FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Daniela Flimmel

Asymptotic inference for stochastic geometry models

Department of Probability and Mathematical Statistics

Supervisor of the doctoral thesis: doc. RNDr. Zbyněk Pawlas, Ph.D.
Study programme: Mathematics
Study branch: P4M9

Prague 2021

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

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 ....

Author's signature

I would like to express my deep gratitude to my supervisor doc. Zbyněk Pawlas for all the help and guidance. Especially, I am grateful to him for enabling me to travel abroad to meet and cooperate with leading experts in the field. Thanks to him, I had a chance to cooperate with prof. Joseph Yukich and prof. Lothar Heinrich. I am grateful to them for giving me a great inspiration in my research and for the possibility to learn from them. I also want to thank Lehigh University and University of Augsburg for their hospitality during my visits.

I would also like to thank prof. Viktor Beneš as he was very supportive to me during my studies both academically and personally. This applies as well to all members of the Department of Probability and Mathematical Statistics.

My special thanks belong to my beloved family, especially my husband and my daughter for being patient, supportive and for sleeping well at night.

Title: Asymptotic inference for stochastic geometry models

Author: Daniela Flimmel

Department: Department of Probability and Mathematical Statistics

Supervisor: doc. RNDr. Zbyněk Pawlas, Ph.D., Department of Probability and Mathematical Statistics

Abstract: We compare three methods used in stochastic geometry in order to investigate asymptotic behaviour of random geometrical structures in large domains or in a large intensity regime. Namely, we describe in detail the Malliavin–Stein method, the method of stabilization and the method of cumulants. Then, we discuss some of its possible variants, combinations or extensions. Each method is supplemented with numerous examples concerning limit behaviour of different kinds of point processes, random tessellations and graphs or particle processes. Specially, for a geometric characteristic of the typical cell in a weighted Voronoi tessellation, we use the minus-sampling technique to construct an unbiased estimator of the average value of this characteristic and using the method of stabilization, we establish variance asymptotic and the asymptotic normality of such estimator. Next, we study asymptotic properties of a cylinder process in the plane derived by a Brillinger-type mixing point process. We prove a weak law of large numbers as well as a formula of the asymptotic variance for the area of the process. Under comparatively stronger assumptions, we also derive a central limit theorem for the cylinder process using the method of cumulants.

Keywords: Malliavin calculus, Stein's method, stabilization method, cumulant method, point processes, random tessellations, random graphs, particle processes, Poisson point processes, Brillinger-mixing point processes, Gibbs point processes

Contents

Introduction

Modelling of geometrical structures is of great interest within the scope of stochastic geometry since the structures appear in profusion in the natural sciences like geology, material sciences and astronomy as well as in technical sciences for example when studying communication networks (including social, transportation and wireless networks). Therefore, popular models include among others unions of random sets (particles), random tessellations or random graphs. Frequently, but it is not a rule, those structures are derived from a simple point process making it easier to track.

Many questions arising in stochastic geometry may be understood in terms of the behaviour of statistics of large random geometric structures. However, these random structures tend to have a problematic finite size description. Therefore, a natural way how to overcome this difficulty is to let the system size grow to infinity and study its asymptotic behaviour. This thesis aims to give a survey of selected limit techniques used in stochastic geometry as well as explore recent development and collect results providing laws of large numbers and central limit theorems for functionals of these random structures. The limiting regimes are first, increasing intensity of the defining point process and second, unboundedly growing observation window. It depends on the situation which approach is more suitable. For instance, we prefer to let the observation window grow when studying the volumes or lengths. On the other hand, for some scale-invariant functionals, one could let the intensity increase.

The structure of the thesis goes as follows:

- The first chapter introduces standard notation and theory of point processes in a general setting. Special situations are presented including spatial point processes, processes of compact subsets (particles) in $\mathbb{R}^d$ and processes with marks. Afterwards, further geometric structures are discussed such as random tessellations and random graphs. The theory is supplemented by various examples. The thesis aims to be self-contained. Therefore, the first chapter covers all the theory of point processes needed in the subsequent parts.

- The following three chapters each cover one asymptotic method. Each method is briefly explained. Proofs of the main results are included if they are short or interesting for the theory. Some of the proofs had to be adjusted to the setting of this thesis, some others were completed as they had only hints in the literature. Many examples of application are then presented, emphasizing the most recent ones and the author's own. The reader can also find references for further application. The methods are namely:

  Malliavin–Stein's method: The second chapter is devoted to the approach to probabilistic approximations that combines Stein's method with infinite-dimensional integration by parts formulae based on the use of Malliavin-type operators. The first stones of the method were built in a seminal paper Nualart and Peccati [2005], where the authors established central limit theorem called *the fourth moment theorem* for sequences of multiple stochastic integrals of a fixed order. Since this paper, a significant development appeared in Nourdin and Peccati [2009a], where by bringing together Stein's method with the Malliavin calculus, the authors were able to associate quantitative bounds to the fourth moment theorem. The basic idea of the approach is that, in order to assess the discrepancy between some Gaussian law and the distribution of a non-linear functional of a Gaussian field, one can apply infinite-dimensional integration by parts formulae from the Malliavin calculus of variations (see e.g. Malliavin [1997]) to the general bounds associated with the so-called Stein's method (see Stein [1972]) for probabilistic approximations. In particular, the Malliavin–Stein approach covers the ideas from Chatterjee [2009], where Stein's method was combined with finite-dimensional integration by parts formulae for Gaussian vectors, in order to deduce second order Poincaré inequalities.

Within the framework of this thesis, however, we use a version of Malliavin calculus for functionals of Poisson processes using the Fock space representation as was introduced in Peccati and Reitzner [2016]. The method then leads to central limit theorems as well as computing explicit rates of convergence for models in stochastic geometry.

An interested reader is strongly recommended to visit this webpage governed by Professor Nourdin which provides a constantly updated list of all existing papers written around the Malliavin–Stein method.

Method of stabilization: Another important tool of geometric limit theory is the concept of stabilization presented in the third chapter. Roughly speaking, a functional of some random structure stabilizes if its behaviour at a given location depends only on the environment within a certain finite but possibly random distance. This approach allows one to study statistics which may be expressed as a sum of spatially dependent terms having short-range interactions but complicated long-range dependence.

The motivation behind implementing this theory in the scope of stochastic geometry originated in a desire to understand the asymptotics of the classical Euclidean optimization problems including the traveling salesman problem initiated in Beardwood et al. [1959] by showing the limit law of the length of the shortest tour through n i.i.d points in the unit cube in $\mathbb{R}^d$. Independently, a similar result was shown in Miles [1970] for the total edge length of some planar tessellations driven by a homogeneous Poissonian input. The modern theory of stabilization used in stochastic geometry was introduced in its present form in Penrose and Yukich [2001], Penrose and Yukich [2002] and Penrose and Yukich [2003] based on previous work of Kesten and Lee [1996].

Method of cumulants: We use the classical result from probability theory by Marcinkiewicz [1939] stating that normal distribution is the only one with finitely many non-zero cumulants (semi-invariants). In the fourth chapter, we show that cumulants of a random variable derived from a random geometric structure can be expanded in terms of cumulant measures of the defining point process. Under suitable restrictions on those measures, the cumulants of higher orders can be shown to tend to zero yielding central limit theorems.

For a more thorough acquaintance with the cumulant method, the reader is referred to Saulis and Statulevičius [1991].

As far as we know, the above described methods are the most frequent ones when proving some asymptotic results in stochastic geometry. Nevertheless, the list is still far from being exhaustive. It would probably take more than one thesis to cover all available methods conscientiously. We shall at least mention limit theorems for geometric functionals enjoying variants of subadditivity or superadditivity properties (a detail survey of this subject is available in the monograph Yukich [1998]), techniques dealing with associated random fields (e.g. Bulinski and Shashkin [2007]) with the application on the volume of the excursion sets in Bulinski et al. [2021] or ideas based on constructing a clan of ancestors for Gibbsian inputs invented in Fernández et al. [1998].

At last, let us clear the connections between this thesis and author's publications.

- D. Flimmel and V. Beneš. *Gaussian approximation for functionals of Gibbs particle processes.* Kybernetika, 54:765-777, 2018.

  The paper follows recent development in the limit theory of functionals of Gibbs point processes in the Euclidean space in order to generalize results to Gibbs processes of geometrical objects (particles). First, the authors verified that the existence of a stationary Gibbs particle process is guaranteed under analogous conditions as stated by Dereudre [2017] for Gibbs point processes. Next, it was found that the methodology of Torrisi [2017] based on the Malliavin–Stein method can be applied to Gibbs particle processes. Based on these results, Gaussian approximation was derived for an innovation of a stationary Gibbs planar segment process. Namely two functionals were investigated: the normalized number of segments observed in a window and normalized total length of segments hitting the window.

  The results are presented here in order to demonstrate applications of Malliavin–Stein's method. However, we do not include all results or proofs since they already appeared in the master thesis Flimmel [2017].

- D. Flimmel, Z. Pawlas, and J. E. Yukich. *Limit theory for unbiased and consistent estimators of statistics of random tessellations.* Journal of Applied Probability, 57:679–702, 2020.

  The paper focuses on stationary generalized weighted Voronoi tessellations of $\mathbb{R}^d$ observed within a bounded observation window tending to the whole space. Given a geometric characteristic of the typical cell, we use the minus-sampling technique to construct an unbiased estimator of the average value of this geometric characteristic. Under mild conditions on the weights of the cells, we establish variance asymptotics and the asymptotic normality of the unbiased estimator as the observation window tends to the whole space using the stabilization properties of the generating point process. Moreover, the weak consistency is shown for this estimator. Specially, apart from already known results for Voronoi tessellations, stabilization properties are shown for Laguerre and Johnson–Mehl tessellations generated by a Poisson point process.

- D. Flimmel and L. Heinrich. *On the variance of the area of planar cylinder processes driven by Brillinger-mixing point processes.* Submitted to Electronic Journal of Probability.

  We study some asymptotic properties of cylinder processes in the plane defined as union sets of dilated straight lines (appearing as mutually overlapping infinitely long strips) derived from a stationary independently marked point process on the real line, where the marks describe thickness and orientation of individual cylinders. We observe such cylinder process in a domain ρK unboundedly growing to the whole space with $\rho \to \infty$. Provided the unmarked point process satisfies a Brillinger-type mixing condition and the thickness of the typical cylinder has a finite second moment we prove a (weak) law of large numbers as well as a formula of the asymptotic variance for the area of the cylinder process in ρK. Due to the long-range dependencies of the cylinder process, this variance increases proportionally to ρ^3. The main technique used in this paper is the expansion of the first two cumulants of the studied random variable in order to connect it with factorial cumulant measures of the defining point process.

  This paper is a starting point for deriving a central limit theorem using the cumulant method by showing that under similar assumptions, all the cumulants of orders three and higher converge to zero with $\rho \to \infty$. There have been some attempts, but so far it has led to an inadequate strengthening of the assumptions. Some results are demonstrated at the end of Chapter 4 to illustrate the application of the method. However, these results are not published and the authors agreed to continue the cooperation in order to obtain more promising results.

In conclusion, new results obtained during the authors PhD study are, namely:

Method of stabilization

- Theorem 3.12 and 3.13 proving the (asymptotic) unbiasedness and consistency of some estimators of a geometric characteristic of the typical cell in the weighted Voronoi tessellation,

- Theorem 3.14 showing the asymptotic variance and central limit theorem for the estimators mentioned above. The result is valid if the weighted Voronoi tessellation is generated by a stationary Poisson point process and any general weight function,

- Theorem 3.15 and 3.16 providing applications of the above listed results leading to limit theory for unbiased estimators of first, the distribution function of the volume and second, the Hausdorff measure of the boundary of the typical cell in a weighted Poisson–Voronoi tessellation, where the weight function relates either to Voronoi, Laguerre or Johnson–Mehl tessellation,

- Proposition 3.1 and 3.2 showing the stabilization properties of scores and cells of Voronoi, Laguerre and Johnson–Mehl tessellations generated by a Poisson input. In fact, these results extend some older results in McGivney and Yukich [1999], Penrose and Yukich [2001, 2003],
- Lemmas 3.3, 3.4, 3.5, 3.6, 3.7 and 3.8 serve as auxiliary results in the proofs of the above stated results.

Method of cumulants

- Lemma 4.1 connecting the Choquet functional of the cylinder process and the probability generating functional of its generating point process. This is a first important step to connect the cumulants of the area of the planar cylinder process with the cumulant measures of its generating point process. Special cases are presented in Corollary 4.1 and Example 4.2,
- Lemma 4.3 and Corollary 4.2 showing the convergence of the expected value of the area of the cylinder process generated by a Brillinger-mixing point process in $\mathbb{R}^1$. Theorem 4.3 then asserts a planar mean-square ergodic theorem,
- Theorem 4.4 providing the exact asymptotic variance of the area of the cylinder process generated by a point process with a strong version of Brillinger-mixing property,
- Theorem 4.5 is a unpublished result giving a central limit theorem for the area of the cylinder process driven by a point process satisfying much more strict assumptions on the factorial cumulant measures than Theorem 4.4,
- Lemmas 4.2, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10 and 4.11 are essential parts of the proofs of the above stated results.

1. Random geometric objects

The first chapter serves as an introduction to the theory of point processes as well as a recapitulation of definitions and results used in the subsequent chapters. It is mainly based on monographs Daley and Vere-Jones [2003] and Daley and Vere-Jones [2008], Rataj [2006], Baddeley [2007], Schneider and Weil [2008], Chiu et al. [2013] and Last and Penrose [2017].

1.1 Random measures and point processes in general setting

The aim of this section is to introduce the concept of the point process in a general way as a special type of random measure on a locally compact space. This notion covers processes of points in $\mathbb{R}^d$, compact sets, convex bodies, curves, lines, etc. Special examples will be discussed in the subsequent sections. If not stated otherwise, we assume that all random elements throughout this thesis are defined on a common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and by $\mathbb{E}$, Var, resp. Cov we denote the expectation, variance, resp. covariance w.r.t. $\mathbb{P}$.

Locally finite measures

Let $\mathbb{X}$ be a locally compact, separable space equipped with a metric ρ. Without loss of generality, we assume that every bounded closed set is compact with respect to ρ. Further in the text, we will often use the following standard notation:

$\mathcal{B}(\mathbb{X})$	...	Borel σ-field of subsets of $\mathbb{X}$,
$\mathcal{B}_b(\mathbb{X})$	...	bounded Borel sets,
$\mathcal{F}(\mathbb{X})$	...	closed sets,
$\mathcal{C}(\mathbb{X})$	...	compact sets.

Definition 1.1 (Finite measure).
A measure μ on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ is *finite*, if $\mu(\mathbb{X}) < \infty$.

Definition 1.2 (Locally finite measure).
A measure μ on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$ is *locally finite*, if $\mu(B) < \infty$ for all $B \in \mathcal{B}_b(\mathbb{X})$.

The notation for sets of measures will be used as follows:

$\mathbf{M}(\mathbb{X})$	...	space of all locally finite measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$,
$\mathbf{M}_f(\mathbb{X})$	...	space of all finite measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$,
$\mathbf{N}(\mathbb{X})$	...	space of all locally finite integer-valued measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$,
$\mathbf{N}_f(\mathbb{X})$	...	space of all finite integer-valued measures on $(\mathbb{X}, \mathcal{B}(\mathbb{X}))$.

If it does not lead to confusion, we use the simplified notation $\mathbf{M}, \mathbf{M}_f, \mathbf{N}, \mathbf{N}_f$ for the latter spaces. The elements of $\mathbf{N}$ are often called *counting measures*.

Moreover, denote by $\mathcal{M}$ the smallest σ-field on $\mathbf{M}$ which makes all the projections $\mu \mapsto \mu(B)$ measurable for all Borel sets B. Lemma 3.1.2 in Schneider and Weil [2008] shows that $\mathbf{N} \in \mathcal{M}$. By $\mathcal{N}$, we denote the trace σ-field of $\mathcal{M}$ on $\mathbf{N}$, i.e.

$$\mathcal{N} = \{M \cap \mathbf{N} : M \in \mathcal{M}\}.$$

For a measure $\mu \in \mathbf{M}$, the support $\mathrm{supp}(\mu)$ is the smallest closed set A in $\mathbb{X}$ such that $\mu(\mathbb{X} \backslash A) = 0$. Specially, if $\mu \in \mathbf{N}$, then

$$\mathrm{supp}(\mu) = \{x \in \mathbb{X} : \mu(\{x\}) \geqslant 1\}.$$

Example 1.1.
Take $k \in \mathbb{N} \cup \{0, \infty\}$ and $x_1, \ldots, x_k \in \mathbb{X}$. Define a counting measure μ by

$$\mu = \sum_{i=1}^{k} \delta_{x_i},$$

where for $x \in \mathbb{X}$, δ_x is the Dirac measure, i.e. for $A \in \mathcal{B}(\mathbb{X})$

$$\delta_x(A) := \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

Remark. Since we assumed $\mathbb{X}$ to be locally compact and separable, all measures in $\mathbf{M}_f(\mathbb{X})$ are regular (Štěpán [1987], Lemma I.7.1) and tight (Štěpán [1987], Lemma I.7.3).

Random measures

Definition 1.3 (Random measure).
A *random measure* on $\mathbb{X}$ is a measurable mapping

$$\Psi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbf{M}, \mathcal{M}).$$

The image measure $P_\Psi = \mathbb{P}\,\Psi^{-1}$ is the *distribution of the random measure* Ψ.

For Ψ a random measure and $B \in \mathcal{B}(\mathbb{X})$, we use the notation $\Psi(B)$ for the mapping $\omega \mapsto \Psi(\omega)(B)$.

Definition 1.4 (Intensity measure).
Let Ψ be a random measure on $\mathbb{X}$. The measure on $\mathbb{X}$ defined by

$$\alpha(B) = \mathbb{E}\left[\Psi(B)\right], \quad B \in \mathcal{B}(\mathbb{X}),$$

is called the *intensity measure* of the random measure Ψ.

Remark. Since $\Psi(B) \geqslant 0, \alpha(B)$ is always defined, but α does no longer need to be locally finite. However, the intensity measures of random measures and point processes in this text will always be assumed to be locally finite.

Theorem 1.1 (Campbell).
Let Ψ be a random measure on $\mathbb{X}$ with intensity measure α and let $f : \mathbb{X} \to \mathbb{R}$ be a non-negative, measurable function. Then

$$\mathbb{E} \int_{\mathbb{X}} f \,\mathrm{d}\Psi = \int_{\mathbb{X}} f \,\mathrm{d}\alpha.$$

The intensity measure is also known under the name *first moment measure*. Let us now introduce the higher moment measures in a similar way.

Definition 1.5 (m-th moment measure).
Let Ψ be a random measure on $\mathbb{X}$ and let $m \in \mathbb{N}$. The *m-th moment measure* $\alpha^{(m)}$ of Ψ is the Borel measure on $\mathbb{X}^m$ for which

$$\alpha^{(m)}(B_1 \times \cdots \times B_m) := \mathbb{E}\,\Psi^m(B_1 \times \cdots \times B_m) = \mathbb{E}\,\Psi(B_1) \cdots \Psi(B_m),$$

where $B_1, \ldots, B_m \in \mathcal{B}(\mathbb{X})$.

Here, the product measure Ψ^m is a random measure on $\mathbb{X}^m$ and therefore, $\alpha^{(m)}$ is the intensity measure of Ψ^m in the sense of Definition 1.4.

For $m \in \mathbb{N}$, define the space

$$\mathbb{X}^m_{\neq} := \{(x_1, \ldots, x_m) \in \mathbb{X}^m : x_i \text{ are pairwise distinct}\}.$$

Definition 1.6 (m-th factorial moment measure).
Let $m \in \mathbb{N}$. For a random measure Ψ, we define the *m-th factorial moment measure* as the Borel measure $\alpha^{[m]}$ on $\mathbb{X}^m$ for which

$$\alpha^{[m]}(B_1 \times \cdots \times B_m) := \mathbb{E}\,\Psi^m((B_1 \times \cdots \times B_m) \cap \mathbb{X}^m_{\neq}),$$

where $B_1, \ldots, B_m \in \mathcal{B}(\mathbb{X})$.

Theorem 1.2 (Campbell's theorem for higher-order moment measures).
Let Ψ be a random measure and $f : \mathbb{X}^m \to \mathbb{R}$ be a non-negative measurable function. Then

$$\mathbb{E} \int_{\mathbb{X}^m} h(x_1, \ldots, x_m)\Psi^m(\mathrm{d}(x_1, \ldots, x_m)) = \int_{\mathbb{X}^m} f(x_1, \ldots, x_m)\alpha^{(m)}(\mathrm{d}(x_1, \ldots, x_m))$$

and

$$\mathbb{E} \int_{\mathbb{X}^m_{\neq}} h(x_1, \ldots, x_m)\Psi^{[m]}(\mathrm{d}(x_1, \ldots, x_m)) = \int_{\mathbb{X}^m_{\neq}} f(x_1, \ldots, x_m)\alpha^{[m]}(\mathrm{d}(x_1, \ldots, x_m)),$$

where $\Psi^{[m]}$ is the restriction of Ψ^m on $\mathbb{X}^m_{\neq}$.

Point processes

A point process is a special example of a random measure. A simple point process is a random measure that can be described by a locally finite sum of Dirac measures, i.e. a random collection of isolated points producing no multiplicities in $\mathbb{X}$. It can also be defined as a random closed set in $\mathbb{X}$ which is almost surely locally finite. Multiplicities in a process can also be treated using marks attached to each point (see Section 1.3).

Definition 1.7 (Point process).
A *point process* on $\mathbb{X}$ is a measurable mapping

$$\mu : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbf{N}, \mathcal{N}).$$

Remark. The measurable space $(\mathbf{N}, \mathcal{N})$ is often called *the outcome space of a point process* on $\mathbb{X}$.

Definition 1.8 (Simple point process).
A point process μ on $\mathbb{X}$ is called *simple* if

$$\mathbb{P}(\mu(\{x\}) \leqslant 1,\ \forall x \in \mathbb{X}) = 1.$$

Definition 1.9 (Distribution of a point process).
Let μ be a point process on $\mathbb{X}$. By the *distribution of the point process* μ, we understand the probability measure P_μ on the space $(\mathbf{N}, \mathcal{N})$ given by

$$P_\mu(A) = \mathbb{P}(\mu \in A) = \mathbb{P}(\{\omega \in \Omega : \mu(\omega) \in A\}), \quad A \in \mathcal{N}.$$

Notation. Point processes can be considered either as random measures or as random sets of discrete points. Due to this interpretation, we will often treat them accordingly and for $\mu \in \mathbf{N}$ write $x \in \mu$ instead of $x \in \operatorname{supp}(\mu)$. At the same time, we denote by $\mu(B) = n$ the fact that the set B contains n points of μ. Among other reasons, it allows us to simplify the notation for the mean values and write

$$\mathbb{E} \sum_{x \in \mu} f(x)$$

instead of

$$\int_{\mathbf{N}} \int_{\mathbb{R}^d} f(x) \phi(\mathrm{d}x) P_\mu(\mathrm{d}\phi).$$

Moreover, if μ is simple, we will write $\mu = \{x_1, x_2, \ldots\}$.

The characteristics of random measures can be defined in the same way for point processes. However, some definitions and results have simpler interpretation. For example, if μ is a simple point process, then the intensity measure α evaluated in some set $B \in \mathcal{B}(\mathbb{X})$ is the mean number of points of μ lying in B.

Remark. If μ is a simple point process, then the Campbell's theorem takes form

$$\mathbb{E} \sum_{x \in \mu} f(x) = \int_{\mathbb{X}} f \mathrm{d}\alpha. \tag{1.1}$$

The next theorem is a version of Campbell's theorem for the m-th moment measure $\alpha^{(m)}$ and the m-th factorial moment measure $\alpha^{[m]}$ of a simple point process μ and it is formulated in Schneider and Weil [2008].

Theorem 1.3 (Campbell's theorem for a simple point process).
Let μ be a simple point process in $\mathbb{X}$ and let $f : \mathbb{X}^m \to \mathbb{R}$ be a non-negative measurable function. Then

$$\mathbb{E} \sum_{(x_1,\ldots,x_m) \in \mu^m} f(x_1, \ldots, x_m) = \int_{\mathbb{X}^m} f(x_1, \ldots, x_m) \alpha^{(m)}(\mathrm{d}(x_1, \ldots, x_m))$$

and

$$\mathbb{E} \sum_{(x_1,\ldots,x_m) \in \mu^m \cap \mathbb{X}^m_{\neq}} f(x_1, \ldots, x_m) = \int_{\mathbb{X}^m} f(x_1, \ldots, x_m) \alpha^{[m]}(\mathrm{d}(x_1, \ldots, x_m)).$$

Using Theorem 1.3, we can see the relation between the measures $\alpha^{(2)}$ and $\alpha^{[2]}$. For a simple point process μ on $\mathbb{X}$ and $B_1, B_2 \in \mathcal{B}(\mathbb{X})$, we have

$$\alpha^{(2)}(B_1 \times B_2) = \mathbb{E}\left(\sum_{(x_1,x_2)\in\mu^2\cap\mathbb{X}^2_{\neq}} \mathbf{1}_{B_1\times B_2}(x_1,x_2) + \sum_{x\in\mu} \mathbf{1}_{B_1}(x)\mathbf{1}_{B_2}(x) \right).$$

Hence,

$$\alpha^{(2)}(B_1 \times B_2) = \alpha^{[2]}(B_1 \times B_2) + \alpha(B_1 \cap B_2). \tag{1.2}$$

A recurrent relation between higher-order moment measures and factorial moment measures can be found. Before stating it, we recall the notion of Stirling number of the second kind. We follow the definition of Daley and Vere-Jones [2003], Section 5.2. For any integer n and k, we define the factorial powers of n by

$$n^{[k]} := \begin{cases} n(n-1)\cdots(n-k+1), & \text{for } k = 0,\dots,n, \\ 0, & \text{for } k > n. \end{cases}$$

Definition 1.10 (Stirling number of the second kind).
The *Stirling number of the second kind* $\left\{ {k \atop l} \right\}$ is defined by the relation

$$n^k = \sum_{l=1}^{k} \left\{ {k \atop l} \right\} n^{[l]}$$

whenever $n \geqslant k$.

Alternatively, the Stirling number of the second kind can be defined using an explicit formula

$$\left\{ {k \atop l} \right\} = \begin{cases} \frac{1}{l!}\sum_{j=0}^{l}(-1)^j \binom{l}{j}(l-j)^k, & \text{if } k \geqslant l, \\ 0, & \text{if } k < l. \end{cases}$$

Having the notion of Stirling numbers of the second kind, we can now describe a connection between moment measures and factorial moment measures. For $A \in \mathcal{B}(\mathbb{X})$ and $k \in \mathbb{N}$ we have

$$\alpha^{(k)}(\times_{i=1}^{k} A) = \sum_{l=1}^{k} \left\{ {k \atop l} \right\} \alpha^{[l]}(\times_{i=1}^{l} A).$$

Definition 1.11 (Probability generating functional).
Denote by $\mathcal{G}(\mathbb{X})$ the class of Borel functions $w : \mathbb{X} \to [0,1]$ with $1 - w$ vanishing outside some bounded set. For a point process μ on $\mathbb{X}$, we define the *probability generating functional* G_μ by

$$G_\mu(w) = \mathbb{E}\left(\prod_{x\in\mu} w(x) \right), \quad w \in \mathcal{G}(\mathbb{X}).$$

As in the case of ordinary random variables, the probability generating functional is associated with the moment structure of the point process.

Theorem 1.4 (Proposition 9.5.VI in Daley and Vere-Jones [2008]). *Let G_μ be the probability generating functional of a point process μ on $\mathbb{X}$ whose k-th moment measure exists for $k \in \mathbb{N}$. Then for $1 - w \in \mathcal{G}(\mathbb{X})$ and $\rho \to 0$,*

$$G_\mu(1-\rho w) = 1+\sum_{j=1}^{k} \frac{(-\rho)^j}{j!} \int_{\mathbb{X}^j_{\neq}} w(x_1)\dots w(x_j)\, \alpha^{[j]}(\mathrm{d}x_1 \times\cdots\times \mathrm{d}x_j)+o(\rho^k). \quad (1.3)$$

Theorem 1.5 (Corollary 9.5.VII in Daley and Vere-Jones [2008]). *Under the conditions of Theorem 1.4, if the $(k+1)$-th moment measure of μ exists, then the remainder in (1.3) is bounded by*

$$\frac{\rho^{k+1}}{(k+1)!} \int_{\mathbb{X}^{k+1}_{\neq}} w(x_1)\cdots w(x_{k+1})\alpha^{[k+1]}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_{k+1}).$$

Similar expression as in Theorem 1.4 leads to the definition of factorial cumulant measures associated with a point process μ when expanding the logarithm of the probability generating functional instead of the probability generating functional itself.

Theorem 1.6 (Corollary 9.5.VIII in Daley and Vere-Jones [2008]). *Under the conditions of Theorem 1.4, the probability generating functional can be expressed using the factorial cumulant measures $\gamma^{[j]}$ for $\rho \to 0$, as*

$$\log G_\mu(1-\rho w) = \sum_{j=1}^{k} \frac{(-\rho)^j}{j!} \int_{\mathbb{X}^j_{\neq}} w(x_1)\cdots w(x_j)\gamma^{[j]}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_j) + o(\rho^k).$$

Factorial cumulant measures form a useful tool in expressing and studying dependencies among distant parts of a point pattern. The relation between factorial moment measures and factorial cumulant measures is based on the general relationship between mixed moments and mixed cumulants (see e.g. Chapter 4 in Baccelli et al. [2020]) and can serve as an alternative definition for the measures. For $k \in \mathbb{N}$, we have

$$\gamma^{[k]}(B_1\times\cdots\times B_k) = \sum_{j=1}^{k}(-1)^{j-1}(j-1)! \sum_{K_1\cup\cdots\cup K_j=\{1,\dots,k\}} \prod_{r=1}^{j} \alpha^{[|K_r|]}(\times_{s\in K_r} B_s), \quad (1.4)$$

$$\alpha^{[k]}(B_1 \times \cdots \times B_k) = \sum_{j=1}^{k} \sum_{K_1\cup\cdots\cup K_j=\{1,\dots,k\}} \prod_{r=1}^{j} \gamma^{[|K_r|]}(\times_{s\in K_r} B_s), \quad (1.5)$$

where $B_1, \dots, B_k \in \mathcal{B}(\mathbb{X})$ and the sum $\sum_{K_1\cup\cdots\cup K_j=\{1,\dots,k\}}$ is taken over all partitions of the set $\{1, \dots, k\}$ into j non-empty sets $K_1, \dots, K_j$ and $|K_i|$ denotes the cardinality of the set K_i. Note that $\gamma^{[k]}$ is a locally finite, signed measure on $[\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)]$.

Example 1.2.
The first factorial cumulant measure coincides with the intensity measure (first

moment measure), whereas the second factorial cumulant measure equals to

$$\begin{aligned}\gamma^{[2]}(B_1 \times B_2) &= \alpha^{[2]}(B_1 \times B_2) - \alpha(B_1)\alpha(B_2)\\ &= \alpha^{(2)}(B_1 \times B_2) - \alpha(B_1 \cap B_2) - \alpha(B_1)\alpha(B_2)\\ &= \operatorname{cov}(\mu(B_1), \mu(B_2)) - \alpha(B_1 \cap B_2),\end{aligned}$$

where in the second equality, we used relation (1.2).

