. *Annals of Probability*, 41(6), 2013.

D. Ruelle. Superstable interactions in classical statistical mechanics. *Comm. Math. Phys.*, 18(2), 1970.

R. Schneider and W. Weil. *Stochastic and Integral Geometry*. Springer, Berlin, 2008.

T. Schreiber and J. Yukich. Limit theorems for geometric functionals of Gibbs point processes. *Annales de l'Institut Henri Poincaré*, 49(4), 2013.

G.L. Torrisi. Probability approximation of point processes with Papangelou conditional intensity. *Bernoulli*, 23(4A), 11 2017.

J. Večeřa. Central limit theorem for Gibbsian U-statistics of facet processes. *Applications of Mathematics*, 61(4), 2016.

J. Večeřa and V. Beneš. Interaction Processes for Unions of Facets, the Asymptotic Behaviour with Increasing Intensity. *Methodology and Computing in Applied Probability*, (4), 2016.

J. Večeřa and V. Beneš. Approaches to asymptotics for U-statistics of Gibbs facet processes. *Statistics & Probability Letters*, 122, 2017.

A. Xia and J. E. Yukich. Normal approximation for statistics of Gibbsian input in geometric probability. *Advances in Applied Probability*, 47(4), 12 2015.

A. Zemel, F. Rehfeldt, A.E.X. Brown, D.E. Discher, and S.A. Safran. Optimal matrix rigidity for stress-fibre polarization in stem cells. *Nature Physics*, 6(4), 2010.

List of Figures

List of Tables

List of publications

1. J. Večeřa and V. Beneš. Interaction Processes for Unions of Facets, the Asymptotic Behaviour with Increasing Intensity. *Methodology and Computing in Applied Probability*, (4), 2016
2. J. Večeřa. Central limit theorem for Gibbsian U-statistics of facet processes. *Applications of Mathematics*, 61(4), 2016
3. J. Večeřa and V. Beneš. Approaches to asymptotics for U-statistics of Gibbs facet processes. *Statistics & Probability Letters*, 122, 2017
4. V. Beneš, J. Večeřa, B. Eltzner, C. Wollnik, F. Rehfeldt, V. Kralova, and S. Huckemann. Estimation of parameters in a planar segment process with a biological application. *Image Analysis & Stereology*, 36(1), 2017
5. V. Beneš, J. Večeřa, and M. Pultar. Planar Segment Processes with Reference Mark Distributions, Modeling and Estimation. *Methodology and Computing in Applied Probability*, 2019. doi: 10.1007/s11009-017-9608-x. Preprint ArXiv 1701.01893
6. V. Beneš, Ch. Hofer-Temmel, G. Last, and J. Večeřa. Decorrelation of a class of Gibbs particle processes and asymptotic properties of U-statistics. 2019+. Preprint ArXiv 1903.06553