Besides the probability generating functional, the most useful transform of a random measure is the Laplace functional.

Definition 1.12 (Laplace functional).
Let Ψ be a random measure and f a non-negative measurable function of bounded support on $\mathbb{X}$. Then the *Laplace functional* is defined by

$$L_\Psi(f) = \mathbb{E}\left[\exp\left(-\int_{\mathbb{X}} f(x)\Psi(\mathrm{d}x)\right)\right].$$

Analogically, the k-th cumulant measures of the point process μ can be defined either via expanding the Laplace functional or based on the relation with the moment measures up to k-th order. The next theorem shows the Taylor series expansion of the Laplace functional about $f \equiv 0$.

Theorem 1.7 (Proposition 4.2.2 in Baccelli et al. [2020]).
Let Ψ be a random measure on $\mathbb{X}$ and $f : \mathbb{X} \to \mathbb{R}_+$ be a measurable function such that the mapping $(x_1, \dots, x_k) \mapsto f(x_1)\cdots f(x_k)$ is integrable with respect to $\alpha^{(k)}$, i.e.

$$\int_{\mathbb{X}^k} f(x_1)\cdots f(x_k)\alpha^{(k)}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_k) < \infty.$$

Then, for $t \in \mathbb{R}_+$

$$L_\Psi(tf) = 1 + \sum_{j=1}^{k} \frac{(-t)^j}{j!} \int_{\mathbb{X}^j} f(x_1)\cdots f(x_j)\alpha^{(j)}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_j) + \frac{t^k}{k!}\epsilon_k(t),$$

where $|\epsilon_k(t)| \leqslant \int_{\mathbb{X}^k} f(x_1)\cdots f(x_k)\alpha^{(k)}(\mathrm{d}x_1 \times\cdots\times \mathrm{d}x_k)$ and $\lim_{t\to 0}\epsilon_k(t) = 0$. Moreover, for $t \to 0$,

$$\log L_\Psi(tf) = \sum_{j=1}^{k} \frac{(-t)^j}{j!} \int_{\mathbb{X}^j} f(x_1)\cdots f(x_j)\gamma^{(j)}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_j) + o(t^k),$$

where $\gamma^{(j)}$ is the j-th cumulant measure of Ψ.

Take $k \in \mathbb{N}$. The cumulant measures and moment measures are related by (cf. Baccelli et al. [2020], Chapter 4)

$$\gamma^{(k)}(B_1 \times \cdots \times B_k) = \sum_{j=1}^{k}(-1)^{j-1}(j-1)! \sum_{K_1\cup\cdots\cup K_j=\{1,\dots,k\}} \prod_{r=1}^{j} \alpha^{(|K_r|)}(\times_{s\in K_r} B_s),$$

$$\alpha^{(k)}(B_1 \times \cdots \times B_k) = \sum_{j=1}^{k} \sum_{K_1\cup\cdots\cup K_j=\{1,\dots,k\}} \prod_{r=1}^{j} \gamma^{(|K_r|)}(\times_{s\in K_r} B_s),$$

where $B_1, \dots, B_k \in \mathcal{B}(\mathbb{X})$.

Example 1.3.
The first cumulant measure also coincides with the intensity measure, the second cumulant measure is equal to the *covariance measure*

$$\gamma^{(2)}(A \times B) = \alpha^{(2)}(A \times B) - \alpha(A)\alpha(B) = \mathrm{cov}(\mu(A), \mu(B)).$$

Palm distributions

In this section we focus on a type of conditional distributions for random measures which are formally defined in terms of so-called *Palm distributions*, first introduced in Palm [1943].

Definition 1.13 (Campbell measure).
Campbell measure of a random measure Ψ on $\mathbb{X}$ is defined as

$$\mathbf{C}(B \times \mathcal{U}) = \mathbb{E}\left[\Psi(B)\mathbf{1}\{\Psi \in \mathcal{U}\}\right], \quad B \in \mathcal{B}(\mathbb{X}), \mathcal{U} \in \mathcal{M}.$$

It represents a refinement of the intensity measure $\alpha(B) = \mathbf{C}(B \times \mathcal{M})$. By Lemma 3.1.1 in Baccelli et al. [2020], the Campbell measure is a unique σ-finite measure on $\mathbb{X} \times \mathcal{M}$. Moreover, assuming the intensity measure α of the random measure Ψ is locally finite, then the measure disintegration theorem allows to disintegrate the Campbell measure in a way that

$$\mathbf{C}(B \times \mathcal{U}) = \int_B \mathbb{P}^x(\mathcal{U})\alpha(\mathrm{d}x), \quad B \in \mathcal{B}(\mathbb{X}), \mathcal{U} \in \mathcal{M}, \tag{1.6}$$

where $P^{\cdot}(\cdot)$ is a probability kernel from $\mathbb{X}$ to $\mathcal{M}$. Note that the family $\{\mathbb{P}^x(\cdot)\}_{x \in \mathbb{X}}$ is unique α-almost everywhere.

Definition 1.14 (Palm distribution).
If $\mathbb{P}^{\cdot}(\cdot)$ is the probability kernel defined by (1.6), then $\mathbb{P}^x(\cdot)$ is called *Palm distribution* of Ψ at point $x \in \mathbb{X}$ and $\{\mathbb{P}^x(\cdot)\}_{x \in \mathbb{X}}$ is a *family of Palm distributions* of Ψ.

In another words, $\mathbb{P}^{\cdot}(\mathcal{U})$ is the Radon–Nikodym derivative of $\mathbf{C}(\cdot \times \mathcal{U})$ with respect to the intensity measure, i.e.

$$\mathbb{P}^{\cdot}(\mathcal{U}) = \frac{\mathrm{d}\mathbf{C}(\cdot \times \mathcal{U})}{\mathrm{d}\alpha(\cdot)}.$$

Heuristically speaking, if μ is a point process with intensity measure α, then $\mathbb{P}^x(\mathcal{U})$ interprets as a conditional probability that $\mu \in \mathcal{U}$ given that there is a point of the process μ located at x. Note that for a point process without a fixed atom at this particular location, the probability of the condition is null. Hence, the basic discrete definition of the conditional probability does not apply. Moreover, we would have that

$$\mathbb{P}^x(\{\nu \in \mathcal{N}; \nu(\{x\}) \geqslant 1\}) = 1, \quad \text{for } \alpha\text{-a.a. } x \in \mathbb{X}.$$

For a simple point process, we may define a modified version of the Campbell measure and the Palm distribution by removing a point that may be present at location x.

Definition 1.15 (Reduced Palm distribution).
For a point process μ, we define the *reduced Palm distribution* at point $x \in \mathbb{X}$ by

$$\mathbb{P}_x^!(\mathcal{U}) := \mathbb{P}^x(\mathcal{U} + \delta_x), \quad \mathcal{U} \in \mathcal{N},$$

where $\mathcal{U} + \delta_x = \{\nu + \delta_x; \nu \in \mathcal{U}\}$.

Theorem 1.8 (Refined Campbell theorem).
For a simple point process μ with intensity measure α, it holds that

$$\mathbb{E} \sum_{x \in \mu} h(x, \mu) = \int_{\mathbb{X} \times \mathcal{N}} h(x, \nu) \mathbf{C}(\mathrm{d}(x, \nu)) = \int_{\mathbb{X}} \int_{\mathcal{N}} h(x, \nu) \mathbb{P}^x(\mathrm{d}\nu) \alpha(\mathrm{d}x),$$

$$\mathbb{E} \sum_{x \in \mu} h(x, \mu \backslash \{x\}) = \int_{\mathbb{X}} \int_{\mathcal{N}} h(x, \nu) \mathbb{P}_x^!(\mathrm{d}\nu) \alpha(\mathrm{d}x)$$

for any non-negative measurable $h : \mathbb{X} \times \mathcal{N} \to \mathbb{R}$.

An important tool, in particular in the analysis of spatial point process (see Section 1.2), is the Papangelou conditional intensity, first introduced in Papangelou [1974]. The definition relies on the so-called reduced Campbell measure.

Definition 1.16 (reduced Campbell measure).
The *reduced Campbell measure* $\mathbf{C}^!$ of a point process μ is defined by

$$\mathbf{C}^!(B \times \mathcal{U}) = \mathbb{E}\left[\mu(B) \mathbf{1}\{\mu - \delta_x \in \mathcal{U}\}\right], \quad B \in \mathcal{B}(\mathbb{X}), \mathcal{U} \in \mathcal{N}.$$

Definition 1.17 (Papangelou conditional intensity).
Let μ be a point process on $\mathbb{X}$ and suppose its reduced Campbell measure $\mathbf{C}^!$ is absolutely continuous with respect to the product measure $\rho \otimes P_\mu$. Then any Radon–Nikodym density λ^* of $\mathbf{C}^!$ relative to $\rho \otimes P_\mu$ is called the *Papangelou conditional intensity* of μ.

For any non-negative measurable $f : \mathbb{X} \times \mathbf{N}(\mathbb{X}) \to \mathbb{R}$, we have

$$\mathbb{E}\left[\sum_{x \in \mu} f(x, \mu \backslash \{x\})\right] = \int_{\mathbb{X}} \mathbb{E}\left[\lambda^*(x, \mu) f(x, \mu)\right] \rho(\mathrm{d}x).$$

The Papangelou conditional intensity $\lambda^*(x, \mathbf{x})\, \rho(\mathrm{d}x)$ for a point configuration $\mathbf{x} \in \mathbf{N}(\mathbb{X})$ has an intuitive interpretation of the conditional probability of observing one point in the infinitesimally small set conditional on that μ agrees with $\mathbf{x}$ outside this set.

Examples

Example 1.4 (Binomial point process).
Let Q be a probability measure on $\mathbb{X}$ and let $n \in \mathbb{N}$. Assume that $X_1, \ldots, X_n$ are independent random elements on $\mathbb{X}$ following the same law Q. Then

$$\mu = \delta_{X_1} + \cdots + \delta_{X_n}$$

is a point process on $\mathbb{X}$ called the binomial point process with sample size n and sampling distribution Q.

Having μ the binomial point process with sample size n and sampling distribution Q, then, for $B \in \mathcal{B}(\mathbb{X})$, $\mu(B)$ follows the binomial distribution, i.e.

$$\mathbb{P}(\mu(B) = k) = \binom{n}{k} Q(B)^k (1 - Q(B))^{n-k}, \quad k = 0, \ldots, n.$$

The intensity measure of the process μ is then

$$\alpha(B) = \mathbb{E} \sum_{i=1}^{n} \mathbf{1}\{X_i \in B\} = \sum_{i=1}^{n} \mathbb{P}(X_i \in B) = n\, Q(B).$$

Example 1.5 (Poisson point process).
Let α be a locally finite non-atomic measure on $\mathbb{X}$. The *Poisson point process* on $\mathbb{X}$ with intensity measure α is a point process η on $\mathbb{X}$ satisfying the following conditions

1. for every compact set $B \subset \mathbb{X}$, $\eta(B)$ is a Poisson distributed random variable with parameter $\alpha(B)$;
2. if $B_1, \ldots, B_n$, $n \in \mathbb{N}$, are pairwise disjoint compact subsets of $\mathbb{X}$, then $\eta(B_1), \ldots, \eta(B_n)$ are independent random variables.

Remark. For a Poisson point process η on $\mathbb{X}$ with intensity measure α and $k \in \mathbb{N}$, the k-th factorial moment measure equals α^k (cf. Corollary 3.2.4 in Schneider and Weil [2008]). Moreover, as shown in Example 4.2 in Chiu et al. [2013], the probability generating functional equals to

$$G_\eta(w) = \exp\left(- \int_{\mathbb{X}} (1 - w(x))\alpha(\mathrm{d}x) \right). \tag{1.7}$$

From the latter expression, we conclude that (cf. Baccelli et al. [2020], Example 4.1.13) the first factorial cumulant measure of η equals its intensity measure and for $k \geqslant 2$, the k-th factorial cumulant measure is null.

From now on, we will use the notation η exclusively for the Poisson point process, the underlying space shall always be specified. For the fundamental properties including the existence of a general Poisson process, see Last and Penrose [2017].

Theorem 1.9 (Slivnyak–Mecke formula).
Let η be a Poisson point process on $\mathbb{X}$ with intensity measure α, take $m \in \mathbb{N}$ and let $f : \mathbf{N} \times \mathbb{X}^m \to \mathbb{R}$ be a non-negative measurable function. Then

$$\begin{aligned} &\mathbb{E}\left[\sum_{(x_1,\ldots,x_m)\in\mathbb{X}^m_{\neq}} f(\eta, (x_i)_{i=1,\ldots,m}) \right] \\ &= \int_{\mathbb{X}} \cdots \int_{\mathbb{X}} \mathbb{E}\, f\left(\eta + \sum_{i=1}^{m} \delta_{x_i}, (x_i)_{i=1,\ldots,m} \right) \alpha(\mathrm{d}x_1) \cdots \alpha(\mathrm{d}x_m). \end{aligned}$$

Example 1.6 (Cox process).
Let Ψ be a random measure with a.s. no atoms. The Cox process (or the doubly stochastic Poisson process) μ directed by Ψ is a Poisson process on $\mathbb{X}$ with random intensity measure Ψ, i.e. the distribution of μ is given by

$$\mathbb{P}(\mu(B) = k) = \int_{\mathbf{M}(\mathbb{X})} e^{-\nu(B)} \frac{\nu(B)^k}{k!} P_\Psi(\mathrm{d}\nu), \quad k \in \mathbb{N} \cup \{0\}, B \in \mathcal{B}(\mathbb{X}).$$

The Cox process is indeed a well defined point process on $\mathbb{X}$ (see Section 6.2 in Daley and Vere-Jones [2003]). The intensity measure of a Cox process μ is

$$\mathbb{E}\,\mu(B) = \int \mathbb{E}\,\eta_\alpha(B) P_\Psi(\mathrm{d}\alpha) = \int \alpha(B) P_\Psi(\mathrm{d}\alpha) = \mathbb{E}\,\Psi(B),$$

where η_α denotes the Poisson point process with intensity measure α.

Example 1.7 (Cluster point process).
Let μ_P be a point process on $\mathbb{X}$ (parent point process) and let $\{\xi_x, x \in \mathbb{X}\}$ be a collection of finite point processes (daughter point processes), i.e $\xi_x(\mathbb{X}) < \infty$ a.s. for all $x \in \mathbb{X}$. Take $B \in \mathcal{B}(\mathbb{X})$ and define a point process μ by

$$\mu(B) = \int_{\mathbb{X}} \xi_x(B) \mu_P(\mathrm{d}x).$$

If $\mu(B) < \infty$ a.s. for all $B \in \mathcal{B}(\mathbb{X})$, then μ is called a cluster point process.

A special example of a cluster point process is a *Poisson cluster process*, which is a cluster process such that its daughter processes are mutually independent, independent of the parent process, which is a Poisson point process. If we take $\mathbb{X} = \mathbb{R}^d$ and assume, moreover, μ_P to be stationary with intensity λ_P (see Definitions 1.18 and 1.19) and that the daughter processes have the number of points distributed according to some law N_0 that are placed around the origin with respect to a common density f, then we talk about the *Neyman–Scott process*. Depending on the choice of the distribution of the number of points and density f, we meet special examples such as *Matérn cluster process* (Matérn [1986]), *Gauss–Poisson process*, etc.

Example 1.8 (Finite point process with density with respect to the distribution of the Poisson point process).
A point process μ on $\mathbb{X}$ is called finite if

$$\mathbb{P}(\mu(\mathbb{X}) < \infty) = 1.$$

Let η be a finite Poisson point process on $\mathbb{X}$ with intensity measure α and distribution P_η. Consider a measurable mapping $p : \mathbf{N}_f \to \mathbb{R}_+$ satisfying

$$\int_{\mathbf{N}_f} p(\mathbf{x}) \mathrm{d}P_\eta(\mathbf{x}) = 1.$$

A point process μ with distribution P_μ, such that

$$P_\mu(A) = \int_A p(\mathbf{x}) \mathrm{d}P_\eta(\mathbf{x}), \quad A \in \mathcal{N}(\mathbb{X}),$$

is called the point process with density p with respect to the distribution of the Poisson point process η.

For a point process with density p with respect to the distribution of Poisson point process η on $\mathbb{R}^d$, the following representation of the Papangelou conditional intensity (see Definition 1.17) holds true.

Theorem 1.10 (Theorem 4.1 in Baddeley [2007]).
Let μ be a finite point process in a bounded set $\Lambda \subset \mathbb{R}^d$ with density p with respect to a finite Poisson process η. Assume that the density p is hereditary, i.e. satisfies

$$p(\mathbf{x}) > 0 \Rightarrow p(\mathbf{y}) > 0, \quad \mathbf{x}, \mathbf{y} \in \mathbf{N}_f, \mathbf{y} \subset \mathbf{x}.$$

Then the Papangelou conditional intensity of the point process μ exists and equals

$$\lambda^*(x, \mathbf{x}) = \frac{p(\mathbf{x} \cup \{x\})}{p(\mathbf{x})}, \quad \mathbf{x} \in \mathbf{N}_f, x \in \Lambda, x \notin \mathbf{x}.$$

If $p(\mathbf{x}) = 0$, we set $\lambda^*(x, \mathbf{x}) = 0$.

Theorem 1.11 (Lemma 4.2 in Baddeley [2007]).
Let μ be a finite point process in a bounded set $\Lambda \subset \mathbb{R}^d$ with density p with respect to a finite Poisson process η and a Papangelou conditional intensity λ^*. Then p is completely determined by λ^*.

For example, the *Strauss point process* on $\mathbb{R}^d$ is constructed as a finite point process having density p with respect to the distribution of the Poisson point process η with intensity measure $\alpha(B) = |B|_d$ (d-dimensional Lebesgue measure of B) for all $B \subset \Lambda$ with

$$p(\mathbf{x}) = C\beta^{\#(\mathbf{x})}\gamma^{s(\mathbf{x})},$$

where C is the normalising constant, $\beta > 0, 0 \leqslant \gamma \leqslant 1, r > 0$ are parameters, $\#(\mathbf{x})$ denotes the number of points in $\mathbf{x}$ and

$$s(\mathbf{x}) = \sum_{x,y \in \mathbf{x}} \mathbf{1}\{\|x - y\| < r\}$$

is the number of pairs in a configuration $\mathbf{x}$ in Λ being at most r units apart from each other. By using $\gamma = 0$, one obtains the *hard-core process*. The choice $\gamma = 1$ returns back the Poisson point process. By applying Theorem 1.10, the Papangelou conditional intensity of the Poisson point process with intensity measure $\alpha(B) = \beta|B|_d, B \subset \Lambda$, simplifies to $\lambda^*(x, \mathbf{x}) = \beta$.

1.2 Spatial point processes

Most commonly in applications, we meet the case when $\mathbb{X} = \mathbb{R}^d$ (where usually $d = 2$ or $d = 3$). Spatial point processes are useful as statistical models in the analysis of observed patterns of points, where the points represent the locations of some object of study (trees in a forest, disease cases, etc.). The linearity of the Euclidean space allows defining stationary random measures and point processes.

In the previous section, we defined the factorial moment measures of a random measure (a point process, resp.). Assuming those measures are absolutely continuous with respect to the underlying Lebesgue measure, we can define product densities of a point process.

Another frequent choice is $\mathbb{X} = \mathbb{S}^d$, the d-dimensional unit sphere. The latter case is not discussed here, but we refer the reader e.g. to Cuevas-Pacheco and Møller [2018], Møller et al. [2018] or Møller and Rubak [2016]. For the Borel σ-fields, we use the standard shorter notation $\mathcal{B}^d := \mathcal{B}(\mathbb{R}^d)$, $\mathcal{B} := \mathcal{B}(\mathbb{R})$. In the rest of the thesis, we use the notation $|B|_d$ for the Lebesgue measure of $B \subset \mathbb{R}^d$.

In this section, we extend the list of examples of point processes from the previous section. The examples include cluster point processes, determinantal point processes, Gibbs point processes and Brillinger-mixing point processes.

Definition 1.18 (Stationary random measure).
The random measure Ψ on $\mathbb{R}^d$ is *stationary* if $\Psi \overset{\mathcal{D}}{=} \Psi + x$ for all $x \in \mathbb{R}^d$.

Definition 1.19 (Intensity function, intensity).
Let μ be a point process on $\mathbb{R}^d$ with intensity measure α. If α is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$, then we can write

$$\alpha(B) = \int_B \lambda(x)\mathrm{d}x, \quad B \in \mathcal{B}^d.$$

The function λ is called the *intensity function* of the point process μ. If λ is constant, we talk about the *intensity* of the point process μ.

Remark. If the random measure Ψ is stationary, then its intensity measure α is invariant under translation. The only translation-invariant, locally finite measure on $\mathbb{R}^d$ is, up to the constant, the Lebesgue measure $|\cdot|_d$. Hence, if we assume that α is locally finite, then there is a constant $\lambda \in [0, \infty)$ such that $\alpha(\cdot) = \lambda |\cdot|_d$. The constant λ is called the intensity and it corresponds with Definition 1.19.

Remark. If the point process μ has the intensity function λ, then the Campbell theorem (1.1) reduces to

$$\mathbb{E} \sum_{x \in \mu} f(x) = \int_{\mathbb{R}^d} f(x)\lambda(x)\mathrm{d}x.$$

Definition 1.20 (Product density).
Let $m \in \mathbb{N}$ and let μ be a point process with m-th factorial moment measure $\alpha^{[m]}$. If $\alpha^{[m]}$ is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^{m \times d}$, then the corresponding density is called *m-th order product density*, i.e.

$$\alpha^{[m]}(B_1 \times \cdots \times B_m) = \int_{B_m} \cdots \int_{B_1} \lambda^{[m]}(x_1, \ldots, x_m)\, \mathrm{d}x_1 \cdots \mathrm{d}x_m,$$

where $B_1, \ldots, B_m \in \mathcal{B}^d$.

We interpret the m-th order product density $\lambda^{[m]}(x_1, \ldots, x_m)$ as the probability that exactly one point of the process μ lays in the infinitesimally small ball around $x_i, i = 1, \ldots, m$.

Remark. If μ is stationary, then $\lambda^{[2]}$ depends only on the difference of its arguments, i.e. $\lambda^{[2]}(x_1, x_2) = \lambda_s(x_1 - x_2)$, where λ_s is a suitable function.

Definition 1.21 (Pair correlation function).
Let μ be a point process on $\mathbb{R}^d$ with the product densities of the first two orders $\lambda^{[1]} =: \lambda$ and $\lambda^{[2]}$. Then the function $g : \mathbb{R}^{2d} \to \mathbb{R}$ defined by

$$g(x, y) := \begin{cases} 0, \text{ if } \lambda(x) = 0 \text{ or } \lambda(y) = 0, \\ \frac{\lambda^{[2]}(x,y)}{\lambda(x)\lambda(y)}, \text{ otherwise,} \end{cases}$$

is called the *pair correlation function* of the point process μ.

Example 1.9 (m-th order product density of the Poisson point process).
If the Poisson point process η has the intensity function λ, then its m-th order product density satisfies

$$\lambda^{[m]}(x_1, \ldots, x_m) = \prod_{i=1}^{m} \lambda(x_i).$$

Moreover, the pair correlation function is equal to 1 for any pair of points x, y such that $x \neq y$ and $\lambda(x) > 0, \lambda(y) > 0$.

Examples

In the previous section, we presented the Poisson point process having the property that there are no interactions among the points of the process. On the other hand, cluster point processes can be viewed as point processes, where nearby points attract each other. As the opposite, we present examples of point processes usually used to model repelling data sets. Namely, Gibbs processes and determinantal point processes (DPP's). Briefly, the class of Gibbs point processes is more flexible, although less tractable due to the uniqueness issues. On the other hand, the biggest advantage of DPP is the knowledge of all its moment measures. More involved comparison of these models as well as discussion over their advantages and disadvantages can be seen in Lavancier et al. [2015] and Lavancier et al. [2014]. The link between the models has been studied in Georgii and Yoo [2005].

Example 1.10 (Determinantal point process).
Let μ be a simple point process on $\mathbb{R}^d$ and we assume that its product density functions satisfy

$$\lambda^{[n]}(x_1, \ldots, x_n) = \det\left(C(x_i, x_j)\right)_{1 \leqslant i,j \leqslant n}, \quad n \in \mathbb{N}, (x_1, \ldots, x_n) \in \mathbb{R}^{nd}$$

for some function $C : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$. Then we call μ a *determinantal point process* (DPP) with a kernel C.

Remark. Note that Poisson process is a special case of DPP, where $C(x, y) = 0$ whenever $x \neq y$.

The DPP's were first introduced in Macchi [1975] to model fermions in quantum mechanics, i.e. particles that repel one another. The existence of DPP's is discussed in the latter paper or in Soshnikov [2000].

In fact, it is possible to consider complex-valued function C. Then it would lead to complex-valued joint densities, which is not consistent with the setting of this thesis. We refer to Hough et al. [2009] for more details.

Example 1.11 (Gibbs point process).
The Gibbs point processes cover a wide class of point processes with interaction among the points used largely in statistical physics for modelling systems with a large number of interacting particles. The interaction can be attractive, repulsive or depending on geometrical features. A special case is the Poisson point process showing no interactions. For a general presentation of Gibbs measures, see Georgii [2011]. An intuitive introduction to Gibbs point processes in $\mathbb{R}^d$ can be found in Dereudre [2017]. Another introductory text that also includes the motivation from the statistical mechanics is in Friedli and Velenik [2017].

First, we recall the definition of the Gibbs point process on a bounded set $\Lambda \subset \mathbb{R}^d$. For this reason, denote by $\mathbf{N}_\Lambda$ the space of finite point configurations inside Λ and $\mathcal{N}_\Lambda$ the corresponding trace σ-field. In finite volume, they are defined, intuitively, as modifications of Poisson point processes involving interactions among the points. To model the interactions, we deal with an *energy function* as a measurable function

$$H : \mathbf{N}_f \to \mathbb{R} \cup \{\infty\}$$

which will be assumed to be

- *non-degenerate*, if
$$H(\emptyset) < \infty,$$
- *hereditary*, if for any $\mathbf{x} \in \mathbf{N}_f$ and $x \in \mathbf{x}$
$$H(\mathbf{x}) < +\infty \implies H(\mathbf{x}\backslash\{x\}) < \infty,$$
- *stable*, if there exists a constant $A \geqslant 0$ such that for any $\mathbf{x} \in \mathbf{N}_f$
$$H(\mathbf{x}) \geqslant -A\#(\mathbf{x}),$$
- *invariant under shifts (stationary)*, if for any $\mathbf{x} \in \mathbf{N}_f$ and $x \in \mathbb{R}^d$
$$H(\mathbf{x}) = H(\mathbf{x} + x).$$

Definition 1.22 (Finite volume Gibbs point process).
Let $\Lambda \subset \mathbb{R}^d$ be such that $0 < |\Lambda|_d < \infty$. The *finite volume Gibbs point process* on Λ with activity $z > 0$, inverse temperature $\beta \geqslant 0$ and energy function H is a point process μ having distribution $P_\Lambda^{z,\beta}$ on $\mathbf{N}_\Lambda$ satisfying

$$P_\Lambda^{z,\beta}(\mathrm{d}\mathbf{x}) = \frac{1}{Z_\Lambda^{z,\beta}} e^{-\beta H(\mathbf{x})} P_{\eta_\Lambda^z}(\mathrm{d}\mathbf{x}), \quad \mathbf{x} \in \mathbf{N}_\Lambda, \tag{1.8}$$

where $P_{\eta_\Lambda^z}$ is the distribution of the Poisson point process η_Λ^z in Λ with intensity z and $Z_\Lambda^{z,\beta}$, called the *partition function*, is the normalization constant.

Due to non-degeneracy and stability of H, the partition function $Z_\Lambda^{z,\beta}$ is positive and finite. Hence, $P_\Lambda^{z,\beta}$ is well defined.

The finite volume Gibbs point process is an example of a point process having density p with respect to the distribution of a Poisson point process, where

$$p(\mathbf{x}) = \frac{1}{Z_\Lambda^{z,\beta}} z^{\#(\mathbf{x})} e^{-\beta H(\mathbf{x})}, \quad \mathbf{x} \in \mathbf{N}_\Lambda.$$

Another option is to define the finite volume Gibbs point process using the Papangelou conditional intensity (recall Definition 1.17), which may be easier to understand than the density. As the result of Theorem 1.10, using the conditional intensity eliminates the normalizing constant $Z_\Lambda^{z,\beta}$ needed for the probability density. We then have a finite volume Gibbs process characterized by

$$\lambda^*(x, \mathbf{x}) = z e^{-\beta(H(\mathbf{x} \cup \{x\}) - H(\mathbf{x}))}.$$

The expression in the exponent is called the *local energy* of x in $\mathbf{x}$ and is usually denoted by $h(x, \mathbf{x})$. Note that if $x \in \mathbf{x}$, then $h(x, \mathbf{x}) = 0$.

Let us demonstrate some examples of frequent choices of the energy function.

- **Ising model**: The configuration of the Ising model in a finite discrete set $\Lambda \subset \mathbb{Z}^d$ are elements of the set $\Omega_\Lambda := \{-1, 1\}^\Lambda$. Discrete variables represent magnetic dipole moments of atomic "spins" that can be in one of two states (positive or negative). The spins are arranged in a lattice, allowing each spin to interact with its neighbours. The associated energy of a configuration $\mathbf{x} = \{x_i\}_{i \in \Lambda}$ is defined by

  $$H(\mathbf{x}) = - \sum_{(i,j) \subset \Lambda : i \sim j} x_i x_j - h \sum_{i \in \Lambda} x_i,$$

  where $h \in \mathbb{R}$ is the magnetic field and $i \sim j$ denotes the fact that sites i and j are neighbours. This simplest example does not allow the spins in Λ to interact with other spins located outside Λ. For more options, see Chapter 3 in Friedli and Velenik [2017].

- **Pairwise interaction model**: Let $g : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a Borel measurable function, called the pair potential. We assume that $g(x) = g(-x)$ for any $x \in \mathbb{R}^d$. The pairwise energy function is defined for any $\mathbf{x} \in \mathbf{N}_f$ by

  $$H(\mathbf{x}) = \sum_{\{x,y\} \subset \mathbf{x}} g(x - y).$$

  Often, it is assumed that the pair potential g is a function on $\mathbb{R}_+$ and

  $$H(\mathbf{x}) = \sum_{\{x,y\} \subset \mathbf{x}} g(|x - y|), \quad \mathbf{x} \in \mathbf{N}_f.$$

  See Dereudre [2017] and the reference therein that such energy function H is hereditary, non-degenerate and stable. The choice $g(u) = \mathbf{1}_{[0,r]}(u)$, where r is a given parameter, leads to the *Strauss process*.

- **Energy based on geometrical objects**: The motivation is to provide random configurations such that special geometrical features appear with higher probability than in the case of the original Poisson point process. Here, we show an example of the energy function based on Voronoi tessellation (see the notation in Section 1.5). Let $C(x, \mathbf{x})$ be the Voronoi cell for a configuration $\mathbf{x} \in \mathbf{N}_f$ and $x \in \mathbf{x}$. Then we define the energy function by

$$H(\mathbf{x}) = \sum_{x \in \mathbf{x}} \mathbf{1}_{[C(x,\mathbf{x}) \text{ is bounded}]} \varphi(C(x, \mathbf{x})),$$

  where φ is any function from the space of polytopes in $\mathbb{R}^d$ to $\mathbb{R}$ (e.g. volume, $(d-1)$-dimensional Hausdorff measure of the boundary, number of vertices, etc.).

Of course, the above examples could be freely combined. We can define pair potential which takes into account pairs of Voronoi cells, higher-order interactions, etc.

Theorem 1.12 (GNZ equations).
For any positive measurable function $f : \mathbb{R}^d \times \mathbf{N}_f \to \mathbb{R}$,

$$\int \sum_{x \in \mathbf{x}} f(x, \mathbf{x} \backslash \{x\}) P_\Lambda^{z,\beta}(\mathrm{d}\mathbf{x}) = \int \int_\Lambda f(x, \mathbf{x}) \lambda^*(x, \mathbf{x}) \mathrm{d}x P_\Lambda^{z,\beta}(\mathrm{d}\mathbf{x}).$$

The GNZ equations have been first introduced by Georgii, Nguyen and Zessin (cf. Georgii [1976], Nguyen and Zessin [1979]). Note that they generalize the Slivnyak–Mecke formulas for Poisson point processes (Theorem 1.9).

The finite volume Gibbs model can be defined by writing down its probability density which is no longer possible in the infinite volume case. Also to define the energy of an infinite configuration $\mathbf{x}$ is meaningless. See Section 6.1 in Friedli and Velenik [2017] for an explanation. The main idea is to define a sequence of finite Gibbs point processes $(P_{\Lambda_n}^{z,\beta})_{n \geqslant 1}$ on $\Lambda_n = [-n, n]^d$ and let $n \to \infty$. Then we extract a convergent subsequence and call its accumulation point the infinite volume Gibbs point process.

Definition 1.23 (Local functions).
The function $f : \mathbf{N} \to \mathbb{R}$ is said to be *local* if there exists a bounded set $\Delta \subset \mathbb{R}^d$ such that for all $\mathbf{x} \in \mathbf{N}$, $f(\mathbf{x}) = f(\mathbf{x}_\Delta)$. By $\mathbf{x}_\Delta$ we denote the configuration of $\mathbf{x}$ inside Δ.

Definition 1.24 (Local convergence topology).
The *local convergence topology* on the space of probability measures on $\mathcal{N}$ is the smallest topology such that for any local bounded function $f : \mathbf{N} \to \mathbb{R}$ the function $P \to \int f \mathrm{d}P$ is continuous.

Take the sequence $\Lambda_n = [-n, n]^d$, the finite volume Gibbs processes $(P_{\Lambda_n}^{z,\beta})_{n \geqslant 1}$ given by (1.8) for common parameters $z > 0, \beta \geqslant 0$ and energy function H assumed to be stationary. We use a stationarization to define measures $\bar{P}_{\Lambda_n}^{z,\beta}, n \geqslant 1$, i.e.

$$\int f(\mathbf{x}) \bar{P}_{\Lambda_n}^{z,\beta}(\mathrm{d}\mathbf{x}) = \frac{1}{(2n)^d} \int_{[-n,n]^d} \int f(\mathbf{x} - u) \bar{P}_{\Lambda_n}^{z,\beta}(\mathrm{d}\mathbf{x}) \mathrm{d}u, \quad \text{for all } f : \mathbf{N} \to \mathbb{R}.$$

We interpret $\bar{P}^{z,\beta}_{\Lambda_n}$ as the Gibbs measure where the origin $\mathbf{o}$ is replaced by a random point chosen uniformly in Λ_n.

Theorem 1.13 (Proposition 9 in Dereudre [2017]).
The sequence $(\bar{P}^{z,\beta}_{\Lambda_n})_{n\geqslant 1}$ is tight for the local convergence topology.

Definition 1.25 (Infinite volume Gibbs point process).
Let us denote by $P^{z,\beta}$ one of the accumulation points of the sequence $(\bar{P}^{z,\beta}_{\Lambda_n})_{n\geqslant 1}$. We call the measure $P^{z,\beta}$ the *infinite volume Gibbs point process*.

For intuitive, yet more rigorous explanation of the transition between the finite volume and infinite volume cases, see Chapter 6 in Friedli and Velenik [2017] or Section 2 in Dereudre [2017]. It is possible to characterize the infinite volume Gibbs point process using the GNZ equation similarly as in the finite volume setting. We need a further assumption on the energy function H.

Definition 1.26 (Finite range energy function).
The energy function H has a *finite range (or is finite range)* with $R > 0$ if for every bounded set $\Delta \subset \mathbb{R}^d$ the local energy $H_\Delta(\cdot) := H(\cdot) - H(\cdot_{\Delta^c})$ satisfies for every finite configuration $\mathbf{x} \in \mathbf{N}_f$

$$H_\Delta(\mathbf{x}) = H(\mathbf{x}_{\Delta \oplus B(\mathbf{o},R)}) - H(\mathbf{x}_{\Delta \oplus B(\mathbf{o},R) \setminus \Delta^c}).$$

By $\oplus$, we denote the Minkowski sum.

Theorem 1.14 (GNZ equations in infinite volume case).
Let P be a probability measure on $\mathcal{N}$. Let H be a finite range energy functions and $z > 0, \beta \geqslant 0$ be two parameters. Then P is the infinite volume Gibbs measure with the energy function H, activity z and inverse temperature β if and only if for any positive measurable function $f : \mathbb{R}^d \times \mathbf{N} \to \mathbb{R}$

$$\int \sum_{x \in \mathbf{x}} f(x, \mathbf{x} \setminus \{x\}) P(\mathrm{d}\mathbf{x}) = \int \int_{\mathbb{R}^d} \lambda^*(x, \mathbf{x}) f(x, \mathbf{x}) \mathrm{d}x P(\mathrm{d}\mathbf{x}), \tag{1.9}$$

where the function $\lambda^*(x, \mathbf{x}) = z e^{-\beta h(x,\mathbf{x})}, x \in \mathbb{R}^d, \mathbf{x} \in \mathbf{N}$, is the Papangelou conditional intensity of P.

For conditions ensuring that (1.9) holds, see Ruelle [1969]. It is important to emphasize that (1.9) may not give a unique solution. The uniqueness and the existence of the solution is still an open problem for a large spectrum of Gibbs models. We refer to Dereudre et al. [2010] for several existence theorems. There exist techniques in the literature used in order to obtain uniqueness of Gibbs measures, namely the Dobrushin criterion (Dobruschin [1968]), cluster expansion (Ruelle [1969] or Jansen [2019]), and disagreement percolation (Hofer-Temmel and Houdebert [2019]).

Example 1.12 (Brillinger-mixing point process).
The assumption of stationarity allows us to disintegrate the moment measures and cumulant measures of higher orders. This disintegration then leads to the

definition of reduced versions of the measures associated to μ. For example, for the k-th factorial cumulant measure, we have

$$\gamma^{[k]}(B_1 \times \cdots \times B_k) = \lambda \int_{B_1} \gamma^{[k]}_{red}((B_2 - x) \times \cdots \times (B_k - x))\, \mathrm{d}x, \tag{1.10}$$

where $B_1, \ldots, B_k \in \mathcal{B}^d$ and $B_j - x$ is the translation of the set B_j by $x \in \mathbb{R}^d$.

Definition 1.27 (Reduced factorial cumulant measure).
The measure $\gamma^{[k]}_{red}$ on $\mathbb{R}^{d(k-1)}$ from the expression (1.10) is called *reduced k-th factorial moment measure.*

Remark. The reduced cumulant measures of the second and higher orders are signed measures, hence admit the Hahn–Jordan decomposition (Dudley [2002], Theorem 5.6.1). Therefore, for any $k \geqslant 2$, there exist two measures $\gamma^{[k]}_{red,+}$ and $\gamma^{[k]}_{red,-}$ uniquely determined by $\gamma^{[k]}_{red}$ such that

$$\gamma^{[k]}_{red} = \gamma^{[k]}_{red,+} - \gamma^{[k]}_{red,-}.$$

The total variation measure of $\gamma^{[k]}_{red}$ is then

$$|\gamma^{[k]}_{red}| = \gamma^{[k]}_{red,+} + \gamma^{[k]}_{red,-}$$

and the total variation is $\|\gamma^{[k]}_{red}\|_{TV} := |\gamma^{[k]}_{red}|(\mathbb{R}^{d(k-1)})$.

Reduced factorial cumulant measures are the basis of the Brillinger-mixing property.

Definition 1.28 (Brillinger-mixing process).
A stationary point process μ is Brillinger-mixing if, for $k \geqslant 2$, we have

$$\|\gamma^{[k]}_{red}\|_{TV} < \infty. \tag{1.11}$$

The condition (1.11) expresses some kind of weak correlatedness (or asymptotic uncorrelatedness) of the numbers of points lying in bounded sets having a large (or unboundedly increasing) distance of one another. This type of weak dependence does not necessarily imply ergodicity, see Heinrich [2018], but allows to prove central limit theorems for various stochastic models related with point processes, e.g. in stochastic geometry, statistical physics for $d \geqslant 1$ or in queueing theory for $d = 1$, see e.g. Heinrich and Schmidt [1985]. Brillinger-mixing processes cover a wide class of processes including e.g.

- Neyman–Scott process: if the distribution of the number of points in each cluster have all moments finite, then it is Brillinger-mixing. It remains true for any Poisson-cluster process (Example 4.1 in Heinrich [2013]),
- Cox process: Under some more involved assumptions, the Cox process is also Brillinger-mixing (Heinrich [1988]).
- Determinantal point process: Recently in Biscio and Lavancier [2016], it has been shown that also DPP's are Brillinger-mixing if its kernel C is symmetric continuous real-valued function in $L^2(\mathbb{R}^d)$ with $C(0) = \rho$ and the Fourier transform of C has values in $[0, 1]$. The latter assumption guarantees the existence of such DPP.

1.3 Marked point processes

In many applications, we observe a further random element M_i assigned to each point x_i of the point pattern. The mark M_i carries some additional information and takes values in some mark space $\mathbb{M}$. The class of processes considered in this section includes processes of objects that are characterized by their location and weight. Such processes are covered formally by the general theory, as they can be viewed as point processes on a product space. Nevertheless, marked point processes deserve attention on their own due to their importance in applications.

We will restrict this section to the case when the unmarked point process is a process on $\mathbb{R}^d$ and leave the space $\mathbb{M}$ to be a complete separable locally compact metric space equipped with a σ-field $\mathcal{M}$.

Definition 1.29 (Marked point process).
By a *marked point process* we understand a point process μ_m on $\mathbb{R}^d \times \mathbb{M}$ such that

$$\alpha_m(B \times \mathbb{M}) := \mathbb{E}\, \mu_m(B \times \mathbb{M}) < \infty$$

for all $B \in \mathcal{B}_b^d := \mathcal{B}_b(\mathbb{R}^d)$. The point process μ given by $\mu(B) = \mu_m(B \times \mathbb{M})$ is called the *unmarked point process*. The measure α_m is the *intensity measure* of the marked point process μ_m.

Definition 1.30 (Simple marked point process).
We say that a marked point process μ_m is *simple* if the unmarked point process μ is simple.

The Campbell theorem (see expression 1.1) also holds:

$$\mathbb{E} \sum_{(x,m)\in\mu_m} f(x,m) = \int f(x,m)\alpha_m(\mathrm{d}(x,m))$$

for any non-negative measurable function f on $\mathbb{R}^d \times \mathbb{M}$.

It is clear from Definition 1.29 that every marked point process is a point process on the product space $\mathbb{R}^d \times \mathbb{M}$. Nevertheless, not every point process on $\mathbb{R}^d \times \mathbb{M}$ can be viewed as a marked point process. For example, if $\mathbb{M} = \mathbb{R}$ and we take a stationary Poisson point process on $\mathbb{R}^d \times \mathbb{M} = \mathbb{R}^{d+1}$ with positive intensity λ, then

$$\alpha_m(B \times \mathbb{R}) = \lambda\, |B \times \mathbb{R}|_{d+1} = \infty$$

for arbitrary $B \in \mathcal{B}_b^d$ with $|B|_d > 0$.

Example 1.13 (Marked Poisson point process).
Let α_m be a locally finite non-atomic measure on $\mathbb{R}^d \times \mathbb{M}$. The marked Poisson point process on $\mathbb{R}^d$ with marks in $\mathbb{M}$ and intensity measure α_m is the point process η_m satisfying the following conditions:

1. $\eta_m(B_1 \times L_1), \ldots, \eta_m(B_k \times L_k)$ are mutually independent random variables for pairwise disjoint $B_j \times L_j \in \mathcal{B}_b^d \times \mathcal{B}(\mathbb{M})$ and $k \in \mathbb{N}$,

2. $\eta_m(B \times L)$ is Poisson distributed with mean $\alpha_m(B \times L)$ for any $B \in \mathcal{B}_b^d, L \in \mathcal{B}(\mathbb{M})$.

Remark. A marked Poisson point process η_m with the intensity measure α_m is simple if and only if $\alpha_m(\{x\} \times \mathbb{M}) = 0$ for all $x \in \mathbb{R}^d$.

Definition 1.31 (Independently marked point process).
A marked point process $\mu_m = \sum_i \delta_{(x_i, M_i)}$ is called *independently marked* if $\{M_i\}$ are identically distributed, mutually independent and independent of the unmarked point process $\mu = \sum_i \delta_{x_i}$.

Definition 1.32 (Stationary marked point process).
A marked point process μ on $\mathbb{R}^d$ with marks in $\mathbb{M}$ is *stationary* if its distribution is invariant under shifts of $\mathbb{R}^d$ only, i.e. $(x, m) \mapsto (x + v, m)$ for all $v \in \mathbb{R}^d$.

Theorem 1.15 (Theorem 2.3 in Baddeley [2007]).
Let μ be a stationary marked point process in $\mathbb{R}^d$ such that the corresponding unmarked point process of unmarked points has finite intensity λ. Then the intensity measure α of μ takes the form

$$\alpha(A \times B) = \lambda \, |A|_d \, Q(B)$$

for all $A \in \mathcal{B}, B \in \mathcal{B}(\mathbb{M})$. The probability measure Q on $\mathbb{M}$ is called the distribution of the typical mark.

Remark. (Campbell's theorem for a stationary marked point process) If μ is a stationary marked point process on $\mathbb{R}^d$ with intensity λ, then

$$\mathbb{E} \sum_{(x,m) \in \mu} f(x, m) = \lambda \mathbb{E}_Q \int_{\mathbb{R}^d} f(x, M) \, \mathrm{d}x,$$

where M is a random variable distributed according to Q and $\mathbb{E}_Q$ is the expectation w.r.t Q.

Example 1.14 (Germ-grain model, Boolean model).
A stationary independently marked point process $\mu_m = \{(x_i, \Xi_i), i \geqslant 1\}$ on $\mathbb{R}^d$ with the mark space $\mathbb{M} = \mathcal{C}^{(d)}$ (the space of all non-empty compact sets in $\mathbb{R}^d$ equipped with the Hausdorff metric, see Section 1.4) is called *germ-grain process.* The associated random set

$$\Xi = \bigcup_{i \geqslant 1} (x_i + \Xi_i)$$

is called *germ-grain model.* If the unmarked point process is the Poisson point process, we call the random set Ξ *Boolean model.*

Example 1.15 (Cylinder process in the plane).
Let $g(p, \varphi) := \{(x, y) \in \mathbb{R}^2 : x \cos \varphi + y \sin \varphi = p\}$ be a parametrized line (Hessian normal form), where $p \in \mathbb{R}^1$ stands for the signed distance of the line from the origin $\mathbf{o}$ and $\varphi \in [0, \pi)$ is the angle (measured anti-clockwise) between the normal vector $v(\varphi) = (\cos \varphi, \sin \varphi)^T$ on the line (with direction in the half-plane not containing $\mathbf{o}$ and the x_1-axis).

We describe a cylinder process in $\mathbb{R}^2$ (see Figure 1.1) in terms of its generating stationary, independently marked point process on $\mathbb{R}^1$. For doing this, let (Φ_0, R_0) be the generic random vector taking value in the mark space $[0, \pi] \times [0, \infty)$ that describes the orientation Φ_0 and the cross-section (or base) $\Xi_0 := [-R_0, R_0]$ of the typical cylinder. In addition, we assume that Φ_0 and R_0 are independent with distribution functions G and F, respectively, i.e. $\mathbb{P}(R_0 \leqslant r, \Phi_0 \leqslant \varphi) = F(r)\, G(\varphi)$.

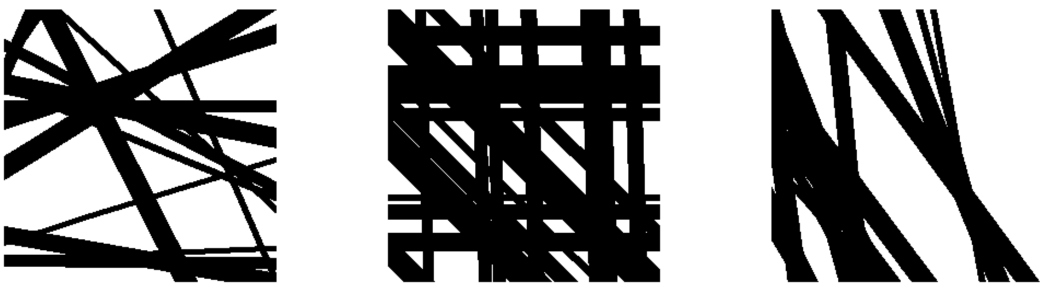
Now we introduce a stationary independently marked point process as locally finite, simple counting measure $\Psi^P_{F,G} := \sum_{i\in\mathbb{Z}} \delta_{[P_i,(\Phi_i,R_i)]}$ defined on the Borel sets of $\mathbb{R}^1 \times [0,\pi] \times [0,\infty)$, whose finite-dimensional distributions are shift-invariant in the first component. The stationary unmarked (or ground) point process $\Psi = \sum_{i\in\mathbb{Z}} \delta_{P_i} \sim P$ with finite and positive intensity $\lambda = \mathbb{E}\,\Psi([0,1]) > 0$ is assumed to be independent of the i.i.d. sequence $\{(\Phi_i, R_i) : i \in \mathbb{Z} := \{0, \pm 1, \pm 2, \ldots\}\}$ of mark vectors. Each triplet $[P_i,(\Phi_i,R_i)], i \in \mathbb{Z}$, determines a random cylinder $g(P_i,\Phi_i) \oplus b(\mathbf{o}, R_i)$, where $b(\mathbf{o}, r) = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leqslant r^2\}$ is the circle in $\mathbb{R}^2$ with radius $r \geqslant 0$ and centre in the origin $\mathbf{o}$ and $\oplus$ stands for Minkowski sum of subsets of $\mathbb{R}^2$. The intensity measure $\Lambda_{F,G}((\cdot) \times [0,\varphi] \times [0,r]) := \mathbb{E}\,\Psi^P_{F,G}((\cdot) \times [0,\varphi] \times [0,r])$ can be expressed for $r \geqslant 0, 0 \leqslant \varphi \leqslant \pi$ as

$$\Lambda_{F,G}((\cdot) \times [0,\varphi] \times [0,r]) = \mathbb{E}\,\Psi(\cdot)\,\mathbb{P}(\Phi_0 \leqslant \varphi, R_0 \leqslant r) = \lambda\,|\cdot|_1\, G(\varphi)\, F(r).$$

Definition 1.33 (Cylinder process in $\mathbb{R}^2$).
By a *cylinder process* $\Xi := \Xi^P_{F,G}$ in the Euclidean plane $\mathbb{R}^2$ derived from the stationary independently marked point process $\Psi^P_{F,G}$, we understand the random union set defined by

$$\Xi := \bigcup_{i\in\mathbb{Z}} \left(g(P_i,\Phi_i) \oplus b(\mathbf{o}, R_i)\right). \tag{1.12}$$

Note that Ξ forms a random set in $\mathbb{R}^2$ which in general is neither closed nor stationary. Cylinder processes have numerous applications (mostly for $d = 2, 3$) among others in material sciences to model materials consisting of long thick fibres or thick membranes, see e.g. Spiess and Spodarev [2011]. Generally, cylinder processes in $\mathbb{R}^d$ are defined as countable unions of dilated affine subspaces of $\mathbb{R}^k$, $k = 1, \ldots, d-1$ see e.g. Weil [1987], Schneider and Weil [2008] or Molchanov [2005].



(a) (b) (c)

Figure 1.1: Three realizations of the cylinder process in the planar unit window. All three cases are generated by a marked point process with uniformly distributed widths. On picture (a), the unmarked point process is a stationary Poisson process with the orientations distributed uniformly in $[0,\pi)$. On picture (b), the unmarked process is a stationary Poisson process and the orientations take three different values with the same probability. On picture (c), the unmarked process is a Thomas point process with the distributions being uniform in $[0,\pi/4)$.

1.4 Particle processes

In this section, we present a special case of point processes, namely the particle processes. Those are point processes in the space of all non-empty compact subsets of $\mathbb{R}^d$. The definitions and results listed in this section are mostly based on Section 4.1 in Schneider and Weil [2008].

To work with particles, we use the following notation.

Notation. For a compact set $K \subset \mathbb{R}^d$, we denote by $B(K)$ the circumscribed ball (circumball) of K and by $c(K)$ the centre (circumcenter) of $B(K)$. We define spaces of particles

$\mathcal{C}^d$	...	set of all compact subsets (particles) in $\mathbb{R}^d$,
$\mathcal{C}^{(d)}$	...	set of all non-empty compact subsets in $\mathbb{R}^d$,
$\mathcal{C}^{(d)}_{\mathbf{o}}$	...	$\{K \in \mathcal{C}^{(d)} : c(K) = \mathbf{o}\}$, where $\mathbf{o}$ is the origin in $\mathbb{R}^d$,
$\mathcal{C}^{(d)}_{\Lambda}$	...	$\{K \in \mathcal{C}^{(d)} : c(K) \in \Lambda\}$, where $\Lambda \subset \mathbb{R}^d$.

Moreover, for $K,\ L \in \mathcal{C}^{(d)}$ we denote by $K \oplus L$ and $\check{K}$ the operations

$$K \oplus L := \{x + y : x \in K,\ y \in L\},$$
$$\check{K} := \{-x : x \in K\}.$$

We equip the space $\mathcal{C}^d$ with the Borel σ-field $\mathcal{B}(\mathcal{C}^d)$ generated by the Fell topology on the space of closed subsets of $\mathbb{R}^d$ restricted to $\mathcal{C}^d$. Moreover, let the space $\mathcal{C}^{(d)}$ be equipped with the Hausdorff metric. Recall that the Hausdorff metric ρ_H on $\mathcal{C}^{(d)}$ is defined for $K, L \in \mathcal{C}^{(d)}$ by

$$\rho_H(K, L) := \max\left\{\sup_{x\in K}\inf_{y\in L} ||x-y||, \sup_{y\in L}\inf_{x\in K} ||x-y||\right\},$$

where $||\cdot||$ denotes the Euclidean distance. The corresponding Borel σ-field $\mathcal{B}(\mathcal{C}^{(d)})$ is a trace of $\mathcal{B}(\mathcal{C}^d)$, (cf. Theorem 2.4.1 in Schneider and Weil [2008]). It can be shown that $\mathcal{C}^{(d)}$ is a Polish space (cf. Theorem A.26 in Last and Penrose [2017]).

In order to define the particle process, we need to specify the outcome space (the space of all particle configurations). It is defined as follows:

- By $\mathbf{N}^d$ we denote the space of all locally finite subsets $\mathbf{x}$ on $\mathcal{C}^{(d)}$, meaning that the cardinality

  $$\#\{L \in \mathbf{x} : L \cap K \neq \emptyset\} < \infty$$

  for all $K \in \mathcal{C}^{(d)}$. The space $\mathbf{N}^d$ is equipped with the σ-field

  $$\mathcal{N}^d = \sigma(\{\mathbf{x} \in \mathbf{N}^d : \#\{K \in \mathbf{x} : K \in B\} = m\},\ B \in \mathcal{B}_b(\mathcal{C}^{(d)}), m \in \mathbb{N}).$$

- For $\Lambda \subset \mathbb{R}^d$ denote by $\mathbf{N}^d_\Lambda$ the system of finite subsets of $\mathcal{C}^{(d)}_\Lambda$ equipped with the trace σ-field $\mathcal{N}(\mathcal{C}^{(d)}_\Lambda)$.

- By $\mathbf{N}^d_f$, we denote the subsystem of $\mathbf{N}^d$ consisting of finite sets.

Definition 1.34 (Particle process).
A *particle process* is a point process on $\mathcal{C}^{(d)}$, i.e. a random element

$$\xi : (\Omega, \mathcal{F}, \mathbb{P}) \longrightarrow (\mathbf{N}^d, \mathcal{N}^d).$$

Its distribution is $P_\xi = \mathbb{P}\xi^{-1}$.

A particle process can also be viewed as a marked point process with mark space $\mathcal{C}^{(d)}$, where the points of the unmarked point process set the location of the particle (using for example the circumcenter).

Definition 1.35 (Intensity measure of a particle process).
The *intensity measure* of a particle process ξ is a Borel measure α on $\mathcal{C}^{(d)}$ defined by $\alpha(B) = \mathbb{E}\,\xi(B), B \in \mathcal{B}(\mathcal{C}^{(d)})$.

Recall that in Section 1.1, we assumed that all intensity measures are locally finite. In the case of particle processes, this assumption transforms to boundedness in the sense

$$\alpha(\{K \in \mathcal{C}^{(d)} : K \cap B \neq \emptyset\}) < \infty, \quad \text{for all } B \in \mathcal{C}^{(d)}. \tag{1.13}$$

Definition 1.36 (Stationary particle process).
A particle process ξ is called *stationary* if $P_{\xi+x} = P_\xi$ for each $x \in \mathbb{R}^d$, where for any $\mathbf{x} \in \mathbf{N}(\mathcal{C}^{(d)})$ we set

$$\mathbf{x} + x = \{K + x \ : K \in \mathbf{x}\},\ K + x \ = \{y + x : y \in K\}.$$

If a particle process ξ is stationary and its intensity measure satisfies (1.13), then it can be decomposed, so that there exist $\beta > 0$ and a probability measure $\mathbb{Q}$ on $\mathcal{C}_{\mathbf{o}}^{(d)}$ such that for all non-negative measurable functions f on $\mathcal{C}^{(d)}$ it holds

$$\int_{\mathcal{C}^{(d)}} f(K)\alpha(\mathrm{d}K) = \beta \int_{\mathbb{R}^d} \int_{\mathcal{C}_{\mathbf{o}}^{(d)}} f(z + K)\mathbb{Q}(\mathrm{d}K)\mathrm{d}z.$$

See Lemma 11.5 in Rataj [2006] for the proof. The constant β is called the intensity of the process ξ and $\mathbb{Q}$ is the reference particle distribution. A random set with distribution $\mathbb{Q}$ is called the typical grain of ξ. Note that the reference particle distribution satisfies

$$\mathbb{Q}(\{K \in \mathcal{C}^{(d)} : c(K) = \mathbf{o}\}) = 1,$$

where $c(K)$ is the centre of the circumscribed ball $B(K)$ of K.

A reference measure of a stationary particle process with reference distribution $\mathbb{Q}$ is a measure λ on $\mathcal{C}^{(d)}$ defined by

$$\lambda(B) = \int_{\mathcal{C}^{(d)}} \int_{\mathbb{R}^d} \mathbf{1}_{[K+x\in B]}\,\mathrm{d}x\,\mathbb{Q}(\mathrm{d}K),\ B \in \mathcal{B}(\mathcal{C}^{(d)}), \tag{1.14}$$

Note that the measure λ is invariant under shifts, i.e. $\lambda(B) = \lambda(B + x)$, $x \in \mathbb{R}^d$.

Example 1.16 (Gibbs particle process).
We construct the Gibbs particle process by the same procedure we used in Section 1.2 starting from a finite volume Gibbs point process defined by probability density. Then, by using the GNZ formula we define the infinite volume Gibbs particle process.

Finite volume Gibbs particle process

Let $H : \mathbf{N}_f^d \to \mathbb{R}_+ \cup \{\infty\}$ be the energy function, i.e. a measurable function that is non-degenerate, hereditary, invariant under shifts and stable (see Section 1.2 for the notion).

In the following we consider a bounded set $\Lambda \subset \mathbb{R}^d$ with $|\Lambda|_d > 0$. Further, let

$$\lambda_\Lambda(B) = \int_{\mathcal{C}^{(d)}} \int_\Lambda \mathbf{1}_{[K+x \in B]} \mathrm{d}x \, \mathbb{Q}(\mathrm{d}K), \ B \in \mathcal{B}(\mathcal{C}_\Lambda^{(d)})$$

and η_Λ be the Poisson process on $\mathcal{C}_\Lambda^{(d)}$ with intensity measure λ_Λ and distribution π_Λ. We define a finite volume Gibbs particle process μ_Λ on Λ with activity $z > 0$, inverse temperature $\beta \geqslant 0$ and energy function H as a particle process with distribution $P_\Lambda^{z,\beta}$ on $\mathbf{N}_\Lambda^d$ given by the Radon–Nikodym density p with respect to π_Λ, where

$$p(\mathbf{x}) = \frac{1}{Z_\Lambda^{z,\beta}} z^{N_\Lambda(\mathbf{x})} \exp(-\beta H(\mathbf{x})), \ \mathbf{x} \in \mathbf{N}_\Lambda^d,$$

$N_\Lambda(\mathbf{x})$ is the number of particles $K \in \mathbf{x}$ with $c(K) \in \Lambda$ and

$$Z_\Lambda^{z,\beta} = \int_{\mathbf{N}_\Lambda^d} z^{N_\Lambda(\mathbf{x})} \exp(-\beta H(\mathbf{x})) \pi_\Lambda(\mathrm{d}\mathbf{x})$$

is the normalizing constant.

The local energy $h : \mathcal{C}^{(d)} \times \mathbf{N}_f^d \to \mathbb{R}$ is defined as

$$h(K, \mathbf{x}) = H(\mathbf{x} \cup \{K\}) - H(\mathbf{x}).$$

The Georgii–Nguyen–Zessin (GNZ) equations follow for any measurable function $f : \mathcal{C}^{(d)} \times \mathbf{N}_f^d \longrightarrow \mathbb{R}_+$:

$$\int\limits_{\mathbf{N}_f^d} \sum_{K \in \mathbf{x}} f(K, \mathbf{x} \backslash \{K\}) P_\Lambda^{z,\beta}(\mathrm{d}\mathbf{x}) = z \int\limits_{\mathbf{N}_f^d} \int\limits_{\mathcal{C}_\Lambda^{(d)}} f(K, \mathbf{x}) e^{-\beta h(K, \mathbf{x})} \lambda_\Lambda(\mathrm{d}K) P_\Lambda^{z,\beta}(\mathrm{d}\mathbf{x}). \quad (1.15)$$

The GNZ equations characterize the finite volume Gibbs particle process, i.e. if any probability measure on $\mathbf{N}_\Lambda^d$ satisfies (1.15) for any f as stated, then it is equal to $P_\Lambda^{z,\beta}$. The function

$$\lambda^*(K, \mathbf{x}) = z \exp(-\beta h(K, \mathbf{x})), \ K \in \mathcal{C}_\Lambda^{(d)}, \ \mathbf{x} \in \mathbf{N}_\Lambda^d$$

is the Papangelou conditional intensity.

Infinite volume Gibbs particle process

Take a sequence of windows $\Lambda_n = [-n, n]^d \subset \mathbb{R}^d$, spaces $\mathcal{C}_{\Lambda_n}^{(d)}$, intensity measures $\lambda_n(\cdot) = \int\int_{\Lambda_n} \mathbf{1}_{[K+x \in \cdot]} \mathrm{d}x \mathbb{Q}(\mathrm{d}K)$ for a fixed probability measure $\mathbb{Q}$ on $\mathcal{C}_\mathbf{o}^d$ such that there is some $R > 0$ so that

$$\mathbb{Q}(\{K \in \mathcal{C}^{(d)} : B(K) \subset B(\mathbf{o}, R)\}) = 1. \quad (1.16)$$

Further, take Poisson particle processes η_{Λ_n} with distributions π_{Λ_n}, Gibbs particle processes μ_{Λ_n} with distributions $P_{\Lambda_n}^{z,\beta}$, $n \in \mathbb{N}$.

A measurable function $f : \mathbf{N}^d \longrightarrow \mathbb{R}$ is called *local* if there is a bounded set $\Delta \subset \mathbb{R}^d$ such that for all $\mathbf{x} \in \mathbf{N}^d$ we have $f(\mathbf{x}) = f(\mathbf{x}_\Delta)$. The *local convergence topology* on the space of probability measures P on $\mathbf{N}^d$ is the smallest topology such that for any local and bounded function $f : \mathbf{N}^d \longrightarrow \mathbb{R}$ the map $P \mapsto \int f \mathrm{d}P$ is continuous.

Define a probability measure $\bar{P}^{z,\beta}_{\Lambda_n}$ such that for any $n \geqslant 1$ and any measurable function $f : \mathbf{N}^d \longrightarrow \mathbb{R}$ it holds

$$\int_{\mathbf{N}^d} f(\mathbf{x}) \bar{P}^{z,\beta}_{\Lambda_n}(\mathrm{d}\mathbf{x}) = (2n)^{-d} \int_{\Lambda_n} \int_{\mathbf{N}^d} f(\mathbf{x} + u) P^{z,\beta}_{\Lambda_n}(\mathrm{d}\mathbf{x}) \mathrm{d}u. \tag{1.17}$$

It can be shown that the sequence $(\bar{P}^{z,\beta}_{\Lambda_n})_{n \geqslant 1}$ is tight for the local convergence topology (cf. Chapter 15 in Georgii [2011]). We denote $P^{z,\beta}$ one of its cluster points.

Due to the stationarization (1.17), $P^{z,\beta}$ is the distribution of a stationary particle process, in order to show that it satisfies the GNZ equations one needs to add an assumption on the energy function. The energy function H has a finite range $r > 0$ if for every bounded set $\Delta \subset \mathbb{R}^d$ the energy H_Δ is a local function on $\Delta \oplus B(0, r)$. The finite range property allows extending of the domain of H_Δ from the space $\mathbf{N}^d_f$ to $\mathbf{N}^d$, since for $\mathbf{x} \in \mathbf{N}^d$ we put

$$H_\Delta(\mathbf{x}) = H_\Delta(\mathbf{x}_{\Delta \oplus B(0,r)}).$$

Lemma 1.1 (Proposition 2.1 in Flimmel and Beneš [2018]).
If the energy function H is non-degenerate, hereditary, stable, invariant under shifts and has a finite range property, then there exists an infinite volume stationary Gibbs particle process $P^{z,\beta}$ with the energy function H.

The stationary Gibbs particle process also satisfies GNZ equations for any measurable function $f : \mathcal{C}^{(d)} \times \mathbf{N}^d \longrightarrow \mathbb{R}_+$:

$$\int_{\mathbf{N}^d} \sum_{K \in \mathbf{x}} f(K, \mathbf{x} \backslash \{K\}) P^{z,\beta}(\mathrm{d}\mathbf{x}) = \int_{\mathbf{N}^d} \int_{\mathcal{C}^{(d)}} f(K, \mathbf{x}) \lambda^*(K, \mathbf{x}) \lambda(\mathrm{d}K) P^{z,\beta}(\mathrm{d}\mathbf{x}), \tag{1.18}$$

where λ comes from (1.14). Conversely, any measure P on $\mathbf{N}^d$ which satisfies (1.18) is a distribution of a stationary Gibbs particle process.

Pairwise interactions

Assume that the energy function is of the form

$$H(\mathbf{x}) = \sum^{\neq}_{\{K,L\} \subset \mathbf{x}} g(K \cap L), \quad \mathbf{x} \in \mathbf{N}^d_f, \tag{1.19}$$

where $g : \mathcal{C}^d \longrightarrow \mathbb{R}_+$ is a measurable function, we assume that it is invariant under shifts and $g(\emptyset) = 0$. The expression $g(K \cap L)$ in (1.19) plays a role of pair potential. Such energy function is non-degenerate, hereditary, stationary and stable. If we restrict ourselves to bounded particles $K \in \mathcal{C}^{(d)} : B(K) \subset B(0, R)$ for some $R > 0$, then H has finite range $r = 2R$.

The corresponding conditional intensity is of the form

$$\lambda^*(K, \mathbf{x}) := z \exp\left\{-\beta \sum_{L \in \mathbf{x}} g(K \cap L)\right\}, \quad K \in \mathcal{C}^{(d)},\ \mathbf{x} \in \mathbf{N}^d, \tag{1.20}$$

where $z > 0$, $\beta \geqslant 0$.

Specially, take $\mathbb{Q}$ being concentrated on the set $S_{\mathbf{o}}^R \subset \mathcal{C}^{(2)}$ (the space of all segments in $\mathbb{R}^2 \cap B(\mathbf{o}, R)$ centered in the origin), which corresponds to $\mathbb{Q}_\phi \otimes \mathbb{Q}_L$, where $\mathbb{Q}_\phi$, $\mathbb{Q}_L$ is the reference distribution of directions, lengths of segments, respectively. Further, let g have a form

$$g(K) = \mathbf{1}\{K \neq \emptyset\}, \quad K \in \mathcal{C}^2.$$

Then we call the corresponding stationary Gibbs process a Gibbs segment process. The conditional intensity is

$$\lambda^*(K, \mathbf{x}) = z e^{-\beta N_{\mathbf{x}}(K)}, \quad K \in S_{\mathbf{o}}^R, \mathbf{x} \in \mathbf{N}^2,$$

where $N_{\mathbf{x}}(K)$ denotes the number of intersections of K with the segments in $\mathbf{x}$. It has to be mentioned that the reference distribution $\mathbb{Q}$ need not coincide with the observed joint length-direction distribution of the process, cf. Beneš et al. [2019].

1.5 Random tessellations

Random tessellations are an important model in stochastic geometry (cf. Chiu et al. [2013] or Schneider and Weil [2008]) and they have numerous applications in engineering and the natural sciences (cf. Okabe et al. [2000]).

By a tessellation we understand a subdivision of the space $\mathbb{R}^d = \bigcup C_i$ into d-dimensional sets C_i with no common interior points. Such geometrical patterns can be observed in many natural situations, such as polycrystalline materials, foam structures, etc. Hence, random tessellation models have been widely used in physics, materials science and chemistry. Depending on the situation, the sets C_i might be called cells, crystals, regions, etc.

Definition 1.37 (Tessellation).
A *tessellation* in $\mathbb{R}^d$ is a countable system T of subsets of $\mathbb{R}^d$ (cells) satisfying the following conditions:

1. T is a locally finite system of non-empty closed sets meaning that

$$\sum_{C \in T} \mathbf{1}_{[C \cap B \neq \emptyset]} < \infty, \quad \text{for all } B \in \mathcal{B}_b^d.$$

2. The cells $C \in T$ are compact, convex and have interior points.

3. The sets of T cover the whole space, i.e.

$$\bigcup_{C \in T} C = \mathbb{R}^d.$$

4. If $C, C' \in T$ and $C \neq C'$, then $\text{int}\, C \cap \text{int}\, C' = \emptyset$.

Remark. Since the cells of the tessellation are assumed to be compact and convex then they are necessarily convex polytopes (Lemma 10.1.1. in Schneider and Weil [2008]).

Definition 1.38 (Faces).
The intersection of a d-dimensional convex polytope P with its supporting hyperplane is called *face*. A face of dimension $k \in \{0, \ldots, d-1\}$ is called a *k-face*. The 0-faces are the *vertices*, the 1-faces are *edges* and the $(d-1)$-faces are *facets*.

For $k \in \{0, \ldots, d-1\}$, denote by $\mathbf{F}_k(P)$ the set of all k-faces of a polytope P and put

$$\mathbf{F}(P) = \bigcup_{k=0}^{d-1} \mathbf{F}_k(P).$$

We are interested in such tessellations where the faces of neighbouring cells overlap.

Definition 1.39 (Face-to-face tessellation).
A tessellation T is called *face-to-face*

$$C \cap C' \in [\mathbf{F}(C) \cap \mathbf{F}(C')] \cup \{\emptyset\}, \quad \forall C, C' \in T.$$

Similarly as for the individual cells, we define the sets of k-faces connected to a tessellation T. If T is face-to-face, define for $k \in \{0, \ldots, d-1\}$

$$\mathbf{F}_k(T) = \bigcup_{C \in T} \mathbf{F}_k(C)$$

and

$$\mathbf{F}(T) = \bigcup_{k=0}^{d-1} \mathbf{F}_k(T).$$

Definition 1.40 (Normal tessellation).
A face-to-face tessellation T is called *normal* if every k-face of T (i.e. element of $\mathbf{F}_k(T)$) is contained in precisely $d-k+1$ cells, $k = 0, \ldots, d-1$.

Note that the condition stated in the latter definition always holds for $k = d-1$ since we assumed T is face-to-face.

Definition 1.41 (Random tessellation).
Denote by $\mathbb{T}$ the set of all face-to-face tessellations of the space $\mathbb{R}^d$. By a *random tessellation*, we understand a particle process ξ on $\mathbb{R}^d$ such that $\xi \in \mathbb{T}$ a.s.

A random tessellation is called face-to-face or normal, if all its realizations are $\mathbb{P}$-almost surely face-to-face or normal, respectively.

Random tessellations can be regarded either as point processes of convex polytopes (special case of particle processes) or can be constructed from random processes of geometric objects (e.g., points, balls or hyperplanes) in space. A detailed treatment of random tessellations can be found in Møller [1994] and in Lautensack [2007]. It is often convenient to represent a random tessellation T as a marked point process in $\mathbb{R}^d$ with an appropriate mark space. We can associate various point processes with T, for example the point processes of vertices, edge midpoints, etc. If these point processes are marked with suitable marks, then we can identify T with the corresponding marked point process.

Examples

Example 1.17 (Hyperplane tessellation).
Denote by $\mathbf{H}$ a locally finite system of hyperplanes in $\mathbb{R}^d$. The cells of a hyperplane tessellation are constructed as closures of the connected components of

$$\left(\bigcup_{H \in \mathbf{H}} H \right)^C .$$

If we take an independently marked stationary Poisson point process $\eta_m = \{(x_i, \theta_i), i \geqslant 1\}$ on $\mathbb{R}$ with marks uniformly distributed in $[0, \pi)$, we can construct a planar Poisson line process, where the points of the unmarked Poisson point process stand for the signed distance from the origin and the marks for the orientation of the lines (infinitely long cylinders). The corresponding hyperplane tessellation is referred to as *Poisson line tessellation* (see Figure 1.2).

The hyperplane tessellation is face-to-face, yet it is not normal.

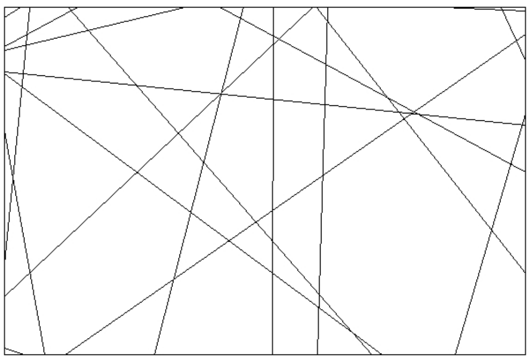


Figure 1.2: A tessellation of $\mathbb{R}^2$ induced by the Poisson line process.

Example 1.18 (Voronoi tessellation).
Let $A \neq \emptyset$ be a locally finite set in $\mathbb{R}^d$. To each $x \in A$, we define a set

$$C(x, A) := \{z \in \mathbb{R}^d : \|z - x\| \leqslant \|z - y\| \text{ for all } y \in A\}.$$

Then $C(x, A)$ represents a set of points in $\mathbb{R}^d$ such that x is their nearest point among points in A. The set $C(x, A)$ is in fact a closed convex set, since it can be written as an intersection of closed half-spaces defined for $x \neq y$ by

$$H_y^+(x) := \left\{ z \in \mathbb{R}^d : \langle z, x - y \rangle \leqslant \frac{1}{2}(\|y\|^2 - \|x\|^2) \right\}.$$

Then,

$$C(x, A) = \bigcap_{y \in A, y \neq x} H_y^+(x).$$

We call the set $C(x, A)$ the *Voronoi cell* of x with respect to A and x is called the *nucleus* of $C(x, A)$.

Theorem 1.16 (Theorem 10.2.1 in Schneider and Weil [2008]).
Let $A \subset \mathbb{R}^d$ be locally finite, non-empty and such that the corresponding Voronoi cells $C(x, A), x \in A$, are bounded. Then the collection $T := \{C(x, A); x \in A\}$ is a face-to-face tessellation.

Theorem 1.17 (Theorem 10.2.2 in Schneider and Weil [2008]).
Let μ be a stationary point process in $\mathbb{R}^d$ (assumed to satisfy $\mu \neq \emptyset$ a.s.) and let $\mathbf{T} := \{C(x, \mu); x \in \mu\}$ be the collection of the corresponding Voronoi cells. Then $\mathbf{T}$ is a stationary face-to-face random tessellation, provided that μ has a locally finite intensity measure.

The tessellation $\mathbf{T} := \{C(x, \mu); x \in \mu\}$ defined in Theorem 1.17 is called the *Voronoi tessellation* generated by the point process μ. If μ is a Poisson point process the set $\mathbf{T}$ is known as *Poisson–Voronoi tessellation*.

Theorem 1.18 (Theorem 10.2.3 in Schneider and Weil [2008]).
Every Poisson–Voronoi tessellation in $\mathbb{R}^d$ is normal.

Example 1.19 (Delaunay triangulation).
Delaunay triangulation can be viewed as dual to Voronoi tessellation (see Figure 1.3). Let $A \subset \mathbb{R}^d$ be a locally finite set such that the convex hull $\text{conv}(A) = \mathbb{R}^d$. Let $T = \{C(x, A); x \in A\}$ be a corresponding Voronoi tessellation. Let e be a vertex of T, i.e. $e \in \mathbf{F}_0(T)$. Then we define the *Delaunay cell* $D(e, A)$ by

$$D(e, A) := \text{conv}\{x \in A; e \in \mathbf{F}_0(C(x, A))\}.$$

Theorem 1.19 (Theorem 10.2.6 in Schneider and Weil [2008]).
Let $A \subset \mathbb{R}^d$ be as above and T the corresponding collection of Voronoi cells. Then

$$D := \{D(e, A); e \in \mathbf{F}_0(T)\}$$

is a face-to-face tessellation.

If we take a point process μ and let $\mathbf{T}$ be the corresponding Voronoi tessellation, we call the collection $\mathbf{D} := \{D(e, \mu); e \in \mathbf{F}_0(\mathbf{T})\}$ from Theorem 1.19 the *Delaunay triangulation* generated by μ. If μ is a Poisson point process, we speak about *Poisson–Delaunay* triangulation.

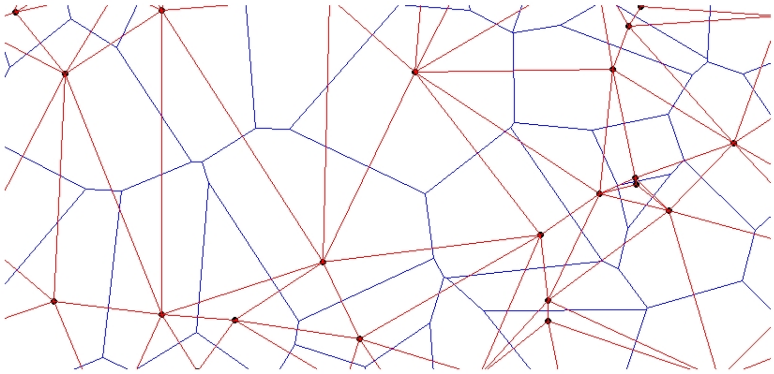


Figure 1.3: A realization of the Poisson–Voronoi tessellation of $\mathbb{R}^2$ (blue) together with the Poisson–Delaunay tessellation of $\mathbb{R}^2$ (red) induced by the same realization of Poisson point process (black dots).

Example 1.20 (Laguerre tessellation).
Laguerre tessellations form a generalization of the Voronoi tessellations where a weight is attached to each of the generating points. First, we define the power of $y \in \mathbb{R}^d$ with respect to a pair $(x, m_x) \in \mathbb{R}^d \times \mathbb{R}$ by

$$\text{pow}(y, (x, m_x)) := \|y - x\|^2 - m_x^2.$$

Now, let $A \subset \mathbb{R}^d \times \mathbb{R}$ be a countable set such that $\min_{(x,m_x)\in A} \text{pow}(y, (x, m_x))$ exists for each $y \in \mathbb{R}^d$. Then we define the *Laguerre cell* of $(x, m_x) \in A$ by

$$C((x, m_x), A) := \{y \in \mathbb{R}^d : \text{pow}(y, (x, m_x)) \leqslant \text{pow}(y, (x', m_{x'})),\ \forall (x', m_{x'}) \in A\}.$$

The point x is called the *nucleus* of the cell $C((x, m_x), A)$. The *Laguerre tessellation* induced by the set A is the set of all non-empty Laguerre cells arising from the points of A. By choosing the weights to be zero for all nuclei, we obtain the Voronoi tessellation.

Note that some nuclei may generate an empty Laguerre cell which was not the case for the Voronoi tessellation. Also, if not empty, the Laguerre cell not necessarily contains its nucleus.

Theorem 1.20 (Theorem 2.2.8 in Lautensack [2007]).
Let A be the set as above satisfying moreover the following conditions:

- *For every $y \in \mathbb{R}^d$ and every $t \in \mathbb{R}$ only finitely many elements $(x, m_x) \in A$ satisfy*
$$\text{pow}(y, (x, m_x)) \leqslant t,$$
- *the convex hull* $\text{conv}\{x : (x, m_x) \in A\} = \mathbb{R}^d$.

Then every Laguerre cell generated by A is compact and the Laguerre tessellation is face-to-face.

If, moreover, the points of A are in general position, i.e.

- *No $k+1$ nuclei are contained in a $(k-1)$-dimensional affine subspace of $\mathbb{R}^d$ for $k = 2, \dots, d$ and*
- *no $d+2$ points have equal power with respect to some point in $\mathbb{R}^d$,*

then all Laguerre cells generated by A have dimension d and the Laguerre tessellation is normal.

Again, we can replace the deterministic set A by a marked point process on $\mathbb{R}^d$ with marks in $\mathbb{R}$. If η_m is a stationary marked Poisson process in $\mathbb{R}^d$ with marks in $\mathbb{R}$, then the Laguerre tessellation induced by η_m almost surely exists and it is referred to as the *Poisson–Laguerre tessellation*.

Let us also mention the possibility to construct the *Delaunay–Laguerre tessellation* and refer to Section 2.3 in Lautensack [2007] for more details.

The importance of Laguerre tessellations can be shown by the following theorem.

Theorem 1.21 (Theorem 2.4.3 in Lautensack [2007]).
Every normal tessellation of $\mathbb{R}^d$ for $d \geqslant 3$ is a Laguerre tessellation.

Example 1.21 (Johnson–Mehl tessellation).
Let $A \subset \mathbb{R}^d \times \mathbb{R}_+$ be locally finite. The idea behind Johnson–Mehl tessellations is that the cell generated by a point $(x, t_x) \in A$ starts to grow with common speed in all directions immediately after being born which happens at time t_x. The tessellation is then created when the cells have no space to grow and fill the whole space. More precisely, we say that a point $y \in \mathbb{R}^d$ is reached by the point x at the time $T(y,(x,t_x))$, where

$$T(y,(x,t_x)) := \|x-y\| - t_x.$$

The *Johnson–Mehl cell* generated by (x, t_x) with respect to A is then defined by

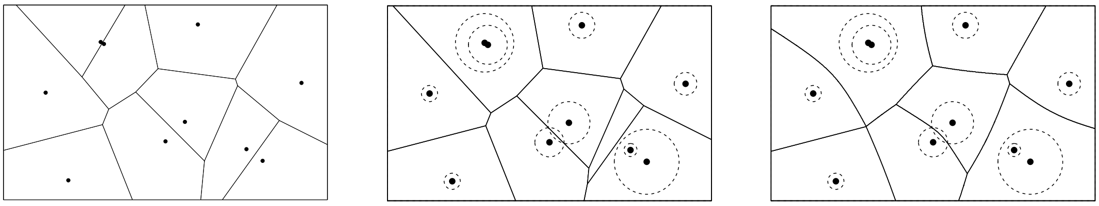
$$C((x,t_x),A) := \{y \in \mathbb{R}^d : T(y,(x,t_x)) \leqslant T(y,(x',t_{x'})),\ \forall (x',t_{x'}) \in A\}.$$

The *Johnson–Mehl tessellation* induced by the set A is then a collection of nonempty Johnson–Mehl cells.

Note that the Johnson–Mehl tessellation does not necessarily satisfy our definition of tessellation, since the cells may not be convex (see Figure 1.4 for a comparison). Hence, we do not assess the normality and face-to-face property. Also, as in the case of Laguerre tessellation, here the Johnson–Mehl cells may not contain their nuclei. Nevertheless, Johnson–Mehl tessellations form another important generalization of Voronoi tessellation and we recommend Møller [1992] for more details.

The Voronoi, Laguerre and Johnson–Mehl tessellation can be covered by the following concept.

Example 1.22 (Weighted Voronoi tessellation).
The cells of the generalized weighted Voronoi tessellations are defined as follows. Let μ_m be a marked point process on $\mathbb{R}^d$ with marks in $\mathbb{M} \subset \mathbb{R}_+$. We introduce



(a) (b) (c)

Figure 1.4: Realization of Voronoi tessellation (a), Laguerre tessellation (b) and Johnson–Mehl tessellation (c) generated based on the same point set. The circles around points in pictures (b) and (c) represent the weights attached to each point. We observe in (b) that the nucleus does not necessarily lie in the cell it is generating. In (c), we see that some points generate an empty cell.

a *weight function* $\rho : \mathbb{R}^d \times (\mathbb{R}^d \times \mathbb{M}) \to \mathbb{R}$ which for each $(x, m_x) \in \mu_m$ generates the *weighted Voronoi cell*

$$C^\rho((x, m_x), \mu_m) := \left\{ y \in \mathbb{R}^d : \rho(y, (x, m_x)) \leqslant \rho(y, (x', m_{x'})),\ \forall (x', m_{x'}) \in \mu_m \right\}.$$

Special choices of the weight function lead to the examples already presented above,

(i) Voronoi cell: $\rho_1(y, (x, m_x)) := \|x - y\|$,

(ii) Laguerre cell: $\rho_2(y, (x, m_x)) := \|x - y\|^2 - m_x^2$,

(iii) Johnson–Mehl cell: $\rho_3(y, (x, m_x)) := \|x - y\| - m_x$.

Notice that larger values of m_x generate larger cells $C((x, m_x), \mu_m)$. The weight functions $\rho_i(\cdot, (x, m_x)), i = 1, 2, 3$ are often called *the power of the point* x; see Section 10.2 in Schneider and Weil [2008]. When μ_m is a marked Poisson point process we shall refer to these tessellations as *weighted Poisson–Voronoi tessellations.*

1.6 Random geometric graphs

In the previous section, we discussed that some random tessellations could be generated by a deterministic rule and random configuration of points. A very similar approach leads to another important class of geometric structures. Random geometric graphs are random structures that can be easily described. A set of points is randomly scattered according to some probability distribution, and any two distinct points are connected by an edge if they satisfy some geometric construction rule (e.g. they are separated by a distance less than a certain specified value). That being said, some of the models for random tessellations can be interpreted as random graphs. These geometric structures form a natural model for systems in nature and society, telecommunication networks (see Zuyev [2010]), pattern recognition (see Toussaint [1982]) and other applications in computer science and optimization.

Classical random graph theory initiated in Erdös and Rényi [1959] defines a *random graph* of n vertices and N edges as a graph chosen randomly among all possible N-edged graphs with n vertices. Another variant is to fix n vertices and connect each pair of vertices at random, independently from other pairs. These variants are both known under the notion of *Erdös–Rényi model*. Here we concentrate on graphs with a random set of vertices induced by a point process and some geometric rule.

Our notion of random geometric graphs goes as follows: A *graph* is a pair $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \subset \mathbb{R}^d$ is a non-empty set of vertices and $\mathcal{E} \subset \mathcal{V}^2_{\neq}$ is a set of (un)directed edges, i.e. pairs of (un)ordered distinct points of $\mathcal{V}$. If we allow the set $\mathcal{V}$ to be a point process (typically Poisson point process), then we speak about a *random geometric graph*. A *path* is a sequence of distinct vertices and corresponding edges, each of which connects two consecutive vertices in the sequence. The *length* of a path is a sum of the Euclidean lengths of the edges in the path. We recommend Penrose [2003] for a survey on random geometric graphs.

In the rest of this section, we present a list of classical examples of construction rules of a random graph.

Example 1.23 (Gilbert graph).
Let $\mathbf{x} \in \mathbf{N}$ be a locally finite point set and let $\delta > 0$ be a fixed parameter. Two distinct points $x, y \in \mathbf{x}$ are connected by an edge if $\|x-y\| \leqslant \delta$. If $\mathbf{x}$ is a realization of a Poisson point process, then the resulting graph is referred to as *Gilbert graph*.

One of the possible generalizations of Gilbert graph is to take a symmetric function $G : \mathbb{R}^d \to [0,1]$ and put an edge between any distinct points $x, y \in \mathbf{x}$ with probability $G(x-y)$.

Example 1.24 (k-nearest neighbour graph).
Let $k \in \mathbb{N}$ be a fixed positive integer. The *k-th nearest neighbour graph* on a locally finite point configuration $\mathbf{x} \in \mathbb{N}$ is a directed graph where an edge is going from x to y in $\mathbf{x}$ if y is one of the k nearest neighbours of x among all the points in $\mathbf{x}$. It can happen that the graph is not well defined due to the higher number of points having exactly the same distance from x. In that case, an additional rule has to be included (e.g. lexicographic ordering). Nevertheless, this event has zero probability for $\mathbf{x}$ being a realization of a suitable point process (Poisson, binomial, etc.)

An undirected version of the k-nearest neighbour graph is sometimes considered and is defined as the k-nearest neighbour graph where all the directions are forgotten. In case, there is a double edge between two points, we take it as a single one.

Example 1.25 (Sphere of influence graph).
The *sphere of influence* of a given point x in a point configuration $\mathbf{x}$ is the largest ball in $\mathbb{R}^d$ centered in x not containing any other point of $\mathbf{x}$. An edge of the *sphere of influence graph* on $\mathbf{x}$ is present between two distinct points $x, y \in \mathbf{x}$ if their spheres of influence overlap. By definition, the sphere of influence graph is undirected.

Some of the examples of random graphs are connected with combinatorial optimization problems. In order to introduce them, let us fix a given vertex set $\mathcal{V} = \{x_1, \ldots, x_n\}$.

Example 1.26 (Travelling salesman).
We say a path over a given vertex set $\mathcal{V} = \{x_1, \ldots, x_n\}$ is *closed* if it is traversing each vertex exactly once. We chose such $\mathcal{E}$ among all the possible sets of edges such that the corresponding path is closed has a minimal length among all closed paths.

Example 1.27 (Minimum spanning tree).
A *cycle* in a graph is a non-empty path in which the only repeated vertices are the first and last vertices. A *spanning tree* of $\mathcal{V}$ is an undirected connected graph, such that it connects all the vertices in $\mathcal{V}$ without any cycles. Note that if there are n vertices in the graph, then each spanning tree has exactly $n-1$ edges.

The *minimum spanning tree* is a spanning tree with a minimal length of the path among all the spanning trees of $\mathcal{V}$.

Example 1.28 (Minimal matching).
Suppose that $\mathcal{V}$ is an even number. A *matching* of $\mathcal{V}$ is an undirected graph requiring every vertex to be matched (i.e. to form an edge with another vertex) and all the edges to have no common vertices. Hence, then there is exactly $\#(\mathcal{V})/2$ edges in the path.

A *minimal matching* of $\mathcal{V}$ is a matching with minimal length of the path among all other matchings of $\mathcal{V}$.

In any of the mentioned optimization problems, instead of the length of edges, we can consider any weight associated to a given edge and minimize with respect to the total weight of the path.

2. Malliavin–Stein's method

Stein's method serves as a way to get explicit estimates in probability theory, typically yielding a normal or Poisson approximation. By approximation in this context, we mean providing estimates of a given distance between the laws of two random variables. Examples of the probability distances include total variation distance, Kolmogorov or Wasserstein distance.

The method was first introduced in Stein [1972] in order to give a convergence speed for the central limit theorem for a sum of dependent random variables satisfying a mixing condition. The main building block is Stein's lemma, which characterizes a normally distributed random variable Z. That is, the fact that $Z \sim N(0,1)$ if and only if

$$\mathbb{E}\left[f'(Z)\right] = \mathbb{E}\left[Zf(Z)\right]$$

for all continuous functions $f : \mathbb{R} \to \mathbb{R}$, for which the above expectations exist. The approach was then extended to Poisson approximation in Chen [1975] leading to the so-called *Chen–Stein method*. According to the Chen–Stein lemma, N has Poisson law with parameter $\lambda > 0$ if and only if for every bounded $f : \mathbb{N} \to \mathbb{N}$,

$$\mathbb{E}\left[Nf(N)\right] = \mathbb{E}\left[\lambda f(N+1)\right].$$

At this work, only the normal approximation is considered. For the Poisson approximation by the Chen–Stein lemma, see e.g. Section 2.4 in Bourguin and Peccati [2016].

The Malliavin calculus of variations was introduced in Malliavin [1978] as an infinite-dimensional differential calculus with operators acting on functionals of Gaussian processes. The theory is based on the infinite-dimensional integration by parts formulae. It was shown recently in Nourdin and Peccati [2009a] and Nourdin and Peccati [2009b] that one can combine Malliavin calculus on the Gaussian space and Stein's method in order to obtain bounds for the normal and non-normal approximation of functionals of Gaussian fields. Later in Peccati et al. [2010], this approach was extended to the normal approximation of functionals of Poisson measures defined on abstract Borel spaces. The idea is to express the estimates arising from Stein's method in terms of Malliavin operators.

The application in stochastic geometry, however, profits from the version of the Malliavin calculus build on the Poisson space. The typical aim is to approximate behaviour of certain functionals of point processes with the standard normal or Poisson distribution. In this chapter, we show how the Malliavin calculus of variations and Stein's method of probability approximations may be combined into a powerful and flexible tool, the *Malliavin–Stein method*, for proving central limit theorems as well as computing explicit rates of convergence for models in stochastic geometry. The application is illustrated on selected models from stochastic geometry at the end of this chapter. For a nice presentation of very recent developments in the theory of Malliavin–Stein's method, we recommend Azmoodeh et al. [2021].

2.1 Stein's method for normal approximation

The aim of this section is to introduce the idea behind Stein's method in the one-dimensional case. The standard modern reference concerning Stein's method for normal approximation is the monograph Chen et al. [2011]. This text is also based on Nourdin and Peccati [2012], Barbour and Chen [2005a] and Barbour and Chen [2005b]. The multi-dimensional case is not discussed here and we recommend Chapter 4 in Nourdin and Peccati [2012] for the matter.

Distances between probability distributions

The goal of Stein's method is to find an upper bound for the difference between the expectations of all functions of a given family of test functions under two given distributions. The choice of the test functions determines the associated metric.

Definition 2.1 (Separating collection of functions).
Let $\mathcal{H}$ be a collection of measurable functions $h : \mathbb{R}^d \to \mathbb{C}$. We say that the collection $\mathcal{H}$ is *separating* if the following holds true: If P, Q are probability measures on $\mathbb{R}^d$ such that $\int |h| \mathrm{d}P, \int |h| \mathrm{d}Q < \infty$ and $\int h \mathrm{d}P = \int h \mathrm{d}Q$ for all $h \in \mathcal{H}$, then $P = Q$.

Definition 2.2 (Distance between probability distributions).
Let $\mathcal{H}$ be a separating collection of functions and P, Q two probability measures on $\mathbb{R}^d$ with $\int |h| \mathrm{d}P, \int |h| \mathrm{d}Q < \infty$ for all $h \in \mathcal{H}$. Then the *distance between laws P and Q* induced by $\mathcal{H}$ is defined by

$$d_{\mathcal{H}}(P, Q) = \sup_{h \in \mathcal{H}} \left| \int h \mathrm{d}P - \int h \mathrm{d}Q \right|.$$

Remark. If $\mathcal{H}$ is a separating collection of functions, then the probability distance $d_{\mathcal{H}}$ induced by $\mathcal{H}$ verifies the usual axioms of a distance (metric) on the space of probability measures P on $\mathbb{R}^d$ such that

$$\int_{\mathbb{R}^d} |h(x)| \mathrm{d}P(x) < \infty$$

for all $h \in \mathcal{H}$. Specifically, for probability measures P, Q, S:

- $d_{\mathcal{H}}(P, Q) = d_{\mathcal{H}}(Q, P)$,
- $d_{\mathcal{H}}(P, Q) = 0$ if and only if $P = Q$,
- $d_{\mathcal{H}}(P, Q) \leqslant d_{\mathcal{H}}(P, S) + d_{\mathcal{H}}(S, Q)$.

Now, we present a list of the typical choices of the separating collections $\mathcal{H}$. For simplicity, we assume the distributions P, Q in the following definitions to be some probability measures on $\mathbb{R}$. The extension to the higher dimensional case is straightforward.

Definition 2.3 (Total variation distance).
The *total variation distance* between two distributions P and Q is defined by

$$d_{TV}(P,Q) := \sup_{h\in\mathcal{H}} \left| \int h \mathrm{d}P - \int h \mathrm{d}Q \right|,$$

where $\mathcal{H} := \{\mathbf{1}_A; A \in \mathcal{B}(\mathbb{R})\}$.

Definition 2.4 (Kolmogorov distance).
The *Kolmogorov distance* between two distributions P and Q is defined by

$$d_K(P,Q) := \sup_{h\in\mathcal{H}} \left| \int h \mathrm{d}P - \int h \mathrm{d}Q \right|,$$

where $\mathcal{H} := \{\mathbf{1}_{(-\infty,z]}; z \in \mathbb{R}\}$.

Remark. The total variation distance of distributions P, Q can be equivalently defined as

$$d_{TV}(P,Q) := \sup\{|P(A) - Q(A)|; A \in \mathcal{B}(\mathbb{R})\}.$$

Analogically, the Kolmogorov distance of P, Q can be rewritten in the form

$$d_K(P,Q) := \sup\{|P(-\infty,z] - Q(-\infty,z]|; z \in \mathbb{R}\}.$$

Clearly, for two probability measures P, Q it holds that

$$0 \leqslant d_K(P,Q) \leqslant d_{TV}(P,Q) \leqslant 1.$$

Definition 2.5 (Wasserstein distance).
The *Wasserstein distance* between two distributions P and Q is defined by

$$d_W(P,Q) := \sup_{h\in\mathrm{Lip}(1)} \left| \int h \mathrm{d}P - \int h \mathrm{d}Q \right|,$$

where $\mathrm{Lip}(1) := \{h : \mathbb{R} \to \mathbb{R}; \|h'\| \leqslant 1\}$ denotes the class of real-valued Lipschitz functions with Lipschitz constant less than or equal to 1. Here by $\|h\|$, we denote $\sup_{x\in\mathbb{R}} |h(x)|$.

The Wasserstein distance is also known as *Kantorovich–Monge–Rubinstein metric*. Note that it ranges in $[0,\infty]$.

Notation. Let P_X and P_Y be the distributions of random variables X and Y, respectively. We will use the notation $d(X,Y)$ for $d(P_X,P_Y)$ whenever d is any distance on the space of all probability measures on $\mathbb{R}$.

Proposition 2.1 (Theorem 3.3 in Chen et al. [2011]).
If X is a real-valued random variable and Z the standard normally distributed random variable, then

$$d_K(X,Z) \leqslant 2\sqrt{d_W(X,Z)}.$$

The following result states that the convergence with respect to d_{TV}, d_K and d_W is stronger than weak convergence. Hence, $d(X_n, X) \to 0$ implies $X_n \xrightarrow{\mathcal{D}} X$ whenever $X, \{X_n, n \in \mathbb{N}\}$ are random variables and d is either total variation, Kolmogorov or Wasserstein distance.

Proposition 2.2 (Proposition C.3.1 in Nourdin and Peccati [2012]). *The topologies induced by the three distances d_{TV}, d_K and d_W on the set of probability measures on $\mathbb{R}$ are strictly stronger than the topology of the convergence in distribution.*

Moreover, the Kolmogorov distance metrizes the convergence in distribution towards real-valued random variables whose distribution function is continuous. However, this result can not be extended to the total variation nor Wasserstein distance.

Proposition 2.3 (Proposition C.3.2 in Nourdin and Peccati [2012]). *Let $X, \{X_n\}_{n\geqslant 1}$ be random variables in $\mathbb{R}$ and let X have a continuous distribution function. Then $X_n \xrightarrow{\mathcal{D}} X$ if and only if $d_K(X_n, X) \to 0$ as $n \to \infty$.*

Stein's lemma

In this section, we always denote by Z the standard normally distributed random variable. For a sequence of random variables X_n, the goal is to find uniform upper bounds of the type $d(X_n, Z) \leqslant \phi(n), n \geqslant 1$ for d being one of the probability distances defined above. The sequence $\{\phi(n)\}_{n\geqslant 1}$ of positive numbers is referred to as the *rate of convergence* if $\phi(n) \to 0$.

Let $f : \mathbb{R} \to \mathbb{R}$ be a bounded measurable function with bounded derivative. Then the following observation can be made:

$$\mathbb{E}\left[f'(Z) - Zf(Z)\right] = 0. \tag{2.1}$$

It can be shown directly by integration by parts

$$\begin{aligned}
\mathbb{E} f'(Z) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(x) e^{-x^2/2} \mathrm{d}x \\
&= \left[\frac{1}{\sqrt{2\pi}} f(x) e^{-x^2/2}\right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x f(x) e^{-x^2/2} \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x f(x) e^{-x^2/2} \mathrm{d}x \\
&= \mathbb{E}\, Z f(Z).
\end{aligned}$$

Moreover, let C_{bd} be the set of continuous and piecewise continuously differentiable functions $f : \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}\,|f'(Z)| < \infty$. Then the equality (2.1) can be generalized into a characterization of the standard normal distribution that forms a base of Stein's method for normal approximation:

Theorem 2.1 (Stein's lemma). *Let W be a real-valued random variable. Then W has a standard normal distribution if and only if*

$$\mathbb{E} f'(W) = \mathbb{E}\, W f(W) \tag{2.2}$$

for all $f \in C_{bd}$.

Proof. We revise the proofs of Lemma 2.1 and Lemma 2.2 in Chen and Shao [2005].

1. Necessity: Take $f \in C_{bd}$ and suppose W has a standard normal distribution. Then

$$\begin{aligned}\mathbb{E}\, f'(W) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f'(w) e^{-w^2/2} \mathrm{d}w \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} f'(w) \left(\int_{-\infty}^{w} (-x) e^{-x^2/2} \mathrm{d}x \right) \mathrm{d}w \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} f'(w) \left(\int_{w}^{\infty} x e^{-x^2/2} \mathrm{d}x \right) \mathrm{d}w.\end{aligned}$$

Using Fubini's theorem, we arrive at

$$\begin{aligned}\mathbb{E}\, f'(W) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \left(\int_{x}^{0} f'(w) \mathrm{d}w \right) (-x) e^{-x^2/2} \mathrm{d}x \\ &\quad + \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \left(\int_{0}^{x} f'(w) \mathrm{d}w \right) x e^{-x^2/2} \mathrm{d}x \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |f(x) - f(0)| x e^{-x^2/2} \mathrm{d}x \\ &= \mathbb{E}\, W f(W).\end{aligned}$$

2. Sufficiency: For a fixed $z \in \mathbb{R}$, we are interested in the ordinary differential equation

$$f'(w) - w f(w) = \mathbf{1}_{-(\infty,z]}(w) - \Phi(z), \tag{2.3}$$

where Φ is a distribution function of the standard Gaussian random variable. Multiplying both sides of (2.3) by $-e^{-w^2/2}$ leads to

$$\left(e^{-w^2/2} f(w) \right)' = -e^{-w^2/2} (\mathbf{1}_{-(\infty,z]} - \Phi(z)).$$

Hence, the solution f_z of (2.3) is given by

$$\begin{aligned}f_z(w) &= e^{w^2/2} \int_{-\infty}^{w} [\mathbf{1}_{(-\infty,z]}(x) - \Phi(z)] e^{-x^2/2} \mathrm{d}x \\ &= -e^{w^2/2} \int_{w}^{\infty} [\mathbf{1}_{(-\infty,z]}(x) - \Phi(z)] e^{-x^2/2} \mathrm{d}x \\ &= \begin{cases} \sqrt{2\pi} e^{w^2/2} \Phi(w)[1 - \Phi(z)], & \text{if } w \leqslant z, \\ \sqrt{2\pi} e^{w^2/2} \Phi(z)[1 - \Phi(w)], & \text{if } w > z. \end{cases}\end{aligned}$$

Clearly, f_z is continuous and piecewise continuously differentiable for all $z \in \mathbb{R}$. We shall prove that $\mathbb{E}\, |f_z'(Z)| < \infty$. By (2.3), we have

$$\begin{aligned}f_z'(w) &= w f_z(w) + \mathbf{1}_{(-\infty,z]}(w) - \Phi(z) \\ &= \begin{cases} w f_z(w) + 1 - \Phi(z), & \text{if } w < z, \\ w f_z(w) - \Phi(z), & \text{for } w > z. \end{cases}\end{aligned} \tag{2.4}$$

The function $w f_z(w)$ is increasing. To see that, take $z \geqslant 0$. The other case is similar, since $f_{-z}(-w) = f_z(w)$. For $w < z$, it is clear, since

$$(w f_z(w))' = \sqrt{2\pi}(1 - \Phi(z)) \left((1 + w^2) e^{w^2/2} \Phi(w) + \frac{w}{\sqrt{2\pi}} \right) \geqslant 0.$$

On the other hand, for $w > 0$, note that

$$\int_w^\infty e^{-x^2/2}\mathrm{d}x \leqslant \int_w^\infty \frac{x}{w} e^{-x^2/2}\mathrm{d}x = \frac{e^{-w^2/2}}{w}. \tag{2.5}$$

Moreover,

$$1 - \Phi(w) \geqslant \frac{we^{-w^2/2}}{(1+w^2)\sqrt{2\pi}}. \tag{2.6}$$

To show (2.6), denote $g_1(w) := (1+w^2)\int_w^\infty e^{-x^2/2}\mathrm{d}x$ and $g_2(w) := we^{-w^2/2}$. We can observe that

$$g_1(0) = \sqrt{\pi/2} > 0 = g_2(0). \tag{2.7}$$

At the same time

$$\lim_{w\to\infty} g_1(w) \geqslant 0 = \lim_{w\to\infty} g_2(w). \tag{2.8}$$

Further, for $w \geqslant 0$,

$$\begin{aligned} g_1'(w) - g_2'(w) &= 2w \int_w^\infty e^{-x^2/2}\mathrm{d}x - (1+w^2)e^{-w^2/2} - (1-w^2)e^{-w^2/2} \\ &= 2w \int_w^\infty e^{-x^2/2}\mathrm{d}x - 2e^{-w^2/2} \\ &\leqslant 0. \end{aligned}$$

In the third inequality, we used (2.5). Since the difference of the derivatives is always non-positive, the graphs of g_1 and g_2 never cross each other. Otherwise, it would contradict the observations (2.7) and (2.8). Thus, $g_1(w) \geqslant g_2(w)$ for all $w \geqslant 0$, from which (2.6) follows immediately.

Now, we use (2.6) to estimate the derivative of $wf_z(w)$ for $w \geqslant z$:

$$(wf_z(w))' = \sqrt{2\pi}\Phi(z)\left((1+w^2)e^{w^2/2}(1-\Phi(w)) - \frac{w}{\sqrt{2\pi}}\right) \geqslant 0.$$

Indeed, the function $wf_z(w)$ is increasing with limits

$$\lim_{w\to-\infty} wf_z(w) = \lim_{w\to-\infty} \sqrt{2\pi}we^{w^2/2}\Phi(w)[1-\Phi(z)] = \Phi(z) - 1 \tag{2.9}$$

and

$$\lim_{w\to\infty} wf_z(w) = \lim_{w\to\infty} \sqrt{2\pi}we^{w^2/2}\Phi(z)[1-\Phi(w)] = \Phi(z). \tag{2.10}$$

By plugging (2.9) and (2.10) into (2.4), we obtain for $w \leqslant z$

$$0 < f_z'(w) = wf_z(w) + 1 - \Phi(z) \leqslant \lim_{x\to\infty} xf_z(x) + 1 - \Phi(z) = 1,$$

and for $w > z$

$$-1 = \lim_{x\to-\infty} xf_z(x) - \Phi(z) \leqslant wf_z(w) - \Phi(z) = f_z'(w) < 0.$$

Finally, $\mathbb{E}\,|f_z'(Z)| \leqslant 1$ and therefore, $f_z \in C_{bd}$. Suppose that (2.2) holds for all $f \in C_{bd}$. Then it is also satisfied by f_z and hence, by (2.3), we have

$$0 = \mathbb{E}\left[f_z'(W) - Wf_z(W)\right] = \mathbb{E}\left[\mathbf{1}_{(-\infty,z]}(W) - \Phi(z)\right] = \mathbb{P}(W \leqslant z) - \Phi(z).$$

Thus, W has a standard normal distribution.

□

Stein's equation

Based on Stein's lemma, if X is a random variable for which $\mathbb{E}[Xf(X) - f'(X)]$ is close to zero for some large class of smooth functions f, is it possible to conclude that X is close (with respect to some distance) to the standard normal distribution? To answer the question, we need to introduce the notion of *Stein's equation* associated with a given function h such that $\mathbb{E}|h(Z)| < \infty$.

Definition 2.6 (Stein's equation).
Let $h : \mathbb{R} \to \mathbb{R}$ be a Borel function such that $\mathbb{E}|h(Z)| < \infty$. Then *Stein's equation associated with* h is the ordinary differential equation

$$f'(x) - xf(x) = h(x) - \mathbb{E}\, h(Z). \tag{2.11}$$

A *solution* to (2.11) is a function f that is absolutely continuous and such that there exists a version of the derivative f' satisfying (2.11) for all $x \in \mathbb{R}$.

Proposition 2.4 (Proposition 3.2.2 in Nourdin and Peccati [2012]).
Every solution to (2.11) has the form

$$f(x) = ce^{x^2/2} + e^{x^2/2} \int_{-\infty}^{x} [h(y) - \mathbb{E}\, h(Z)] e^{-y^2/2} \mathrm{d}y, \quad x \in \mathbb{R},$$

where $c \in \mathbb{R}$.

Notation. Denote by f_h the unique solution f to the Stein's equation associated with a function h that satisfies

$$\lim_{x \to \pm\infty} e^{-x^2/2} f(x) = 0.$$

Note that f_h is given by

$$\begin{aligned} f_h(x) &= e^{x^2/2} \int_{-\infty}^{x} [h(y) - \mathbb{E}\, h(Z)] e^{-y^2/2} \mathrm{d}y \\ &= -e^{x^2/2} \int_{x}^{\infty} [h(y) - \mathbb{E}\, h(Z)] e^{-y^2/2} \mathrm{d}y. \end{aligned} \tag{2.12}$$

Lemma 2.1 (Lemma 2.3 in Chen and Shao [2005]).
For any absolutely continuous function $h : \mathbb{R} \to \mathbb{R}$, the solution f_h given in (2.12) satisfies

$$\|f_h\| \leqslant \min(\sqrt{\frac{\pi}{2}} \|h(\cdot) - \mathbb{E}\, h(Z)\|, 2\|h'\|),$$

$$\|f_h'\| \leqslant \min(2\|h(\cdot) - \mathbb{E}\, h(Z)\|, 4\|h'\|)$$

and

$$\|f_h''\| \leqslant 2\|h'\|.$$

Normal approximation with respect to d_{TV}, d_K and d_W

For a given random variable X, the goal now is to bound

$$\sup_{h \in \mathcal{H}} |\mathbb{E}\, h(X) - \mathbb{E}\, h(Z)|,$$

where $\mathcal{H}$ is some separating collection of functions $h : \mathbb{R} \to \mathbb{R}$ with $\mathbb{E}\,|h(X)| < \infty$ and $\mathbb{E}\,|h(Z)| < \infty$. The trick of Stein's method consists of finding another class of functions $\mathcal{H}'$ such that

$$\sup_{h\in\mathcal{H}} |\mathbb{E}\,h(W) - \mathbb{E}\,h(Z)| \leqslant \sup_{f\in\mathcal{H}'} |\mathbb{E}\,[f'(X) - Xf(X)]|.$$

First, by taking expectation on both sides of (2.11), we arrive at

$$\mathbb{E}\,h(X) - \mathbb{E}\,h(Z) = \mathbb{E}\,[f_h'(X) - Xf_h(X)].$$

Hence,

$$d_{\mathcal{H}}(X, Z) = \sup_{h\in\mathcal{H}} |\mathbb{E}\,[f_h'(X) - Xf_h(X)]| \,.$$

Let us now focus on the examples of $\mathcal{H}$ associated with total variation, Kolmogorov and Wasserstein distances. The following three theorems can be found in Nourdin and Peccati [2012] (cf. Sections 3.3 - 3.5). We took hints from there to complete the proofs.

Theorem 2.2 (Normal approximation w.r.t d_{TV}).
Let $h : \mathbb{R} \to [0, 1]$ be a Borel function. Then, the solution f_h to Stein's equation (2.11) satisfies

$$\|f_h\| \leqslant \sqrt{\frac{\pi}{2}} \quad \textit{and} \quad \|f_h'\| \leqslant 2.$$

In particular, for any integrable random variable X

$$d_{TV}(X, Z) \leqslant \sup_{f\in\mathcal{F}_{TV}} |\mathbb{E}\,f'(X) - \mathbb{E}\,Xf(X)|,$$

where $\mathcal{F}_{TV} = \{f : \mathbb{R} \to \mathbb{R}$ absolutely continuous ; $\|f\| \leqslant \sqrt{\pi/2}, \|f'\| \leqslant 2\}$.

Proof. Take any $h : \mathbb{R} \to [0, 1]$ Borel. Then obviously, $|h(x) - \mathbb{E}\,h(Z)| \leqslant 1$ for all $x \in \mathbb{R}$. The statement then follows directly from Lemma 2.1.

□

The bounds from Theorem 2.2 can be applied for the Kolmogorov distance as well since the relationship $d_K \leqslant d_{TV}$ holds. However, they can be further improved.

Theorem 2.3 (Normal approximation w.r.t. d_K).
Let $z \in \mathbb{R}$. Then the solution f_z satisfies

$$\|f_z\| \leqslant \frac{\sqrt{2\pi}}{4} \quad \textit{and} \quad \|f_z'\| \leqslant 1.$$

In particular, for any integrable random variable X,

$$d_K(X, Z) \leqslant \sup_{f\in\mathcal{F}_K} |\mathbb{E}\,f'(X) - \mathbb{E}\,Xf(X)|,$$

where $\mathcal{F}_K = \{f : \mathbb{R} \to \mathbb{R}$ absolutely continuous; $\|f\| \leqslant \sqrt{2\pi}/4, \|f'\| \leqslant 1\}$.

Proof. Recall that the Kolmogorov distance is induced by the set of functions $\mathcal{H} = \{\mathbf{1}_{(-\infty,z]}; z \in \mathbb{R}\}$. For a special choice $h = \mathbf{1}_{(-\infty,z]}$, $z \in \mathbb{R}$, we denote by f_z the solution $f_{\mathbf{1}_{(-\infty,z]}}$ given by (2.12). Then, for $x \leqslant z$, we have

$$\begin{aligned}
f_z(x) &= e^{x^2/2} \int_{-\infty}^{x} \left[\mathbf{1}_{(-\infty,z]}(y) - \mathbb{E}\,\mathbf{1}_{(-\infty,z]}(Z)\right] e^{-y^2/2} \mathrm{d}y \\
&= e^{x^2/2} \int_{-\infty}^{x} (1 - \Phi(z)) e^{-y^2/2} \mathrm{d}y \\
&= \sqrt{2\pi} e^{x^2/2} \Phi(x)(1 - \Phi(z)) \\
&\leqslant \sqrt{2\pi} e^{z^2/2} \Phi(z)(1 - \Phi(z)).
\end{aligned}$$

It is easy to check that the function $g(z) := \sqrt{2\pi} e^{z^2/2} \Phi(z)(1 - \Phi(z))$ attains its maximum in $z = 0$, therefore

$$f_z(x) \leqslant \frac{\sqrt{2\pi}}{4}.$$

Analogically, having $x > z$,

$$\begin{aligned}
f_z(x) &= e^{x^2/2} \left(\int_{-\infty}^{z} e^{-y^2/2} \mathrm{d}y - \int_{-\infty}^{x} \Phi(z) e^{-y^2/2} \mathrm{d}y \right) \\
&= \sqrt{2\pi} e^{x^2/2} \Phi(z)(1 - \Phi(x)) \\
&\leqslant \sqrt{2\pi} e^{x^2/2} \Phi(x)(1 - \Phi(x)) \\
&\leqslant \frac{\sqrt{2\pi}}{4}.
\end{aligned}$$

Moreover, by f_z' we denote the corresponding version of the derivative of f_z satisfying the Stein's equation

$$f_z'(x) = x f_z(x) + \mathbf{1}_{(-\infty,z]}(x) - \Phi(z).$$

The estimate $\|f_z'\| \leqslant 1$ was already shown in the proof of Theorem 2.1.

□

Theorem 2.4 (Normal approximation w.r.t. d_W).
Let X be a square-integrable random variable. Then

$$d_W(X, Z) \leqslant \sup_{f \in \mathcal{F}_W} |\mathbb{E}\, f'(X) - \mathbb{E}\, X f(X)|,$$

where $\mathcal{F}_W := \{f : \mathbb{R} \to \mathbb{R}$ continuously differentiable; $\|f'\| \leqslant \sqrt{2/\pi}\}$.

Proof. Let h be any Lipschitz function with the Lipschitz constant 1. Then the solution f_h of Stein's equation (2.11) is continuously differentiable directly from the explicit expression (2.12). Further, by Proposition 3.5.1 in Nourdin and Peccati [2012], f_h admits the representation

$$f_h(x) = -\int_0^{\infty} \frac{e^{-t}}{\sqrt{1 - e^{-2t}}} \mathbb{E}\left[h(e^{-t}x + \sqrt{1 - e^{-2t}} Z) Z\right] \mathrm{d}t$$

and $\|f_h'\| \leqslant \sqrt{2/\pi}$. The statement of Theorem 2.4 is then a direct consequence of this result.

□

Remark. 1. In the definition of $\mathcal{F}_{TV}$ and $\mathcal{F}_K$, we assume that for $f \in \mathcal{F}_{TV}$, $f \in \mathcal{F}_K$ resp., there exists a version of f' satisfying the prescribed conditions.

2. The supremum $\sup_{f \in \mathcal{F}} |\mathbb{E} f'(X) - \mathbb{E} X f(X)|$, where $\mathcal{F}$ is either $\mathcal{F}_{TV}$ or $\mathcal{F}_K$, stands for $\sup_{f \in \mathcal{F}} \sup\{|\mathbb{E} g(X) - \mathbb{E} X f(X)|, g \text{ a version of } f'\}$.

3. The requirement that X is integrable is needed so that $\mathbb{E}|X f(X)|$ exists for every $f \in \mathcal{F}_{TV}$ or $f \in \mathcal{F}_K$. The assumption of square-integrability guarantees the existence of $\mathbb{E}|X f(X)|$ for all $f \in \mathcal{F}_W$.

Example 2.1 (Berry–Esseen bounds and CLT).
The Stein's method can be applied to provide a proof of the classical Berry–Esseen theorem. Let $\{X_k; k \geqslant 1\}$ be a sequence of i.i.d. random variables such that $\mathbb{E} X_1 = 0$ and $\mathbb{E} X_1^2 = 1$. Define for $n \in \mathbb{N}$,

$$S_n := \frac{1}{\sqrt{n}} \sum_{k=1}^{n} X_k.$$

Then by the Stein's method

$$d_K(S_n, Z) \leqslant \frac{C \mathbb{E}|X_1|^3}{\sqrt{n}}, \quad n \in \mathbb{N},$$

where $C > 0$ is some universal constant depending neither on n nor $X_k, k \in \mathbb{N}$ (as shown lately in Tyurin [2009], it holds for $C = 0.4785$). As a consequence, if $\mathbb{E}|X_1|^3 < \infty$,

$$S_n \xrightarrow{\mathcal{D}} Z.$$

2.2 Malliavin calculus on the Poisson space

Malliavin calculus (also known as stochastic calculus of variations) extends the calculus of variations from deterministic functions to stochastic processes. In particular, it allows the computation of derivatives of random variables and integration by parts with random variables which is needed in order to explicitly assess the bounds arising from Stein's method.

The aim of this section is to briefly explain the basic elements of Malliavin calculus on the Poisson space and demonstrate how the bounds in the previous section based on Stein's method can be combined with Malliavin operators on abstract Poisson spaces defined via the *Fock space representation*. This section is mainly based on Peccati and Reitzner [2016] and Last and Penrose [2017].

For an introduction to the theory of Malliavin calculus on the Wiener space, we recommend Nualart [2006] or Nourdin and Peccati [2009a].

Malliavin operators

Here and throughout the whole section, we will assume that $(\mathbb{X}, \mathcal{X})$ is some fixed measurable space and η is a Poisson point process on $\mathbb{X}$ (see Example 1.5) with σ-finite intensity measure λ and distribution P_η. By the *Poisson space*, we mean the space $L^2(\mathbf{N}(\mathbb{X}), P_\eta)$ of all measurable functions $f : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ satisfying $\mathbb{E}\, f(\eta)^2 < \infty$. A measurable function of a Poisson process is called *Poisson functional*.

Moreover, for a space X equipped with a measure ρ, we will shorten the notation and write $L^q(\rho)$ for $L^q(X, \rho)$ whenever $q > 0$ and it is clear what is the underlying space. Throughout this chapter, we will moreover use the following notation.

Notation.

$$L^0_\eta(\mathbb{P}) := \{F; F = f(\eta)\ \mathbb{P}\text{-a.s. for some measurable } f : \mathbf{N}(\mathbb{X}) \to \mathbb{R}\},$$
$$L^q_\eta(\mathbb{P}) := \{F \in L^0_\eta(\mathbb{P});\ F \in L^q(\mathbb{P})\}, \quad q > 0.$$

Similarly, denote

$$L^0_\eta(\mathbb{P}\otimes\lambda) := \{G; G = g(\eta, x)\ \mathbb{P}\otimes\lambda\text{-a.s. for some measurable } h : \mathbf{N}(\mathbb{X}) \times \mathbb{X} \to \mathbb{R}\},$$
$$L^q_\eta(\mathbb{P} \otimes \lambda) := \{G \in L^0_\eta(\mathbb{P} \otimes \lambda); G \in L^q(\mathbb{P} \otimes \lambda)\}, \quad q > 0.$$

The functions f, resp. g such that $F = f(\eta)$ for some $F \in L^0_\eta(\mathbb{P})$ and $G = g(\eta, x)$ for some $G \in L^0_\eta(\mathbb{P}\otimes\lambda)$ are called the *representatives* of the functionals F and G, resp.

Next, for $f,\ g \in L^2(\lambda^n)$, we denote the inner product of f and g by

$$\langle f,\ g\rangle_n = \int_{\mathbb{X}^n} fg \, \mathrm{d}\lambda^n.$$

The associated norm is denoted by

$$\|f\|_n = \sqrt{\langle f, f\rangle_n}.$$

For $n = 0$, we put $\langle a, b\rangle_0 = ab$ for $a, b \in \mathbb{R}$. For $n = 1$, to emphasize that $f, g \in L^q(\lambda)$, we write $\langle f, g\rangle_{L^q(\lambda)}$, resp. $\|f\|_{L^q(\lambda)}$.

Definition 2.7 (Difference operator).
Let $F : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ be a measurable function. For $y \in \mathbb{X}$, we define the *difference operator* (or *add-one cost operator*) as a function $D_y F : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ satisfying

$$D_y F(\mathbf{x}) = F(\mathbf{x} + \delta_y) - F(\mathbf{x}), \quad \mathbf{x} \in \mathbf{N}(\mathbb{X}).$$

Iterating this definition, we get for $n \in \mathbb{N},\ n \geqslant 2$, the *difference operator of the n-th order* $D^n_{y_1,\dots,y_n} F : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ defined by

$$D^n_{y_1,\dots,y_n} F = D^1_{y_1} D^{n-1}_{y_2,\dots,y_n} F,$$

where $D^1 = D$ and $D^0 F = F$.

For $F \in L^2_\eta(\mathbb{P})$ and $x \in \mathbb{X}$ define the random variable $D_x F := D_x f(\eta)$ and for $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{X}$ define $D^n_{x_1,\dots,x_n} F := D^n_{x_1,\dots,x_n} f(\eta)$. Moreover, denote by $D^n F$ the mapping $(\omega, x_1, \dots, x_n) \mapsto (D^n_{x_1,\dots,x_n} F)(\omega)$.

By the Slivnyak–Mecke equation (see (1.9)), the definitions of DF and D^nF are $\mathbb{P} \otimes \lambda$-a.s. independent of the choice of the representative f.

Remark. The n-th order difference operator can be expressed by

$$D^n_{y_1,\dots,y_n} F(\mathbf{x}) = \sum_{J \subset \{1,\dots,n\}} (-1)^{n-\#(J)} F\left(\mathbf{x} + \sum_{j \in J} \delta_{y_j}\right), \tag{2.13}$$

where $\#(J)$ denotes the number of elements in J. It shows that $D^n_{y_1,\dots,y_n} F$ is a symmetric mapping in $y_1 \dots, y_n \in \mathbb{X}$ and that $(\mathbf{x}, y_1, \dots, y_n) \mapsto D^n_{y_1,\dots,y_n} F(\mathbf{x})$ is measurable.

Notation. For all $F : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ and $n \in \mathbb{N}$, we define the expectation of the n-th order difference operator of F as a function T_nF on $\mathbb{X}^n$, where

$$T_nF(y_1, \dots, y_n) = \mathbb{E}[D^n_{y_1,\dots,y_n} F(\eta)]$$

and set $T_0F = \mathbb{E}[F(\eta)]$, whenever these expectations exist. Otherwise, we put $T_nF(y_1, \dots, y_n) = 0$. Note that the mapping $T_nF : \mathbb{X}^n \to \mathbb{R}$ is again symmetric and measurable.

Definition 2.8 (Wiener–Itô integral).
For $n \geqslant 1$ and $g \in L^2(\lambda^n)$, we define the *n-th order Wiener–Itô integral of g* as a random variable $I_n(g)$ defined by

$$I_n(g) := \sum_{J \subset \{1,\dots,n\}} (-1)^{n-\#(J)} \int \int g(x_1, \dots, x_n) \eta^{(\#(J))}(\mathrm{d}x_J) \lambda^{n-\#(J)}(\mathrm{d}x_{J^c}),$$

where $x_J := (x_j)_{j \in J}$ and $J^c := \{1, \dots, n\} \backslash J$.

Remark. The Slivnyak–Mecke equation (1.9) combined with Fubini's theorem implies that the integrals in Definition 2.8 are finite and $\mathbb{E}\, I_n(g) = 0$.

If g is, moreover, symmetric (i.e. invariant towards any permutation of the variables), then

$$I_n(g) = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} \int g \mathrm{d}\eta^{(k)} \otimes \lambda^{n-k}.$$

See Chapter 12 in Last and Penrose [2017] for further properties of $I_n(g)$.

The following result based on Itô [1956] and Wiener [1938] is an analogy to an orthogonal expansion into series of polynomials for a square-integrable function of a real variable. It was proved in this form for Poisson functionals in Last and Penrose [2011], Theorem 1.3.

Theorem 2.5 (Wiener–Itô chaos expansion).
Let $f \in L^2(P_\eta)$. Then for all $n \in \mathbb{N}$, $T_nf \in L^2(\lambda^n)$ and

$$f(\eta) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(T_nf),$$

where the series converges in $L^2(\mathbb{P})$. Moreover, if for $n \in \mathbb{N}$, $g_n \in L^2(\lambda^n)$ is a symmetric function such that $f(\eta) = \sum_{n=0}^{\infty} \frac{1}{n!} I_n(g_n)$ with convergence in $L^2(\mathbb{P})$, then $g_0 = \mathbb{E}\, f(\eta)$ and $g_n = T_nf$ λ^n-a.e. on $\mathbb{X}$ and for all $n \in \mathbb{N}$.

For $F \in L^2_\eta(\mathbb{P})$ denote $f_n := \frac{1}{n!}\mathbb{E}\, D^n F$. Then by Theorem 2.5, F can be written as

$$F = \mathbb{E}\, F + \sum_{n=1}^{\infty} I_n(f_n). \tag{2.14}$$

For many models in stochastic geometry, the Wiener–Itô chaotic expansion and associated operators from Malliavin calculus could be further analysed because the expressions usually consist of terms with a natural geometric interpretation.

Example 2.2 (Poisson U-statistic).
Let $n \in \mathbb{N}$ and $f \in L^1(\lambda^n)$ symmetric. Then we define the *Poisson U-statistic of order n* with *kernel function f* by

$$U(f,\eta) := \int f(x_1,\dots,x_n)\eta^{(n)}(\mathrm{d}(x_1,\dots,x_n)) = \sum_{x_1,\dots,x_n\in\eta}^{\neq} f(x_1,\dots,x_n),$$

where in the last expression, we sum over all mutually different points $x_1,\dots,x_n$ of η.

For $k \in \{0,\dots,n\}$ define symmetric functions $f_k \in L^1(\lambda^k)$ by

$$f_k(x_1,\dots,x_k) := \binom{n}{k}\int f(x_1,\dots,x_k,y_1,\dots,y_{n-k})\lambda^{n-k}(\mathrm{d}(y_1,\dots,y_{n-k})).$$

Then the U-statistic with kernel function f admits the following representation (Proposition 12.11 in Last and Penrose [2017]):

$$U = \mathbb{E}\,[U] + \sum_{k=1}^{n} I_k(f_k), \quad \mathbb{P} - a.s. \tag{2.15}$$

The next definition introduces three Malliavin-type operators defined in terms of the chaotic expansions that are involved in the estimates of the distances between Poisson functionals and standard normal distribution (recall Theorems 2.2, 2.3 and 2.4).

Definition 2.9 (Malliavin operators).
We define the operators D, δ, L on $L^2_\eta(\mathbb{P})$ as follows:

1. **Derivative operator D:** Define the *domain* of D by dom$D$ as the set of all $F \in L^2_\eta(\mathbb{P})$ admitting a chaotic decomposition (2.14) such that

$$\sum_{n=1}^{\infty} n n!\|f_n\|_n^2 < \infty. \tag{2.16}$$

   Then for $F \in \mathrm{dom}D$, the random function $x \mapsto D_x F$ is given by

$$D_x F = \sum_{n=1}^{\infty} n I_{n-1}(f(x,\cdot)), \quad x \in \mathbb{X},$$

   where for symmetric function $f \in L^2(\lambda^n)$, $f(x,\cdot)$ indicates the function on $\mathbb{X}^{n-1}$ given by $(x_1,\dots,x_{n-1}) \mapsto f(x,x_1,\dots,x_{n-1})$.

2. **Skorohod integral** δ: For a function $G \in L^2_\eta(\mathbb{P}\otimes\lambda)$, we have that $G(x) := G(\cdot, x) \in L^2(\mathbb{P})$ and hence, from Theorem 2.5

$$G(x) = \sum_{n=0}^{\infty} I_n(g_n(x,\cdot)), \quad \mathbb{P}-a.s.,$$

where $g_n(x, x_1, \ldots, x_n) := \frac{1}{n!}\mathbb{E}\, D^n_{x_1,\ldots,x_n} G(x)$. The *Skorohod integral of* G is defined by

$$\delta(G) := \sum_{n=0}^{\infty} I_{n+1}(g_n).$$

The domain domδ of δ is defined as the set of functions $G \in L^2_\eta(\mathbb{P}\otimes\lambda)$ satisfying

$$\sum_{n=0}^{\infty}(n+1)! \int \tilde{g}_n^2 \mathrm{d}\lambda^{n+1} < \infty,$$

where $\tilde{g}_n$ is the symmetrization of g_n, i.e.

$$\tilde{g}_n(x_1, \ldots, x_{n+1}) := \frac{1}{(n+1)!}\sum_{i=1}^{n+1} \mathbb{E}\, D^n_{x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_{n+1}} G(x_i).$$

3. **Ornstein–Uhlenbeck generator** L**:** The domain of L, dom$L$, is defined as the set of functions $F \in L^2_\eta(\mathbb{P})$ with the chaotic decomposition (2.14) satisfying

$$\sum_{n=1}^{\infty} n^2 n! \|f_n\|_n^2 < \infty.$$

For $F \in \mathrm{dom}L$, we define the *Ornstein–Uhlenbeck generator of* L by

$$LF := -\sum_{n=1}^{\infty} nI_n(f_n).$$

The *inverse of* L is given by

$$L^{-1}F := -\sum_{n=0}^{\infty} \frac{1}{n} I_n(f_n).$$

Let us mention several properties of the Malliavin operators:

1. Note that the derivative operator D transforms random variables into random functions. The condition (2.16) guarantees that the Malliavin derivative of F coincides with the difference operator DF (see e.g. Theorem 6.2 in Nualart and Vives [1990] for the proof).

2. The operator δ is the adjoint of the difference operator D and the following formula holds.

Theorem 2.6 (Integration by parts formula).
Let $F \in \mathrm{dom}D$ *and* $G \in \mathrm{dom}\delta$*. Then,*

$$\mathbb{E} \int (D_x F) G(x) \lambda(\mathrm{d}x) = \mathbb{E}\, F\delta(G).$$

See Theorem 4 in Last [2016] for the proof or Proposition 4.2 in Nualart and Vives [1990] for more general result.

3. The random variable $L^{-1}F$ is well defined for all $F \in L^2_\eta(\mathbb{P})$, $L^{-1}F \in \text{dom}L$ and if $\mathbb{E}\, F = 0$, then $LL^{-1}F = F$.

4. The operators D, δ and L are connected by the following identity.

   Theorem 2.7 (Proposition 3 in Last [2016]).
   Let $F \in \text{dom}L$. Then $F \in \text{dom}D, DF \in \text{dom}\delta$ and

   $$\delta(DF) = -LF.$$

Fock space representation

Definition 2.10 (Fock space).
Let $\mathbf{H}_n$ denote the space of all λ^n-a.e. symmetric functions $f \in L^2(\lambda^n)$ equipped with the inner product $\langle \cdot, \cdot \rangle_n$ and the corresponding norm $||f||_n$. We define the *Fock space* $\mathbf{H}$ as the space of all sequences $f = (f_n)_{n \geqslant 0}, f_n \in \mathbf{H}_n$, i.e. as the direct product of the spaces $\mathbf{H}_n$, i.e.

$$\mathbf{H} = \bigtimes_{n=0}^{\infty} \mathbf{H}_n.$$

with the scalar product defined by

$$\langle f, g \rangle_{\mathbf{H}} = \sum_{i=0}^{\infty} \frac{1}{n!} \langle f_n, g_n \rangle_n, \quad f, g \in \mathbf{H}.$$

Note that $\mathbf{H}$ is a Hilbert space, i.e. a complete metric space with respect to the metric $((u_n)_{n \geqslant 0}, (v_n)_{n \geqslant 0}) \mapsto \sqrt{\langle (u_n - v_n)_{n \geqslant 0}, (u_n - v_n)_{n \geqslant 0} \rangle_{\mathbf{H}}}$. An extensive treatment of the stochastic calculus on $\mathbf{H}$ can be found in Meyer [1995]. Our goal is to prove that the linear mapping $f \mapsto (T_n(f))_{n \geqslant 0}$ is an isometry from $L^2(\lambda^n)$ into the Fock space $\mathbf{H}$. Then it follows from Theorem 2.5 and isometry properties of stochastic integrals that the mapping is in fact a bijection from $L^2(P_\eta)$ to the Fock space.

Theorem 2.8 (Fock space representation).
Let $f, g \in L^2(P_\eta)$. Then $Tf := (T_n f)_{n \geqslant 0} \in \mathbf{H}$ and

$$\mathbb{E}[f(\eta) g(\eta)] = \langle Tf, Tg \rangle_{\mathbf{H}}.$$

Remark. In other words, Theorem 2.8 says that

$$\mathbb{E}\,[f(\eta)g(\eta)] = (\mathbb{E}\,[f(\eta)])\,(\mathbb{E}\,[g(\eta)]) + \sum_{n=1}^{\infty} \frac{1}{n!} \langle T_n f, T_n g \rangle_n . \tag{2.17}$$

In particular,

$$\mathbb{E}\,[f(\eta)^2] = (\mathbb{E}\,[f(\eta)])^2 + \sum_{n=1}^{\infty} \frac{1}{n!} ||T_n f||_n^2. \tag{2.18}$$

Proof. We will follow the proofs of Lemma 18.3, Lemma 18.4, Lemma 18.5 and Theorem 18.6 of Last and Penrose [2017]. The proof consists of four individual steps. In prior to show that the equality holds for arbitrary $f, g \in L^2(P_\eta)$, we will prove it for special space of bounded and measurable functions, which will be proved to be dense in $L^2(P_\eta)$. Then we apply some approximation arguments to prove the theorem.

Step 1 Let $\mathcal{X}_0$ be the system of all measurable sets $B \in \mathcal{B}(\mathbb{X})$ for which $\lambda(B) < \infty$. Denote by $\mathbb{R}_0(\mathbb{X})$ the space of all bounded functions $v : \mathbb{X} \to \mathbb{R}_+$ vanishing outside some $B \in \mathcal{X}_0$. Furthermore, denote by $\mathbf{G}$ the space of all (bounded and measurable) functions $g : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ of the form

$$g(\mu) = a_1 e^{-\mu(v_1)} + \ldots + a_n e^{-\mu(v_n)},$$

where $n \in \mathbb{N}, a_1, \ldots, a_n \in \mathbb{R},\ v_1, \ldots, v_n \in \mathbb{R}_0(\mathbb{X})$ and $\mu(v)$ denotes the integral $\int v \mathrm{d}\mu$ for $\mu \in \mathbf{N}(\mathbb{X})$ and $v \in \mathbb{R}_0(\mathbb{X})$. Let us show that equality (2.17) holds for $f, g \in \mathbf{G}$.

By linearity, it is sufficient to consider functions f and g of the form

$$f(\mu) = \exp[-\mu(v)], \quad g(\mu) = \exp[-\mu(w)]$$

for $v, w \in \mathbb{R}_0(\mathbb{X})$. First, we will calculate $T_n f$ and $T_n g$ for $n \in \mathbb{N}$. For each $\mu \in \mathbf{N}(\mathbb{X})$ and $x \in \mathbb{X}$, we have

$$f(\mu + \delta_x) = \exp\left[-\int_{\mathbb{X}} v(y)(\mu + \delta_x)(\mathrm{d}y)\right] = \exp[-\mu(v)]\exp[-v(x)],$$

and therefore,

$$D_x f(\mu) = \exp[-\mu(v)](\exp[-v(x)] - 1).$$

Iterating this identity, we can get for all $n \in \mathbb{N}$ and all $x_1, \ldots, x_n \in \mathbb{X}$ that

$$D^n_{x_1,\ldots,x_n} f(\mu) = \exp[-\mu(v)] \prod_{i=1}^{n} (\exp[-v(x_i)] - 1). \tag{2.19}$$

Recall that for the Poisson point process η with intensity measure λ, the Laplace functional (see Definition 1.12) takes form

$$L_\eta(u) = \exp\left[-\lambda(1 - e^{-u})\right], \quad u : \mathbb{X} \to \mathbb{R}_+. \tag{2.20}$$

From (2.19) and (2.20), we obtain that

$$T_n f = \exp[-\lambda(1 - e^{-v})] \prod_{i=1}^{n} (\exp[-v(x_i)] - 1).$$

Analogously for g. Since $v, w \in \mathbb{R}_0(\mathbb{X})$ it follows that $T_n f, T_n g \in \mathbf{H}_n, n \geqslant 0$. Using again equality (2.20), we obtain that

$$\mathbb{E}\left[f(\eta)g(\eta)\right] = \exp[-\lambda(1 - e^{-(v+w)})].$$

Now, we can compute the right-hand side of (2.17)

$$\sum_{n=1}^{\infty}\frac{1}{n!}\langle T_n f, T_n g\rangle_n$$
$$= \exp[-\lambda(1-e^{-v})]\exp[-\lambda(1-e^{-w})]\sum_{n=1}^{\infty}\frac{1}{n!}\lambda^n(((e^{-v}-1)(e^{-w}-1))^{\otimes n})$$
$$= \exp[-\lambda(2-e^{-v}-e^{-w})]\exp[\lambda((e^{-v}-1)(e^{-w}-1))]$$
$$= \exp[-\lambda(1-e^{-(v+w)})]$$

and hence the assertion holds true for $f, g \in \mathbf{G}$.

Step 2 We need to prove that the set $\mathbf{G}$ is dense in $L^2(P_\eta)$. Let $\mathbf{W}$ be the space of all bounded measurable $g : \mathbf{N}(\mathbb{X}) \to \mathbb{R}$ that can be approximated in $L^2(P_\eta)$ by functions in $\mathbf{G}$. We want to use the functional version of the monotone class theorem (see Theorem 2.12.9 in Bogachev [2007]). We can see that space $\mathbf{G}$ is closed under uniformly bounded convergence. It also contains the constant functions and it is closed under multiplication. If we denote by $\mathcal{N}'$ the smallest σ-field on $\mathbf{N}(\mathbb{X})$ such that $\mu \mapsto h(\mu)$ is measurable for all $h \in \mathbf{G}$, then according to Theorem 2.12.9 in Bogachev [2007], $\mathbf{W}$ contains any bounded $\mathcal{N}'$-measurable g.

On the other hand we can write for every $C \in \mathcal{X}_0$ that

$$\mu(C) = \lim_{t\to 0_+} t^{-1}(1-e^{-t\mu(C)}), \quad \mu \in \mathbf{N}(\mathbb{X}),$$

such that $\mu \mapsto \mu(C)$ is $\mathcal{N}'$-measurable. Since λ is σ-finite, for any $C \in \mathcal{X}$ there exists a monotone sequence $C_k \in \mathcal{X}_0, k \in \mathbb{N}$ such that $C = \cup C_k$, so that $\mu \mapsto \mu(C)$ is $\mathcal{N}'$-measurable. Thus, $\mathcal{N}' = \mathcal{N}$ and it follows that $\mathbf{W}$ contains all bounded measurable functions. Hence $\mathbf{W}$ is dense in $L^2(P_\eta)$.

Step 3 For further purposes we will show that $f, f^1, f^2, \ldots \in L^2(P_\eta)$ satisfying $f^k \to f$ in $L^2(P_\eta)$ as $k \to \infty$ implies

$$\lim_{k\to\infty}\int_{C^n}\mathbb{E}\left[|D^n_{x_1,\ldots,x_n}f(\eta) - D^n_{x_1\ldots,x_n}f^k(\eta)|\right]\lambda^n(\mathrm{d}(x_1,\ldots,x_n)) = 0 \tag{2.21}$$

for all $n \in \mathbb{N}$ and $C \in \mathcal{X}_0$. According to (2.13), it is sufficient to prove

$$\lim_{k\to\infty}\int_{C^n}\mathbb{E}\left[\left|f\left(\eta+\sum_{i=1}^{m}\delta_{x_i}\right) - f^k\left(\eta+\sum_{i=1}^{m}\delta_{x_i}\right)\right|\right]\lambda^n(\mathrm{d}(x_1,\ldots,x_n)) = 0 \tag{2.22}$$

for all $m \in \{0,\ldots,n\}$. The case of $m = 0$ is obvious. Assuming $m \in \{0,\ldots,n\}$, we apply on the integral inside the limit in (2.21) the Slivnyak–Mecke equation (see (1.9)). Thus,

$$
\begin{aligned}
&\int_{C^n} \mathbb{E}\left[\left|f\left(\eta+\sum_{i=1}^m \delta_{x_i}\right)-f^k\left(\eta+\sum_{i=1}^m \delta_{x_i}\right)\right|\right]\lambda^n(\mathrm{d}(x_1,\dots,x_n)) \\
&= \lambda(C)^{n-m}\int_{C^m} \mathbb{E}\left[\left|f\left(\eta+\sum_{i=1}^m \delta_{x_i}\right)-f^k\left(\eta+\sum_{i=1}^m \delta_{x_i}\right)\right|\right]\lambda^m(\mathrm{d}(x_1,\dots,x_m)) \\
&= \lambda(C)^{n-m}\mathbb{E}\left[\int_{C^m} \left|f(\eta)-f^k(\eta)\right|\eta^m(\mathrm{d}(x_1,\dots,x_m))\right] \\
&\leqslant \lambda(C)^{n-m}\mathbb{E}\left[\left|f(\eta)-f^k(\eta)\right|\eta^{(m)}(C^m)\right] \\
&\leqslant \lambda(C)^{n-m}(\mathbb{E}\left[(f(\eta)-f^k(\eta))^2\right])^{\frac{1}{2}}(\mathbb{E}\left[(\eta^{(m)}(C^m))^2\right])^{\frac{1}{2}},
\end{aligned}
$$

where by $\eta^m(C^m)$, we denote the number of m-tuples of points of η in C while $\eta^{(m)}(C^m)$ restricts on m-tuples with mutually distinct points.

The last bound follows from the Cauchy–Schwarz inequality. Since all moments of the Poisson distribution exist, we obtain (2.22) and hence (2.21).

Step 4 Recall the polarization identity of the scalar product

$$4\langle f,g\rangle_{\mathbf{H}} = \langle f+g,f+g\rangle_{\mathbf{H}} - \langle f-g,f-g\rangle_{\mathbf{H}}.$$

Because of the linearity of the scalar product, it is sufficient to show that (2.18) holds to prove the theorem.

Since the system $\mathbf{G}$ is dense in $L^2(P_\eta)$, for every $f \in L^2(P_\eta)$ there is a sequence $f^k \in \mathbf{G}$ such that $f^k \to f$ in $L^2(P_\eta)$ as $k \to \infty$. In step 3, we proved that $Tf^k, k \in \mathbb{N}$, is a Cauchy sequence in $\mathbf{H}$, hence has a limit $\tilde{f} = (\tilde{f}_n) \in \mathbf{H}$, meaning that

$$\lim_{k\to\infty}\sum_{n=0}^{\infty}\frac{1}{n!}||T_n f^k - \tilde{f}_n||_n^2 = 0. \tag{2.23}$$

Taking the limit in the identity $\mathbb{E}[f^k(\eta)^2] = \langle Tf^k, Tf^k\rangle_{\mathbf{H}}$ yields

$$\mathbb{E}[f(\eta)^2] = \langle \tilde{f},\tilde{f}\rangle_{\mathbf{H}}.$$

Equation (2.23) immediately implies that $\tilde{f}_0 = \mathbb{E}[f(\eta)] = T_0 f$. It remains to show that for any $n \geqslant 1$, we have

$$\tilde{f}_n = T_n f, \quad \lambda^n\text{-a.e.} \tag{2.24}$$

Let $C \in \mathcal{X}_0$ and let $B := C^n$. Denote by $(\lambda^n)_B$ the restriction of the measure λ^n to B. By (2.23) $T_n f^k$ converges to f in $L^2(B,(\lambda^n)_B)$ and hence also in $L^1(B,(\lambda^n)_B)$. Meanwhile, by the definition of T_n and the equality (2.21), $T_n f^k$ converges in $L^1(B,(\lambda^n)_B)$ to $T_n f$. Hence the uniqueness of these limits yields $\tilde{f}_n = T_n f \lambda^n$-a.e. on B. Since λ is assumed to be σ-finite, this implies (2.24) and hence the theorem.

□

Covariance identity

The covariance identity is a direct consequence of the Fock space representation theorem. It will be further used to obtain bounds on the Wasserstein distance between the standard normal distribution and distribution of a Poisson functional.

Assume that we have a square integrable Poisson functional F (i.e. $F \in L^2_\eta(\mathbb{P})$) and $t \in [0,1]$. To obtain the covariance identity, we need to introduce an operator P_tF defined by a combination of t-thinning and independent superposition. Then we will be able to rewrite the Fock space series representation as an integral equation involving only the first order difference operator and the operator P_t.

Definition 2.11 (Operator P_t).
Let for $F \in L^1_\eta(\mathbb{P})$ with a representative f define

$$P_tF := \mathbb{E}\left[\int\limits_{\mathbf{N}(\mathbb{X})} f(\eta_t + \mu)\Pi_{(1-t)\lambda}(\mathrm{d}\mu)\middle|\eta\right], \quad t \in [0,1],$$

where η_t is a t-thinning of η and $\Pi_{\lambda'}$ denotes the distribution of a Poisson process with intensity measure λ'.

Lemma 2.2.
For $F \in L^1_\eta(\mathbb{P})$, the definition of P_tF almost surely does not depend on the choice of the representative f and

$$\mathbb{E}P_tF = \mathbb{E}F. \tag{2.25}$$

Consequently, $P_tF \in L^1_\eta(\mathbb{P})$, whenever $F \in L^1_\eta(\mathbb{P})$.

Proof. The first statement follows directly from the application of the superposition and thinning theorems (see Theorem 3.3 and Corollary 5.9 in Last and Penrose [2017]), since

$$\Pi_\lambda = \mathbb{E}\left[\int\limits_{\mathbf{N}(\mathbb{X})} \mathbf{1}\{\eta_t + \mu \in \cdot\}\Pi_{(1-t)\lambda}(\mathrm{d}\mu)\right].$$

Due to Lemma B.16 in Last and Penrose [2017], which is based on the monotone convergence theorem, there exists a measurable version of $\mathbb{E}\left[f(\eta_t + \mu)|\eta\right]$ such that

$$P_tF = \mathbb{E}\left[\int\limits_{\mathbf{N}(\mathbb{X})} f(\eta_t + \mu)\Pi_{(1-t)\lambda}(\mathrm{d}\mu)\middle|\eta\right] = \int\limits_{\mathbf{N}(\mathbb{X})} \mathbb{E}\left[f(\eta_t + \mu)|\eta\right]\Pi_{(1-t)\lambda}(\mathrm{d}\mu),$$

$\mathbb{P}$-a.s. for all $t \in [0,1]$. We can also see that

$$P_tF = \mathbb{E}\left[f(\eta_t + \eta'_{1-t})|\eta\right] \tag{2.26}$$

where η'_{1-t} is a Poisson process with intensity measure $(1-t)\lambda$, independent of the pair (η, η_t). The equality (2.25) follows from (2.26).

□

Using the conditional version of the Jensen inequality (Proposition B.1 in Last and Penrose [2017]) and equality (2.25), we can determine an estimate for the p-th absolute moment of P_tF.

Lemma 2.3 (Contractivity property).
For any $p \geqslant 1$, $F \in L^p_\eta(\mathbb{P})$ *and* $t \in [0,1]$, *we have*

$$\mathbb{E}\left[|P_tF|^p\right] \leqslant \mathbb{E}\left[|F|^p\right].$$

Proof. Let f be a representative of F and denote $g = |f|^p, G = g(\eta)$. Then,

$$\begin{aligned}\mathbb{E}\left[|P_tF|^p\right] &= \mathbb{E}\left|\mathbb{E}\left[f(\eta_t + \eta'_{1-t})|\eta\right]\right|^p \leqslant \mathbb{E}\left[\mathbb{E}\left[|f(\eta_t + \eta'_{1-t})|^p|\eta\right]\right] \\ &= \mathbb{E}\left[\mathbb{E}\left[g(\eta_t + \eta'_{1-t})|\eta\right]\right] = \mathbb{E}\left[P_tG\right] = \mathbb{E}\left[G\right] = \mathbb{E}\left[|F|^p\right].\end{aligned}$$

□

Lemma 2.4 (Mehler's formula).
Let $F \in L^2_\eta(\mathbb{P})$, $n \in \mathbb{N}$ *and* $t \in [0,1]$. *Then*

$$D^n_{x_1,\dots,x_n}(P_tF) = t^n P_t D^n_{x_1,\dots,x_n}F, \quad \lambda^n\text{-}a.a.\ (x_1,\dots,x_n) \in \mathbb{X}^n,\ \mathbb{P} - a.s.$$

In particular,

$$\mathbb{E}\left[D^n_{x_1,\dots,x_n}P_tF\right] = t^n\mathbb{E}\left[D^n_{x_1,\dots,x_n}F\right], \quad \lambda^n\text{-}a.a.\ (x_1,\dots,x_n) \in \mathbb{X}^n.$$

For the proof of Lemma 2.4, see Lemma 20.1 in Last and Penrose [2017].

Notation. Let for $F \in L^2_\eta(\mathbb{P})$ denote by DF the mapping $(\omega, x) \mapsto (D_xF)(\omega)$. The next theorem will additionally require $DF \in L^2(\Omega \times \mathbb{X}, \mathbb{P} \otimes \lambda)$, i.e.

$$\mathbb{E}\left[\int_{\mathbb{X}} (D_xF)^2\lambda(\mathrm{d}x)\right] < \infty.$$

Theorem 2.9 (Covariance Identity).
For any $F, G \in L^2_\eta(\mathbb{P})$ *such that* $DF, DG \in L^2(\Omega \times \mathbb{X}, \mathbb{P} \otimes \lambda)$, *we have*

$$\mathbb{E}\left[FG\right] - \mathbb{E}\left[F\right]\mathbb{E}\left[G\right] = \mathbb{E}\left[\int_{\mathbb{X}}\int_0^1 (D_xF)(P_tD_xG)\mathrm{d}t\lambda(\mathrm{d}x)\right]. \tag{2.27}$$

Proof. We follow the proof of Theorem 20.2 in Last and Penrose [2017]. Using first the Cauchy–Schwarz inequality and then the contractivity property (Lemma 2.3) we can estimate

$$\left(\mathbb{E}\left[\int_{\mathbb{X}}\int_0^1 |D_xF||P_tD_xG|\mathrm{d}t\lambda(\mathrm{d}x)\right]\right)^2$$
$$\leqslant \mathbb{E}\left[\int_{\mathbb{X}}(D_xF)^2\lambda(\mathrm{d}x)\right]\mathbb{E}\left[\int_{\mathbb{X}}\int_0^1(P_tD_xG)^2\mathrm{d}t\lambda(\mathrm{d}x)\right]$$
$$\leqslant \mathbb{E}\left[\int_{\mathbb{X}}(D_xF)^2\lambda(\mathrm{d}x)\right]\mathbb{E}\left[\int_{\mathbb{X}}(D_xG)^2\lambda(\mathrm{d}x)\right],$$

which is finite by the assumption. Therefore, using Fubini's theorem and Mehler's formula (Lemma 2.4), we obtain that the right-hand side of (2.27) equals

$$\int_{\mathbb{X}}\int_0^1 t^{-1}\mathbb{E}\left[(D_xF)(P_tD_xG)\right]\mathrm{d}t\lambda(\mathrm{d}x). \tag{2.28}$$

We can now apply the Fock space representation (Theorem 2.8) to the expectation inside the integral. For $t \in [0,1]$ and taking into account also Lemma 2.4, we obtain

$$\mathbb{E}\left[(D_xF)(D_xP_tG)\right] = t\mathbb{E}\left[D_xF\right]\mathbb{E}\left[D_xG\right] + \sum_{i=1}^{\infty}\frac{t^{n+1}}{n!}\int_{\mathbb{X}^n}\mathbb{E}\left[D^{n+1}_{x_1,\dots,x_n,x}F\right]\mathbb{E}\left[D^{n+1}_{x_1,\dots,x_n,x}G\right]\lambda^n(\mathrm{d}(x_1,\dots,x_n)).$$

We want to insert this expression into formula (2.28) and use Fubini's theorem (to be justified below). Compute (2.28) as

$$\begin{aligned}
&\int_{\mathbb{X}}\int_0^1 \mathbb{E}\left[D_xF\right]\mathbb{E}\left[D_xG\right]\mathrm{d}t\lambda(\mathrm{d}x)\\
&\quad+\sum_{n=1}^{\infty}\int_0^1\frac{t^n}{n!}\mathrm{d}t\int_{\mathbb{X}}\int_{\mathbb{X}^n}\mathbb{E}\left[D^{n+1}_{x_1,\dots,x_n,x}F\right]\mathbb{E}\left[D^{n+1}_{x_1,\dots,x_n,x}G\right]\lambda^n(\mathrm{d}(x_1,\dots,x_n))\lambda(\mathrm{d}x)\\
&=\int_{\mathbb{X}}\mathbb{E}\left[D_xF\right]\mathbb{E}\left[D_xG\right]\lambda(\mathrm{d}x)\\
&\quad+\sum_{n=1}^{\infty}\frac{1}{(n+1)!}\int_{\mathbb{X}}\int_{\mathbb{X}^n}\mathbb{E}\left[D^{n+1}_{x_1,\dots,x_n,x}F\right]\mathbb{E}\left[D^{n+1}_{x_1,\dots,x_n,x}G\right]\lambda^n(\mathrm{d}(x_1,\dots,x_n))\lambda(\mathrm{d}x)\\
&=\sum_{n=1}^{\infty}\frac{1}{n!}\int_{\mathbb{X}^n}\mathbb{E}\left[D^{n}_{x_1,\dots,x_n}F\right]\mathbb{E}\left[D^{n}_{x_1,\dots,x_n}G\right]\lambda^n(\mathrm{d}(x_1,\dots,x_n)).
\end{aligned}$$

Eventually, by Theorem 2.8, this equals to $\mathbb{E}\left[FG\right]-\mathbb{E}\left[F\right]\mathbb{E}\left[G\right]$, which yields the asserted formula (2.27). The use of Fubini's theorem is justified by identity (2.18) and the Cauchy–Schwarz inequality.

□

A direct consequence of Theorem 2.9 is the following upper bound for the variance. More general variance inequalities were developed in Last and Penrose [2011].

Corollary 2.1 (Poincaré inequality).
Let $F \in L^2(\mathbb{P})_\eta$ *and* $DF \in L^2(\Omega \times \mathbb{X}, \mathbb{P} \otimes \lambda)$*, then*

$$\operatorname{Var} F \leqslant \mathbb{E} \int (D_x F)^2 \lambda(\mathrm{d}x).$$

Remark. Corollary 2.1 forms a Poisson version of the famous *Chernoff–Nash–Poincaré inequality* formulated in the framework of Gaussian analysis stating that for $X = (X_1, \dots, X_d)$ an i.i.d. standard Gaussian vector and f being a smooth function on $\mathbb{R}^d$, we have that

$$\operatorname{Var} f(X) \leqslant \mathbb{E} \|\nabla f(X)\|^2,$$

where ∇f is the gradient of f.

2.3 Normal approximation of Poisson functionals

In this section, we will demonstrate how the bounds of Theorems 2.2, 2.3 and 2.4 can be combined with the Malliavin operators (see Definition 2.9). The connection of Stein's method and Malliavin calculus was first mentioned in Nourdin and Peccati [2009a] in order to derive explicit bounds in the Gaussian and Gamma approximations of random variables in a fixed Wiener chaos of a general Gaussian process. Later in Peccati et al. [2010], the theory was first formulated in the framework of point measures.

Recall that η is assumed to be a Poisson point process with locally finite intensity measure λ and distribution P_η and Z stands for the standard Gaussian random variable.

Theorem 7 and Theorem 8 in Bourguin and Peccati [2016] give bounds on the Wasserstein and Kolmogorov distances in the language of the Malliavin operators D and L^{-1}:

Theorem 2.10 (Bounds on the Wasserstein distance).
Let $F \in \operatorname{dom} D$ *be such that* $\mathbb{E} F = 0$*. Then*

$$d_W(F, Z) \leqslant \sqrt{\frac{2}{\pi}} \mathbb{E} \left| 1 - \left\langle DF, -DL^{-1}F \right\rangle_{L^2(\lambda)} \right| + \int \mathbb{E} \left[|D_x F|^2 |D_x L^{-1} F| \right] \lambda(\mathrm{d}x).$$

Theorem 2.11 (Bounds on the Kolmogorov distance).
Let $F \in \operatorname{dom} D$ *be such that* $\mathbb{E} F = 0$*. Then*

$$\begin{aligned} d_K(F, Z) \leqslant{}& \mathbb{E} \left| 1 - \langle DF, -DL^{-1}F \rangle_{L^2(\lambda)} \right| + \frac{\sqrt{2\pi}}{8} \mathbb{E} \left[\langle |DF|^2, |DL^{-1}F| \rangle \right] \\ &+ \frac{1}{2} \mathbb{E} \left[\langle |DF|^2, |F \times DL^{-1}F| \rangle \right] \\ &+ \sup_{z \in \mathbb{R}} \mathbb{E} \left[\langle (DF) D\mathbf{1}_{[F>z]}, |DL^{-1}F| \rangle_{L^2(\lambda)} \right], \end{aligned}$$

where $D_x\mathbf{1}_{[F>z]} = \mathbf{1}_{[F_x>z]} - \mathbf{1}_{[F>z]}, x \in \mathbb{X}$.

Working with the operator L^{-1} can be rather difficult. One option is to use the chaotic expansion (2.14) of $L^{-1}F$. It was illustrated in Lachièze-Rey and Reitzner [2016] and Schulte and Thäle [2016] for the case where F is a Poisson U-statistic. In particular, it was stated that such F lives in a finite sum of Wiener chaoses.

Another possibility is to use the following representation of $L^{-1}F$ in terms of the operator P_t (see Definition 2.11): For $F \in L^2_\eta(\mathbb{P})$, we have

$$L^{-1}F = -\int_0^1 \frac{1}{s} P_s F \mathrm{d}s.$$

In combination with the covariance identity (Theorem 2.9), one can obtain the following bounds for the Wasserstein distance.

Theorem 2.12 (Theorem 21.2 in Last and Penrose [2017]).
Let $F \in L^2_\eta(\mathbb{P})$ *satisfy* $DF \in L^2(\Omega \times \mathbb{X}, \mathbb{P} \otimes \lambda)$ *and* $\mathbb{E}[F] = 0$. *Then*

$$\begin{aligned} d_W(F, Z) \leqslant &\mathbb{E}\left[\left|1 - \int_{\mathbb{X}} \int_0^1 (P_t D_x F)(D_x F) \mathrm{d}t \lambda(\mathrm{d}x)\right|\right] \\ &+ \mathbb{E}\left[\int_{\mathbb{X}} \int_0^1 |P_t D_x F| (D_x F)^2 \mathrm{d}t \lambda(\mathrm{d}x)\right]. \end{aligned}$$

Nevertheless, the operator P_t might still be difficult to manage. One can again use the covariance identity (Theorem 2.9) and the contractivity property (Lemma 2.3) and determine a bound depending only on the random functions DF and D^2F, which could be evaluated directly for some simple choices of the Poisson functionals (see Example 2.3). The following result was shown in Last et al. [2016].

Theorem 2.13 (Second order Poincaré inequality).
Suppose that $F \in L^2_\eta(\mathbb{P})$ *satisfies* $DF \in L^2(\Omega \times \mathbb{X}, \mathbb{P} \otimes \lambda)$, $\mathbb{E}[F] = 0$ *and moreover,* $\mathrm{Var}[F] = 1$. *Denote*

$$\alpha_{F,1} := 2\left[\int_{\mathbb{X}^3} (\mathbb{E}[(D_{x_1}F)^2(D_{x_2}F)^2])^{1/2} (\mathbb{E}[\Delta_{x_1,x_2,x_3}(F)])^{1/2} \lambda^3(\mathrm{d}(x_1,x_2,x_3))\right]^{1/2},$$

$$\alpha_{F,2} := \left[\int_{\mathbb{X}^3} \mathbb{E}[\Delta_{x_1,x_2,x_3}(F)] \lambda^3(\mathrm{d}(x_1,x_2,x_3))\right]^{1/2},$$

$$\alpha_{F,3} := \int_{\mathbb{X}} \mathbb{E}[|D_x F|^3] \lambda(\mathrm{d}x),$$

where $\Delta_{x_1,x_2,x_3}(F) = (D^2_{x_1,x_3}F)^2 (D^2_{x_2,x_3}F)^2$. *Then the upper bound on the Wasserstein distance can be expressed in terms of the constants* $\alpha_{F,1}, \alpha_{F,2}, \alpha_{F,3}$ *as*

$$d_W(F, Z) \leqslant \alpha_{F,1} + \alpha_{F,2} + \alpha_{F,3}.$$

Theorem 2.14 (Theorem 1.2 in Last et al. [2016]).
Assume the assumptions of Theorem 2.13 hold and denote, moreover,

$$\alpha_{F,4} := \frac{1}{2}\left[\mathbb{E}\,F^4\right]^{1/4}\int\limits_{\mathbb{X}}\left[\mathbb{E}\,(D_xF)^4\right]^{3/4}\lambda(\mathrm{d}x),$$

$$\alpha_{F,5} := \left[\int\limits_{\mathbb{X}}\mathbb{E}\,(D_xF)^4\lambda(\mathrm{d}x)\right]^{1/2},$$

$$\alpha_{F,6} := \left[\int\limits_{\mathbb{X}^2} 6\left[\mathbb{E}\,(D_{x_1}F)^4\right]^{1/2}\left[\mathbb{E}\,(D^2_{x_1,x_2}F)^4\right]^{1/2} + 3\mathbb{E}\,(D^2_{x_1,x_2}F)^4\lambda^2(\mathrm{d}(x_1,x_2))\right]^{1/2}.$$

Then

$$d_K(F,Z) \leqslant \alpha_{F,1}+\alpha_{F,2}+\alpha_{F,3}+\alpha_{F,4}+\alpha_{F,5}+\alpha_{F,6}.$$

Remark. The contents of Theorem 2.13 and Theorem 2.14 form a starting point of the *method of stabilization* presented in the subsequent chapter.

Example 2.3 (CLT for non-homogeneous Poisson processes).
Let η be a Poisson point process on $\mathbb{R}_+$, whose intensity measure λ satisfies $0 < \lambda([0,t]) < \infty$ for all sufficiently large t and $\lambda([0,\infty)) = \infty$.

We will define Poisson functionals F_t, $t > 0$ as the normalized difference between the actual number of points of point process η in the interval $[0,t]$ and the expected number of points in this interval, i.e.

$$F_t(\eta) = \frac{\eta([0,t]) - \lambda([0,t])}{\sqrt{\lambda([0,t])}}.$$

We want to use Theorem 2.13 to prove the central limit theorem. First, we have to verify its assumptions. We can observe that all moments of F_t exist, since Poisson distribution has all moments finite. Furthermore, since

$$\mathbb{E}\left[\eta([0,t])\right] = \mathrm{Var}\left[\eta([0,t])\right] = \lambda([0,t]),$$

the assumptions on the variance and the expectation are evidently satisfied.

Take an arbitrary point $x \in \mathbb{R}_+$. Then for the difference operator of the functional F_t, it holds from the definition that

$$D_xF_t(\eta) = \frac{(\eta+\delta_x)([0,t]) - \eta([0,t])}{\sqrt{\lambda([0,t])}} = \frac{\mathbf{1}[x \in [0,t]]}{\sqrt{\lambda([0,t])}}.$$

The difference operator of F_t is no longer random, which implies that the assumption of square integrability of DF holds and moreover, the difference operators of the higher orders are zero.

It remains to plug the difference operator of F_t into the formulae for the constants $\alpha_{F,1}, \alpha_{F,2}, \alpha_{F,3}$ in Theorem 2.13, i.e.

$$\alpha_{F,1} = 0,$$
$$\alpha_{F,2} = 0,$$
$$\alpha_{F,3} = \int \mathbb{E}[|D_xF|^3]\lambda(\mathrm{d}x) = \frac{1}{(\lambda([0,t]))^{\frac{3}{2}}}\int \mathbf{1}[x \in [0,t]]\lambda(\mathrm{d}x) = \frac{1}{\sqrt{\lambda([0,t])}}.$$

Thus,

$$d_W(F(t), Z) \leqslant \frac{1}{\sqrt{\lambda([0,t])}}. \tag{2.29}$$

The right-hand side of (2.29) tends to zero as t goes to infinity, hence $F(t) \xrightarrow{\mathcal{D}} Z$ as $t \to \infty$.

2.4 Selected asymptotic results in stochastic geometry

In this section, we present several asymptotic results from the literature that are based on Malliavin–Stein's method which was discussed in the previous sections.

Example 1: Normal approximation and CLT in the Boolean model

Recall that the Boolean model (Example 1.14) is a random set

$$\Xi = \bigcup_{i \geqslant 1} (x_i + \Xi_i),$$

where $\eta_m = \{(x_i, \Xi_i), i \geqslant 1\}$ is a stationary independently marked Poisson point process on $\mathbb{R}^d$ with marks in the space $C^{(d)}$ of non-empty compact sets in $\mathbb{R}^d$ and the mark distribution $\mathbb{Q}$. Let $\lambda \in (0, \infty)$ be the intensity of the unmarked Poisson point process η on $\mathbb{R}^d$ and $\alpha_m = \lambda|\cdot|_d \otimes \mathbb{Q}$ the intensity measure of η_m.

Moreover, we assume that $\mathbb{Q}$ satisfies the integrability condition

$$\int |K \oplus B_r(\mathbf{o})|_d \mathbb{Q}(\mathrm{d}K) < \infty, \quad \forall r \geqslant 0. \tag{2.30}$$

Condition (2.30) guarantees that Ξ is a random element of the space $\mathcal{F}^d$ of closed subsets of $\mathbb{R}^d$ (Proposition 17.5 in Last and Penrose [2017]).

Let W be a fixed observation window with $|W|_d \in (0, \infty)$. If $\mathbb{Q}$ is concentrated on convex sets, then by observing Ξ in W, we see a finite union of convex sets. Hence, $\Xi \cap W$ is amendable to additive translation-invariant functionals φ, such as intrinsic volumes (Example 2.4 below).

The aim is to study random variables of the type

$$\varphi(\Xi \cap rW),$$

where φ is a suitable geometric function defined on compact sets and $r > 0$ is large. First, let $\varphi = |\cdot|_d$. Then $F_W := |\Xi \cap W|_d$ is a Poisson functional with $\mathbb{E}\, F_W = p|W|_d$, where

$$p = \mathbb{P}(\mathbf{o} \in \Xi) < 1$$

is the *volume fraction* of Ξ. The following Lemma gives the asymptotic variance of F_{rW} when $r \to \infty$.

Lemma 2.5 (Proposition 22.1 in Last and Penrose [2017]).
We have that

$$\mathrm{Var}\,[F_W] = (1-p)^2 \int |W \cap (W+x)|_d (e^{\lambda \beta_d(x)} - 1)\mathrm{d}x,$$

where $\beta_d(x) = \int |K \cap (K+x)|_d \mathbb{Q}(\mathrm{d}K)$*. Moreover, if the boundary of* W *denoted by* ∂W *satisfies* $|\partial W|_d = 0$*, then*

$$\lim_{r\to\infty} \frac{1}{|rW|_d} \mathrm{Var}\,[F_{rW}] = (1-p)^2 \int (e^{\lambda\beta_d(x)} - 1)\mathrm{d}x.$$

The following bound on the Wasserstein distance between a suitably normalized version of F_W and the standard normal random variable Z can be proved. For this purpose, we use the following notation:

$$\phi_{d,k} := \int (|K|_d)^k \mathbb{Q}(\mathrm{d}K), \quad k \in \mathbb{N},$$

and

$$c_W := \frac{|W|_d}{(1-p)^2} \left[\int |W \cap (W+x)|_d (e^{\lambda\beta_d(x)} - 1)\mathrm{d}x \right]^{-1}.$$

Theorem 2.15 (Theorem 22.2 in Last and Penrose [2017]).
Let $\hat{F}_w := (\mathrm{Var}\,[F_W])^{-1/2}(F_W - \mathbb{E}\,F_W)$ *and assume that* $\phi_{d,3} < \infty$ *and* $\phi_{d,1} > 0$*. Then*

$$d_W(\hat{F}_W, Z) \leqslant (|W|_d)^{-1/2}[2(\lambda\phi_{d,2})^{3/2}c_W + \lambda^{3/2}\phi_{d,2}c_W + \lambda\phi_{d,3}c_W^{3/2}].$$

Remark. The assumption $\phi_{d,1} > 0$ guarantees that $\mathrm{Var}\,[F_W] > 0$ (see Exercise 22.2 in Last and Penrose [2017]).

Corollary 2.2 (Corollary 22.3 in Last and Penrose [2017]).
Assume that $|\partial W|_d = 0, \phi_{d,1} > 0$ *and* $\phi_{d,3} < \infty$*. Then*

$$\hat{F}_{r^{1/d}W} \xrightarrow{\mathcal{D}} Z, \quad \text{as } r \to \infty.$$

Remark. The rate of convergence in Corollary 2.2 is $r^{-1/2}$.

Under additional assumptions of W being convex and $\mathbb{Q}$ being concentrated on the system $\mathcal{K}^{(d)}$ of compact convex sets with positive Lebesgue measure one can extend Theorem 2.15 and Corollary 2.2 for the class of geometric functions.

Definition 2.12 (Geometric function).
A function $\varphi : \mathcal{C}^d \to \mathbb{R}$ is *geometric* if it is translation-invariant (meaning that $\varphi(K+x) = \varphi(K)$ for all $x \in \mathbb{R}^d$ and $K \in \mathcal{C}^d$), additive and satisfies

$$M(\varphi) := \sup\{|\varphi(K)| : K \in \mathcal{K}^{(d)}, K \subset [-1/2, 1/2]^d\} < \infty.$$

Example 2.4 (Intrinsic volumes).
Define the Minkowski sum of $K, L \subset \mathbb{R}^d$ by $K \oplus L := \{x + y; x \in K, y \in L\}$. The

Steiner formula states that for $K \subset \mathbb{R}^d$ convex, the volume of the set $K \oplus B_r(\mathbf{o})$ (*the parallel set* of K) can be expressed as a polynomial of degree d:

$$|K \oplus B_r(\mathbf{o})|_d = \sum_{j=0}^{d} r^{d-j} \kappa_{d-j} V_j(K),$$

where κ_k is the volume of the k-dimensional unit ball and V_k is the k-th *intrinsic volume*.

For example, if $K \neq \varnothing$, then $V_0(K) = 1$ and $V_d(K) = |K|_d$. Moreover, if K has a non-empty interior, then

$$V_{d-1}(K) = \frac{1}{2}\mathcal{H}_{d-1}(\partial K),$$

the $(d-1)$-dimensional Hausdorff measure of the boundary of K.

Remark. The intrinsic volumes $V_0, \ldots, V_d$ are geometric functions on $\mathcal{C}^d$.

Given a geometric function φ, in Last and Penrose [2017], Theorem 2.13 was applied to the Poisson functional $F_{W,\varphi} := \varphi(\Xi \cap W)$.

Theorem 2.16 (Theorem 22.7 in Last and Penrose [2017]). *Suppose φ is a geometric function such that $\sigma_{W,\varphi} := (\mathrm{Var}\,[F_{W,\varphi}])^{1/2} > 0$. Denote $\bar{V}(K) := |K \oplus B_1(\mathbf{o})|_d$ and assume that*

$$\int \bar{V}(K)^3 \mathbb{Q}(\mathrm{d}K) < \infty.$$

Let $\hat{F}_{W,\varphi} := \sigma_{W,\varphi}^{-1}(F_{W,\varphi} - \mathbb{E}\, F_{W,\varphi})$. Then

$$d_W(\hat{F}_{W,\varphi}, Z) \leqslant c_1 \sigma_{W,\varphi}^{-2} [\bar{V}(W)]^{1/2} + c_2 \sigma_{W,\varphi}^{-3} \bar{V}(W),$$

where c_1, c_2 do not depend on W.

Corollary 2.3 (Theorem 22.8 in Last and Penrose [2017]). *Assume that the conditions of Theorem 2.16 hold and moreover,*

$$\liminf_{r \to \infty} \frac{\sigma_{W,\varphi}^2}{r} > 0.$$

Then there exists $c > 0$ such that

$$d_W(\hat{F}_{W,\varphi}) \leqslant c r^{-1/2}.$$

In particular, $\hat{F}_{W,\varphi} \xrightarrow{\mathcal{D}} Z$.

Example 2: Normal approximation and CLT for geometric U-statistics

Recall that a *Poisson U-statistic* of order k with kernel function $f : \mathbb{X}^k \to \mathbb{R}$ is defined by

$$U(f, \eta) = \int f(x_1, \ldots, x_k) \eta^{(k)}(\mathrm{d}(x_1, \ldots, x_k)),$$

where η is a Poisson point process on $\mathbb{X}$ with non-atomic locally finite intensity measure λ. U-statistics play an important role in stochastic geometry since many interesting functionals can be expressed as U-statistics, for instance the intrinsic volumes. In Example 2.2, it was shown that a U-statistics can be expressed as a finite sum of multiple Wiener–Itô integrals (see the chaotic representation (2.15)). Therefore, to study the asymptotic properties of a U-statistic is equivalent to study the individual Wiener–Itô integrals. The Malliavin operators in this case are usually easy to handle. For instance, if $F = I_k(f)$ is a multiple Wiener–Itô integral of order $k \geqslant 1$ and $f \in L^2(\lambda^k)$, then

$$D_x F = k I_{k-1}(f(x, \cdot)), \quad x \in \mathbb{X}$$

and

$$LF = -k I_k(f), \quad L^{-1} F = -k^{-1} I_k(f).$$

Directly from the chaotic representation (2.15) and Slivnyak–Mecke equation (1.9), one can show the following result.

Proposition 2.5 (Proposition 1 in Lachièze-Rey and Reitzner [2016]). *Let $U(f,\eta)$ be a Poisson U-statistic of order k with symmetric kernel function $f \in L^1(\lambda^k)$. Then $\mathbb{E}\,|U(f,\eta)| < \infty$ and*

$$\mathbb{E}\, U(f,\eta) = \int f(x_1, \ldots, x_k) \lambda^k(\mathrm{d}(x_1, \ldots, x_k)).$$

If, moreover, $U(f,\eta) \in L^2(P_\eta)$, then

$$\operatorname{Var}[U(f,\eta)] = \sum_{n=1}^{k} n! \|f_n\|_n^2,$$

where f_n are given in (2.15).

To study the asymptotic behaviour of U-statistics, we denote by η_t the Poisson point process on $\mathbb{X}$ with intensity measure $\lambda_t = t\lambda$ and put $F_t = U(f, \eta_t)$ where $f \in L^1(\lambda)$ is some fixed symmetric function such that F_1 has a finite variance. Then $F_t - \mathbb{E}\, F_t$ admits the chaotic representation

$$F_t - \mathbb{E}\, F_t = \sum_{n=1}^{k} I_n(g_{n,t}),$$

where the stochastic integrations are with respect to λ_t^n and

$$g_{n,t}(x_1, \ldots, x_n) = t^{k-n} \binom{k}{n} \int f(x_1, \ldots x_n, y_1, \ldots, y_{k-n}) \lambda^{k-n}(\mathrm{d}(y_1, \ldots, y_{k-n})).$$

Define a constant n_1 by

$$n_1 := \inf\{n; \|g_{n,t}\|_n \neq 0\}.$$

In fact, n_1 is the so-called *Hoeffding rank* of the U-statistic $U(f, \eta_t)$ (see e.g. Vitale [1992] for the theory of Hoeffding decomposition). The following asymptotic result can be proved using the Malliavin–Stein's method.

Theorem 2.17 (Theorem 7.3 in Lachièze-Rey and Peccati [2013]). *Let* $\hat{F}_t := (\mathrm{Var}\,[F_t])^{-1/2}(F_t - \mathbb{E}\,F_t)$. *Then there exist constants* c_1, c_2, c_3 *not depending on* t *such that*

$$c_1 t^{2k-n_1} \leqslant \mathrm{Var}\,(F_t) \leqslant c_2 t^{2k-n_1}.$$

If $n_1 = 1$, *then* $U(f, \eta_t)$ *follows a central limit theorem and*

$$d_W(\hat{F}_t, Z) \leqslant c_3 t^{-1/2},$$
$$d_K(\hat{F}_t, Z) \leqslant c_3 t^{-1/2}.$$

Remark. If $n_1 > 1$, then $U(f, \eta_t)$ does not follow a central limit theorem. See Section 2.1 in Lachièze-Rey and Reitzner [2016] for the discussion over the speed of the convergence depending on the choice of the kernel function f.

Next, we present one example of U-statistic counting the intrinsic volumes of the intersections in the flat process. For more examples, see Chapter 4 in Lachièze-Rey and Reitzner [2016].

Example 2.5 (Intersection process).
Denote by $A(d, i)$ the affine Grassmanian (the space of all i-dimensional spaces in $\mathbb{R}^d$, endowed with the usual hit-and-miss topology and Borel σ-field).

Let η_t be a Poisson point process on $A(d, i)$ with intensity measure $\lambda_t = t\lambda$ for some locally finite and non-atomic measure λ on $A(d, i)$. Then we call η_t the *Poisson flat process.*

Take a compact, convex observation window $W \subset \mathbb{R}^d$ with interior points. Suppose we observe only what is happening inside W, i.e. we understand η_t as a point process on $[W] := \{L \in A(d, i); L \cap W \neq \emptyset\}$.

Denote by $n_t^{(k)}$ the process of the intersections of k flats of η_t for $k \leqslant d/(d-i)$ and define the U-statistic Φ_t by

$$\Phi_t = \Phi_t(W, i, k, j) = \frac{1}{k!} \sum_{(L_1,\dots,L_k) \in \eta^k_{t,\neq}} V_j(L_1 \cap \cdots \cap L_k \cap W)$$

for $j = 0, \dots, d - k(d-i), i = 0, \dots d-1$ and $k = 1, \dots, m$, where m is the greatest integer with $m \leqslant d/(d-i)$.

Theorem 2.18 (Theorem 10 in Lachièze-Rey and Reitzner [2016]). *There is a constant* $c = c(W, i, k, j)$ *such that for* $t \geqslant 1$,

$$d_W(\hat{\Phi}_t, Z) \leqslant ct^{-1/2} \quad \textit{and} \quad d_K(\hat{\Phi}_t, Z) \leqslant ct^{-1/2}.$$

Example 3: Normal approximation for point processes with Papangelou conditional intensity

The Malliavin–Stein's method can work successfully also for functionals of point processes having a Papangelou conditional intensity (Definition 1.17). In Torrisi [2017], bounds on the Wasserstein distance between the standard normal distribution and distribution of so-called *innovations* are proved using the techniques from Malliavin–Stein's method for point processes having Papangelou conditional intensity. Consequently, these bounds are derived for Gibbs point processes in $\mathbb{R}^d, d \in \mathbb{N}$.

Definition 2.13 (Innovation of a point process).
Let μ be a point process on $\mathbb{X}$ with Papangelou conditional intensity λ^* and intensity measure λ. Then we define the *innovation* of the point process μ as a random variable

$$I_\mu(\varphi) := \sum_{x \in \mu} \varphi(x, \mu - \delta_x) - \int_{\mathbb{X}} \varphi(x, \mu)\lambda^*(x, \mu)\lambda(\mathrm{d}x)$$

for any measurable $\varphi : \mathbb{X} \times \mathbf{N} \to \mathbb{R}$, for which $|I_{\mathbf{x}}(\varphi)| < \infty$ for μ-a.a. $\mathbf{x} \in \mathbf{N}$.

Remark. It follows from the Georgii–Nguyen–Zessin formula (Theorem 1.12) that $\mathbb{E}[I_\mu(\varphi)] = 0$ for any innovation defined above.

Theorem 2.19 (Theorem 3.1 in Torrisi [2017]).
Let $\varphi : \mathbb{X} \times \boldsymbol{N} \to \mathbb{R}$ be a measurable function satisfying

$$\mathbb{E}\left[\int_{\mathbb{X}} |\varphi(x, \mu)|\lambda^*(x, \mu)\lambda(\mathrm{d}x)\right] < \infty \text{ and } \mathbb{E}\left[\int_{\mathbb{X}} |\varphi(x, \mu)|^2\lambda^*(x, \mu)\lambda(\mathrm{d}x)\right] < \infty.$$

Then,

$$d_W(I_\mu(\varphi), Z) \leqslant \sqrt{\frac{2}{\pi}}\mathbb{E}\left[\left|1 - \int_{\mathbb{X}} \varphi(x, \mu)D_x I_\mu(\varphi)\lambda^*(x, \mu)\lambda(\mathrm{d}x)\right|\right]$$
$$+\mathbb{E}\left[\int_{\mathbb{X}} |\varphi(x, \mu)||D_x I_\mu(\varphi)|^2\lambda^*(x, \mu)\lambda(\mathrm{d}x)\right].$$

Remark. An advantage of Theorem 2.19 is that it allows the function φ to depend also on a given realization of the point process μ. That gives us an opportunity to study important functionals as the volume of intersections between particles in this realisation, etc. However, the terms are usually difficult to evaluate.

The following result simplifies considerably the bound in Theorem 2.19, but with the price that the function φ no longer depends on a given realization, hence it is only function on $\mathbb{X}$.

Theorem 2.20 (Corollary 3.5 in Torrisi [2017]).
Let $\varphi : \mathbb{X} \to \mathbb{R}$ be a measurable function such that

$$\int_{\mathbb{X}} |\varphi(x)|\mathbb{E}\left[\lambda^*(x, \mu)\right]\lambda(\mathrm{d}x) < \infty \text{ and } \int_{\mathbb{X}} |\varphi(x)|^2\mathbb{E}\left[\lambda^*(x, \mu)\right]\lambda(\mathrm{d}x) < \infty.$$

Then,

$$d_W(I_\mu(\varphi), Z) \leqslant$$

$$\sqrt{\frac{2}{\pi}}\sqrt{1 - 2\int_{\mathbb{X}} |\varphi(x)|^2 \mathbb{E}\left[\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x) + \int_{\mathbb{X}^2} |\varphi(x)\varphi(y)|^2 \alpha_2(x,y,\mu)\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)}$$

$$+ \int_{\mathbb{X}} |\varphi(x)|^3 \mathbb{E}\left[\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x)$$

$$+ \sqrt{\frac{2}{\pi}}\int_{\mathbb{X}^2} |\varphi(x)\varphi(y)|\mathbb{E}\left[|D_x\lambda^*(y,\mu)|\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)$$

$$+ 2\int_{\mathbb{X}^2} |\varphi(x)|^2\varphi(y)|\mathbb{E}\left[|D_x\lambda^*(y,\mu)|\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)$$

$$+ \int_{\mathbb{X}^3} |\varphi(x)\varphi(y)\varphi(z)|\mathbb{E}\left[|D_x\lambda^*(y,\mu)D_x\lambda^*(z,\mu)|\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)\lambda(\mathrm{d}z).$$

Moreover, if we add the assumption of repulsivity of the point process μ, we can express the bound of Theorem 2.20 using the product densities (Definition 1.9) up to the third order.

Definition 2.14 (Repulsive point process).
The point process μ on $\mathbb{X}$ with the Papangelou conditional intensity λ^* is said to be *repulsive* if

$$\lambda^*(x, \mathbf{x}) \geqslant \lambda^*(x, \mathbf{y}) \quad \text{for } \mathbf{x}, \mathbf{y} \in \mathbf{N}(\mathbb{X}), \mathbf{x} \subset \mathbf{y}, x \in \mathbb{X}.$$

Notation. In what follows we use the following notation. Define functions $\alpha_2 : \mathbb{X}^2 \times \mathbf{N} \to \mathbb{R}$ and $\alpha_3 : \mathbb{X}^3 \times \mathbf{N} \to \mathbb{R}$ by

$$\alpha_2(x, y, \mu) := \mathbb{E}[\lambda^*(x,\mu)\lambda^*(y,\mu)],$$
$$\alpha_3(x, y, z, \mu) := \mathbb{E}[\lambda^*(x,\mu)\lambda^*(y,\mu)\lambda^*(z,\mu)],$$

for $x, y, z \in \mathbb{X}$ and the point process μ on $\mathbb{X}$.

Corollary 2.4 (Corollary 3.7 in Torrisi [2017]).
Let $\varphi : \mathbb{X} \to \mathbb{R}$ be a measurable function such that

$$\int_{\mathbb{X}} |\varphi(x)|\mathbb{E}\left[\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x) < \infty \textit{ and } \int_{\mathbb{X}} |\varphi(x)|^2\mathbb{E}\left[\lambda^*(x,\mu)\right]\lambda(\mathrm{d}x) < \infty.$$

If, moreover, μ is repulsive, we have

$$d_W(I_\mu(\varphi), Z) \leqslant$$

$$\sqrt{\frac{2}{\pi}}\sqrt{1 - 2\int_{\mathbb{X}} |\varphi(x)|^2 \lambda^{[1]}(x)\lambda(\mathrm{d}x) + \int_{\mathbb{X}^2} |\varphi(x)\varphi(y)|^2 \alpha_2(x, y, \mu)\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)}$$

$$+ \int_{\mathbb{X}} |\varphi(x)|^3 \lambda^{[1]}(x)\lambda(\mathrm{d}x) + \sqrt{\frac{2}{\pi}} \int_{\mathbb{X}^2} |\varphi(x)\varphi(y)|(\alpha_2(x, y, \mu) - \lambda^{[2]}(x, y))\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)$$

$$+ 2\int_{\mathbb{X}^2} |\varphi(x)|^2 \varphi(y)|(\alpha_2(x, y, \mu) - \lambda^{[2]}(x, y))\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)$$

$$+ \int_{\mathbb{X}^3} |\varphi(x)\varphi(y)\varphi(z)|(\alpha_3(x, y, z, \mu) - \lambda^{[3]}(x, y, z))\lambda(\mathrm{d}x)\lambda(\mathrm{d}y)\lambda(\mathrm{d}z).$$

Corollary 2.4 may be useful to provide explicit bounds in the normal approximation of innovations of repulsive point processes for which the first three correlation functions are explicitly known. This is the imminent case of determinantal point processes (recall Example 1.10 or see Section 7 in Torrisi [2017]).

Example 4: Normal approximation and CLT in the Gibbs setting

The general bound of Theorem 2.20 can be used in the normal approximation of the innovation of a Gibbs point process on $\mathbb{R}^d$ with pair potential where the exact form of the conditional intensity is known.

Recall that a *pair potential* is a Borel measurable function $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ such that $\phi(x) = \phi(-x)$. For $x \in \mathbf{N}$ and $u \in \mathbb{R}^d$, we define the *relative energy* of interaction between the point u and the configuration $\mathbf{x}$ by

$$E(u, \mathbf{x}) = \begin{cases} \sum_{y \in \mathbf{x}} \phi(u - y), & \text{if } \sum_{y \in \mathbf{x}} |\phi(u - y)| < \infty, \\ +\infty, & \text{otherwise.} \end{cases}$$

A point process μ on $\mathbb{R}^d$ is called the Gibbs point process with activity $\tau > 0$ and pair potential ϕ if its Papangelou conditional intensity takes form

$$\lambda^*(u, \mathbf{x}) = \tau \exp\{-E(u, \mathbf{x})\}, \quad u \in \mathbb{R}^d, \ \mathbf{x} \in \mathbf{N}.$$

Moreover, it will be assumed that μ is stationary, *inhibitory*, i.e. $\phi \geqslant 0$ and *finite range* meaning that $1 - e^{-\phi}$ has compact support.

Theorem 2.21 (Theorem 5.3 in Torrisi [2017]).
Let μ be a stationary Gibbs point process with activity $\tau > 0$ and pair potential $\phi : \mathbb{R}^d \to [0, +\infty]$, and suppose

$$\varphi \in L^1(\mathbb{R}^d, |\cdot|_d) \cap L^2(\mathbb{R}^d, |\cdot|_d).$$

If, moreover, μ has finite range, then for any $p,\ q,\ p',\ q' > 1$ such that $\frac{1}{p} + \frac{1}{q} = \frac{1}{p'} + \frac{1}{q'} = 1$,

$$d_W(I_\mu(\varphi), Z) \leqslant \sqrt{\frac{2}{\pi}}\sqrt{1 - 2c_1\|\varphi\|^2_{L^2(\mathbb{R}^d,|\cdot|_d)} + \tau c_2\|\varphi\|^4_{L^2(\mathbb{R}^d,|\cdot|_d)}} + c_2 A,$$

where

$$\begin{aligned} A :=& \|\varphi\|^3_{L^3(\mathbb{R}^d,|\cdot|_d)} + \sqrt{\frac{2}{\pi}}\tau\|\varphi\|^2_{L^2(\mathbb{R}^d,|\cdot|_d)}\|1 - e^{-\phi}\|_{L^1(\mathbb{R}^d,|\cdot|_d)} \\ &+ 2\tau\|\varphi\|^2_{L^q(\mathbb{R}^d,|\cdot|_d)}\|1 - e^{-\phi}\|_{L^1(\mathbb{R}^d,|\cdot|_d)} \\ &+ \tau^2\|\varphi\|_{L^{pp'}(\mathbb{R}^d,|\cdot|_d)}\|\varphi\|_{L^{pq'}(\mathbb{R}^d,|\cdot|_d)}\|\varphi\|_{L^q(\mathbb{R}^d,|\cdot|_d)}\|1 - e^{-\phi}\|^2_{L^1(\mathbb{R}^d,|\cdot|_d)} \end{aligned}$$

and

$$c_1 := \frac{\tau}{1 + \tau\|1 - e^{-\phi}\|_{L^1(\mathbb{R}^d,|\cdot|_d)}}, \qquad c_2 := \frac{\tau}{2 - \exp\{-\tau\|1 - e^{-\phi}\|_{L^1(\mathbb{R}^d,|\cdot|_d)}\}}.$$

Example 2.6 (Hard-core process).
Take $r > 0$ fixed and let $\phi : \mathbb{R}^d \to \mathbb{R} \cup \{\infty\}$ be a real function defined by

$$\phi(x) = \begin{cases} 0, & \text{if } \|x\| > r, \\ +\infty, & \text{if } \|x\| \leqslant r. \end{cases}$$

Set the relative energy between point $u \in \mathbb{R}^d$ and system of points $\mathbf{x} \in \mathbf{N}$ as

$$\begin{aligned} E(u, \mathbf{x}) =& \begin{cases} \sum_{y\in\mathbf{x}} \phi(u - y), & \text{if } \sum_{y\in\mathbf{x}} |\phi(u - y)| < \infty, \\ +\infty, & \text{otherwise} \end{cases} \\ =& \begin{cases} 0, & \text{if } \|u - y\| > r,\ \forall y \in \mathbf{x}, \\ +, \infty & \text{otherwise.} \end{cases} . \end{aligned}$$

Then we define the hard-core point process with pair potential by its Papangelou conditional intensity

$$\lambda^*(u, \mathbf{x}) = \tau \exp\{-E(u, \mathbf{x})\} = \tau \mathbf{1}_{[\|u-y\|>r,\ \forall y\in\mathbf{x}]}, \quad u \in \mathbb{R}^d, \mathbf{x} \in \mathbf{N}.$$

Theorem 2.22 (Theorem 4.8 in Flimmel [2017]).
Consider for each $n \in \mathbb{N}$ a stationary hard-core point process $\mu^{(n)}$ in $\mathbb{R}^d$ with activity $\tau_n > 0$ such that $\tau_n \to \tau$ as $n \to \infty$, and with pair potential

$$\phi_n(y) = \begin{cases} 0, & \text{if } \|x\| > r_n, \\ +\infty, & \text{if } \|x\| \leqslant r_n, \end{cases}$$

where $r_n \geqslant 0,\ r_n \to 0$ as $n \to \infty$. Let $K_n,\ n \in \mathbb{N}$, be bounded Borel sets in $\mathbb{R}^d$ such that $|K_n|_d \to \infty$ as $n \to \infty$. Define functions

$$\varphi_n(x) = \frac{1}{\sqrt{\tau_n|K_n|_d}} \cdot \mathbf{1}_{K_n}(x), \quad n \in \mathbb{N}, x \in \mathbb{R}^d.$$

Then,

$$d_W(I_{\mu^{(n)}}(\varphi_n), Z) \to 0$$

as $n \to \infty$.

Proof. It is clear that the point processes $\mu^{(n)}$ have finite ranges. Also, for every $n \in \mathbb{N}$

$$\int\limits_{\mathbb{R}^d} |\varphi_n(x)| \mathrm{d}x = \int\limits_{\mathbb{R}^d} \left| \frac{1}{\sqrt{\tau_n |K_n|_d}} \cdot \mathbf{1}_{K_n}(x) \right| \mathrm{d}x = \sqrt{\frac{|K_n|_d}{\tau_n}} < \infty$$

and

$$\int\limits_{\mathbb{R}^d} |\varphi_n(x)|^2 \mathrm{d}x = \int\limits_{\mathbb{R}^d} \left| \frac{1}{\sqrt{\tau_n |K_n|_d}} \cdot \mathbf{1}_{K_n}(x) \right|^2 \mathrm{d}x = \frac{1}{\tau_n} < \infty.$$

Hence, the assumptions of Theorem 2.21 are satisfied and so we can compute bounds on the Wasserstein distance between the standard normal distribution Z and the innovation $I_{\mu^{(n)}}(\varphi_n)$ for each $n \in \mathbb{N}$.

First, we need to compute the L^1 norm of the function $1 - e^{-\phi_n}$:

$$\|1 - e^{-\phi_n}\|_{L^1(\mathbb{R}^d, |\cdot|_d)} = \int\limits_{\mathbb{R}^d} |1 - e^{-\phi_n(x)}| \mathrm{d}x = \int\limits_{B_{r_n}(\mathbf{o})} 1 \mathrm{d}x = |B_{r_n}(\mathbf{o})|_d.$$

Set $p = q = p' = q' = 2$ and compute for given $n \in \mathbb{N}$ the constants $A^{(n)}, c_1^{(n)}$ and $c_2^{(n)}$ from Theorem 2.21:

$$c_1^{(n)} = \frac{\tau_n}{1 + \tau_n \|1 - e^{-\phi_n}\|_{L^1(\mathbb{R}^d, |\cdot|_d)}} = \frac{\tau_n}{1 + \tau_n |B_{r_n}(\mathbf{o})|_d},$$

$$c_2^{(n)} = \frac{\tau_n}{2 - \exp\{-\tau_n \|1 - e^{-\phi_n}\|_{L^1(\mathbb{R}^d, |\cdot|_d)}\}} = \frac{\tau_n}{2 - \exp\{-\tau_n |B_{r_n}(\mathbf{o})|_d\}}$$

and

$$\begin{aligned} A^{(n)} &= \|\varphi_n\|^3_{L^3(\mathbb{R}^d, |\cdot|_d)} + \sqrt{\frac{2}{\pi}} \tau_n \|\varphi_n\|^2_{L^2(\mathbb{R}^d, |\cdot|_d)} \|1 - e^{-\phi_n}\|_{L^1(\mathbb{R}^d, |\cdot|_d)} \\ &\quad + 2\tau_n \|\varphi_n\|^2_{L^2(\mathbb{R}^d, |\cdot|_d)} \|1 - e^{-\phi_n}\|_{L^1(\mathbb{R}^d, |\cdot|_d)} \\ &\quad + \tau_n^2 \|\varphi_n\|^2_{L^4(\mathbb{R}^d, |\cdot|_d)} \|\varphi_n\|_{L^2(\mathbb{R}^d, |\cdot|_d)} \|1 - e^{-\phi_n}\|^2_{L^1(\mathbb{R}^d, |\cdot|_d)} \\ &= \frac{1}{\tau_n^{3/2} \sqrt{|K_n|}} + \sqrt{\frac{2}{\pi}} |B_{r_n}(\mathbf{o})| + 2|B_{r_n}(\mathbf{o})|_d \\ &\quad + \frac{\sqrt{\tau_n}}{\sqrt{|K_n|_d}} (|B_{r_n}(\mathbf{o})|_d)^2. \end{aligned}$$

We can see that

$$c_1^{(n)} \to \tau, \qquad c_2^{(n)} \to \tau, \qquad A^{(n)} \to 0$$

as $n \to \infty$. Altogether, using bound from Theorem 2.21, we arrive at

$$\begin{aligned} d_W(I_{\mu^{(n)}}(\varphi_n), Z) &\leqslant \sqrt{\frac{2}{\pi}} \sqrt{1 - 2c_1^{(n)} \|\varphi_n\|^2_{L^2(\mathbb{R}^d, |\cdot|_d)} + \tau_n c_2^{(n)} \|\varphi_n\|^4_{L^2(\mathbb{R}^d, |\cdot|_d)}} + c_2^{(n)} A^{(n)} \\ &= \sqrt{\frac{2}{\pi}} \sqrt{1 - 2c_1^{(n)} \frac{1}{\tau_n} + c_2^{(n)} \frac{1}{\tau_n}} + c_2^{(n)} A^{(n)}, \end{aligned}$$

which tends to 0 as n approaches $+\infty$.

□

Example 2.7 (Strauss process).
The Strauss process is a Gibbs point process μ with activity τ and range of interaction $r > 0$ with the pair potential

$$\phi(x) := (-\log u)\mathbf{1}_{[\|x\| \leqslant r]}, \quad u \in [0,1], x \in \mathbb{R}^d.$$

Then ϕ satisfies the assumptions of Theorem 2.21 and

$$\|1 - e^{-\phi}\|_{L^1(\mathbb{R}^d, |\cdot|_d)} = (1-u)\kappa_d r^d,$$

where κ_d is the volume of the unit ball in $\mathbb{R}^d$.

The result of Theorem 5.3 in Torrisi [2017] was later extended in Flimmel and Beneš [2018] to the case $\mathbb{X} = \mathcal{C}^{(d)}, d \in \mathbb{N}$ (recall the notation of Section 1.4).

Assume μ is a stationary Gibbs particle process with the energy of the form (1.19). The explicit form of the Papangelou conditional intensity is given by (1.20).

The innovation of a Gibbs particle process μ is of the form

$$I_\mu(\varphi) = \sum_{K \in \mu} \varphi(K, \mu \backslash \{K\}) - \int_{\mathcal{C}^{(d)}} \varphi(K,\mu)\lambda^*(K,\mu)\lambda(\mathrm{d}K)$$

for any measurable $\varphi : \mathcal{C}^{(d)} \times \mathbf{N}^d \to \mathbb{R}$, for which $I_{\mathbf{x}}(\varphi)$ is defined and finite μ-a.e. on $\mathbf{N}^d$.

Theorem 2.23 (Theorem 3.3 in Flimmel and Beneš [2018]).
Let μ be a stationary Gibbs particle process given by the conditional intensity of the form (1.20) with activity $\tau > 0$, inverse temperature $\beta \geqslant 0$, reference particle distribution $\mathbb{Q}$ satisfying (1.16), and with pair potential g which is bounded from above by some positive constant a. Let $\varphi : \mathcal{C}^{(d)} \to \mathbb{R}$ be a measurable function that does not depend on $\boldsymbol{x} \in \boldsymbol{N}^d$ and

$$\varphi \in L^1(\mathcal{C}^{(d)}, \lambda) \cap L^2(\mathcal{C}^{(d)}, \lambda).$$

Then

$$\begin{aligned} d_W(I_\mu(\varphi), Z) \leqslant & \sqrt{\frac{2}{\pi}}\sqrt{1 - 2\tau(1-\beta b)\|\varphi\|^2_{L^2(\mathcal{C}^{(d)},\lambda)} + \tau^2\|\varphi\|^4_{L^2(\mathcal{C}^{(d)},\lambda)}} \\ & + \tau\|\varphi\|^3_{L^3(\mathcal{C}^{(d)},\lambda)} + \sqrt{\frac{2}{\pi}}\tau^2\|\varphi\|^2_{L^1(\mathcal{C}^{(d)},\lambda)}|1 - e^{-\beta a}| \\ & + 2\tau^2\|\varphi\|^2_{L^2(\mathcal{C}^{(d)},\lambda)}\|\varphi\|_{L^1(\mathcal{C}^{(d)},\lambda)}|1 - e^{-\beta a}| \\ & + \tau^3\|\varphi\|^3_{L^1(\mathcal{C}^{(d)},\lambda)}|1 - e^{-\beta a}|^2. \end{aligned}$$

Example 2.8 (Gibbs planar segment process).
Take $\mathbb{Q}$ being concentrated on the set $S^R_{\mathbf{o}} \subset \mathcal{C}^{(2)}$ (the space of all segments in $\mathbb{R}^2 \cap B_R(\mathbf{o})$ centered in the origin). Theorem 2.23 can be applied to the special case of planar segment process to derive central limit theorems for two functionals: the normalized number of segments observed in a window and normalized total length of segments hitting the window. We take windows forming a convex averaging sequence (cf. Daley and Vere-Jones [2003]), i.e. monotone increasing sequence of convex bounded Borel sets converging to $\mathbb{R}^2$.

Theorem 2.24 (Theorem 3.4 in Flimmel and Beneš [2018]).
Consider for each $n \in \mathbb{N}$ a stationary Gibbs planar segment process $\xi^{(n)}$ with the conditional intensity

$$\lambda_n^*(K, \mathbf{x}) = \tau_n \exp\left\{-\beta_n \sum_{L \in \mathbf{x}} \boldsymbol{1}\{K \cap L \neq \emptyset\}\right\}, \quad K \in S, \mathbf{x} \in \boldsymbol{N}^2,$$

where $\tau_n > 0$ and $\beta_n \geqslant 0$. Moreover, suppose that $\beta_n \to 0$ and $0 < c_1 < \tau_n < c_2 < \infty$, $n \in \mathbb{N}$, for some constants c_1, c_2 and that the common reference particle distribution $\mathbb{Q}$ for all $\xi^{(n)}$ has the uniform directional distribution. Let $\{W_n,\ n \in \mathbb{N}\}$ be a convex averaging sequence in $\mathbb{R}^2$ such that $|W_n|_d = o(\beta_n^{-1})$ (i.e. $|W_n|_d \beta_n$ tends to zero with growing n). For $n \in \mathbb{N}$ and $K \in S$, define

$$\varphi_n(K) = \frac{1}{\sqrt{\tau_n |W_n|_d}} \cdot \boldsymbol{1}\{K \cap W_n \neq \emptyset\}.$$

Further

$$\psi_n(K) = \frac{l(K)}{\sqrt{\mathbb{E}_L l^2}} \varphi_n(K),$$

where $l(K)$ denotes the length of the segment K, l is a random variable that follows the law of $\mathbb{Q}_L$ and $\mathbb{E}_L$ denotes the expectation with respect to $\mathbb{Q}_L$. Then

$$d_W(I_{\xi^{(n)}}(\varphi_n), Z) \to 0, \qquad d_W(I_{\xi^{(n)}}(\psi_n), Z) \to 0$$

as $n \to \infty$.

3. Method of stabilization

In this chapter, we present a recent method for studying the limit behaviour of geometric structures evincing local form of dependency. Typical examples of application include random graphs (Section 1.6), germ-grain models (Example 1.14) and other geometric structures based on marked point processes such as weighted Voronoi tessellations (Example 1.22) or Delaunay triangulation (Example 1.19), where the unmarked point process displays only local dependencies.

The asymptotic properties are usually investigated via geometric functionals called *scores* or *score functions* of the type

$$\xi(x, \mathbf{x}),$$

where ξ is a real-valued function defined on pairs $(x, \mathbf{x})$ of $x \in \mathbf{x}$ and a locally finite point configuration $\mathbf{x}$ in $\mathbb{R}^d$ (or some more general space).

We shall write $H(\mathbf{x})(:= H^\xi(\mathbf{x}))$ for the total sum over all $x \in \mathbf{x}$, i.e.

$$H(\mathbf{x}) := \sum_{x \in \mathbf{x}} \xi(x, \mathbf{x}). \tag{3.1}$$

The statistic $H(\mathbf{x})$ typically describes a global property of a geometric structure generated by $\mathbf{x}$ and the value $\xi(x, \mathbf{x})$ represents the interaction of x with respect to $\mathbf{x}$.

For a general point process μ, $H(\mu)$ is a sum of mutually dependent terms. A lot of functionals in stochastic geometry are in the form (3.1), e.g. the total edge length of a random graph, statistics of Voronoi set approximation, etc. The object of our interest is then the asymptotic behaviour of $H(\mu \cap W_\lambda)$, where W_λ is a suitable observation window tending to $\mathbb{R}^d$. Alternatively, we let the observation window be fixed and let the intensity of a stationary point process μ tend to infinity. Often, the values of $\xi(x, \mathbf{x})$ and $\xi(y, \mathbf{x})$, $x \neq y$, are not unrelated but, loosely speaking, become more related as the distance between x and y becomes smaller. This dependency cause problems when developing the limit theory for H on random point sets.

The *stabilization method* is a tool that allows one to study statistics of this type, which may be expressed as a sum of spatially dependent terms, where the short-range interactions can be controlled. Roughly speaking, a geometric functional stabilizes if its behaviour at a given point is locally determined by a certain finite, possibly random, neighbourhood of this point. Or, in another words, any local modification (e.g. insertion of a point into the underlying point process) has only a local effect.

The notion of stabilization comes from Lee [1997, 1999] and Avram and Bertsimas [1993]. The modern theory of stabilization in the context of central limit theorems in stochastic geometry was introduced in Penrose and Yukich [2001, 2002] and Baryshnikov and Yukich [2005]. The laws of large numbers for stabilizing functionals of point processes on $\mathbb{R}^d$ are investigated in Penrose and Yukich [2003]. For a survey on limit theorems in stochastic geometry with a particular focus on stabilization we refer to Schreiber [2010].

Notation. For a score function ξ we will write $\xi(x, \mathbf{x})$ for $\xi(x, \mathbf{x} \cup \{x\})$ whenever $\mathbf{x} \in \mathbf{N}$ and $x \in \mathbb{R}^d$.

The most commonly understood definition of a stabilizing functional goes as follows:

Definition 3.1 (Stabilizing score function, radius of stabilization). A score function ξ is said to be *stabilizing* with respect to a point process μ if and only if for each $x \in \mathbb{R}^d$ there exists an a.s. finite random variable $R_x := R_x^{\xi}(\mu)$ such that

$$\xi(x, \mu) = \xi(x, (\mu \cap B_{R_x}(x)) \cup \mathcal{A}), \quad \text{a.s.}$$

for all finite $\mathcal{A} \subset \mathbb{R}^d \backslash B_{R_x}(x)$, where as usual $B_r(x)$ is a ball in $\mathbb{R}^d$ with radius r around x. The random variable R_x is called the *radius of stabilization*.

Definition 3.1 says that the value of $\xi(x, \mu)$ is almost surely fully determined by the configuration of μ inside $B_{R_x}(x)$. Naturally, the concept of stabilization plays the most significant role in the study of Poisson functionals, since the configuration of a Poisson point process η inside $B_R(x)$ does not depend on the configuration outside $B_R(x)$ for any $x \in \eta$ and $R > 0$. Nevertheless, also the cases of binomial and Gibbsian inputs are considered in the literature (see Penrose and Yukich [2001] for the binomial case and Schreiber and Yukich [2013] for the Gibbsian case).

The strength of the stabilization is characterized by the tail distribution of the radii of stabilization, i.e. the probabilities $\mathbb{P}(R_x > r), r > 0$.

Definition 3.2 (Exponentially stabilizing score functions). We say that ξ is *exponentially stabilizing* if there exist constants $c_{stab}, \alpha_{stab} \in (0, \infty)$ such that for all $x \in \mathbb{R}^d$ and $r > 0$,

$$\mathbb{P}(R_x > r) \leqslant c_{stab} \exp\left(-\frac{1}{c_{stab}} r^{\alpha_{stab}}\right).$$

Yet, there are several different variations available for proving central limit theorems based on stabilization, all having slightly different notions of the stabilization property. We shall mention at least three approaches and briefly compare their assumptions and results. Namely, it is

1. *add-one cost stabilization*: an approach based on martingale differences presented in Penrose and Yukich [2001],

2. *moment approach*: a combination of cumulant method (see Chapter 4) and stabilization introduced in Baryshnikov and Yukich [2005],

3. *Stein's method*: a combination of the Malliavin–Stein bounds (see Chapter 2) and stabilization presented in Last et al. [2016].

Even though every approach generates a large scale of applications, especially in the random graphs theory, we will demonstrate the usage on the nearest neighbour graph as it is the most frequently discussed example in the literature. The rest of the chapter is organized as follows: Sections 3.1-3.3 each discusses an individual approach. Section 3.4 is devoted to the stabilization property of random tessellations, where the results of the previous sections are used.

3.1 Add-one cost stabilization

The add-one cost version of stabilization is based on the martingale method developed in Kesten and Lee [1996] and Lee [1997], where the authors studied random Euclidean minimal spanning trees. The main idea of the method is that the value of $\xi(x, \mathbf{x})$ is unaffected by the external configuration of $\mathbf{x}$ (i.e. by points beyond some stabilization distance from x) and moreover, the value of ξ for the points outside the stabilization region of x is not affected by the presence of x. Therefore, this version of stabilization is sometimes referred to as *external stabilization*.

The presentation below is mostly based on Penrose and Yukich [2001], where the authors derived a general central limit theorem for functionals of random graphs as defined in Section 1.6. The functionals of interest include the total edge length, the total number of any type of component, etc. The limit behaviour of these functionals is investigated in two regimes:

1. increasing non-random number of random points being distributed uniformly in a fixed region or
2. increasing observation window in $\mathbb{R}^d$.

Definition 3.3 (Add-one cost).
The *add-one cost* $D_x(\mathbf{x})$ of a point configuration $\mathbf{x} \not\ni x$ with respect to geometric functional ξ is given as the increment caused by inserting a point $x \in \mathbb{R}^d$ into $\mathbf{x}$, i.e.

$$D_x H(\mathbf{x}) := H(\mathbf{x} \cup \{x\}) - H(\mathbf{x}).$$

Note that the add-one cost is exactly the difference operator of the first order (see Definition 2.7) and itself is a geometric statistic. It will be assumed that H is translation-invariant, so that

$$H(\mathbf{x} + y) = H(\mathbf{x})$$

for all $\mathbf{x} \subset \mathbb{R}^d$ and $y \in \mathbb{R}^d$. Then it is enough to investigate the case when a point is inserted in the origin. Thus, we denote

$$\Delta(\mathbf{x}) := D_{\mathbf{o}}(H)(\mathbf{x}) = H(\mathbf{x} \cup \{\mathbf{o}\}) - H(\mathbf{x})$$

for a point set $\mathbf{x} \in \mathbb{R}^d$.

Remark. The add-one cost method does not require the functional H to be in the form (3.1). Nevertheless, most of the examples in Penrose and Yukich [2001] have such form and the add-one cost then equals

$$D_x H(\mathbf{x}) = \sum_{y \in \mathbf{x} \cup \{x\}} \xi(y, \mathbf{x} \cup \{x\}) - \sum_{y \in \mathbf{x}} \xi(y, \mathbf{x}).$$

Assumptions

Notation. In what follows, we denote for $r > 0$ and $x \in \mathbb{R}^d$ the d-cube centered in x by $Q_r(x) := [-r, r]^d + x$. Moreover, for $F \subset \mathbb{R}^d$, let ∂F denote the boundary of F, i.e. the intersection of the closure of F with the closure of its complement and set $\partial_r F := \cup_{x \in \partial F} Q_r(x)$.

Let us fix $d \geqslant 1$ and $\lambda > 0$. We suppose that $(W_n)_{n \geqslant 1}$ is a sequence of bounded observation windows in $\mathbb{R}^d$ satisfying the following conditions:

- $|W_n|_d = n/\lambda$ for all $n \in \mathbb{N}$,
- $W_n \nearrow \mathbb{R}^d$,
- the *vanishing relative boundary* condition:

$$\lim_{n\to\infty} \frac{|\partial_r W_n|_d}{n} = 0, \quad \text{for all } r > 0,$$

- the *polynomial boundedness* condition: There exists a constant β_1 such that the diameter of W_n satisfies $\operatorname{diam}(W_n) \leqslant \beta_1 n^{\beta_1}$ for all $n \in \mathbb{N}$.

Denote by $\mathcal{W}$ the collection of all regions $A \subset \mathbb{R}^d$ having the form $A = \{W_n + x : x \in \mathbb{R}^d, n \in \mathbb{N}\}$. Moreover, let W_0 be a fixed bounded Borel set in $\mathbb{R}^d$ satisfying $|W_0|_d = 1$ and $|\partial W_0|_d = 0$ and denote by $\mathcal{W}_0$ the collection of all regions of the form $A = aW_0 + x$ with $a \geqslant 1, x \in \mathbb{R}^d$.

Definition 3.4 (Regularity condition).
We say that a collection of sets $\mathcal{B}$ is *regular* if there exists $\delta > 0$ such that for all $r \in [1, \infty)$, whenever $B \in \mathcal{B}$ and $x, y \in W$ with $\|x - y\| = r$, we have

$$|B_{r/4}(x) \cap B| \geqslant \delta r^{\delta}.$$

For example, the collection of all boxes $B \subset \mathbb{R}^d$ of the form $\prod_{i=1}^{d}[a_i, b_i]$ with $b_i \geqslant a_i + 1$ for each i is regular. Similarly, the collection of all balls or ellipsoids is regular. As for $\mathcal{W}_0$, the sufficient condition for being regular is that $r^{-d}|B_r(x) \cap W_0|_d$ is bounded away from zero, uniformly over $x \in W_0$ and $r \in (0, 1]$ meaning that W_0 has a reasonably smooth boundary.

We will take into consideration the following types of underlying point processes for our random structures:

- Let $\mathcal{X}_n = \delta_{X_1} + \cdots + \delta_{X_n}$, where $X_1, X_2, \ldots$ are i.i.d. random variables uniformly distributed in W_0,
- let $\mu_n = \delta_{Y_1} + \cdots + \delta_{Y_n}$ denote the binomial point process on W_n, where $Y_1, \ldots, Y_n$ are i.i.d. random variables uniformly distributed in W_n,
- let η_n be a homogeneous Poisson process on W_n with intensity λ. By η we denote the Poisson process with intensity λ on $\mathbb{R}^d$.

The point process $\mathcal{X}_n$ corresponds to the first scenario of the increasing non-random number of points in a fixed region while μ_n and η_n represent the second scenario of the increasing observation window.

Definition 3.5 (Strongly stabilizing functional, radius of stabilization).
The functional H is *strongly stabilizing* if there exists a.s. finite random variable S (a *radius of stabilization* of H) and $\Delta(\infty)$ such that with probability 1,

$$\Delta((\eta \cap B_S(\mathbf{o})) \cup \mathcal{A}) = \Delta(\infty), \quad \text{for all finite } \mathcal{A} \subset \mathbb{R}^d \backslash B_S(\mathbf{o}).$$

Hence S is a radius of stabilization if the add-one cost is almost surely not affected by inserting points into η outside the ball $B_S(\mathbf{o})$. In other words, the functional H is strongly stabilizing for the add-one cost method if and only if its add-one cost stabilizes in the sense of Definition 3.1.

Definition 3.6 (Weakly stabilizing functional).
The functional H is *weakly stabilizing* on $\mathcal{W}$ if there is a random variable $\Delta(\infty)$ such that

$$\Delta(\eta \cap A_n) \xrightarrow[n\to\infty]{a.s.} \Delta(\infty)$$

as $A_n \to \mathbb{R}^d$, $A_n \in \mathcal{W}$.

Roughly speaking, strong or weak stabilization says that the first order difference operator has a behaviour which is determined by local data.

Clearly, the strong stabilization implies weak stabilization on $\mathcal{W}$. Nevertheless, there is no general result comparing the strong stabilization or the weak stabilization with the point stabilization of Definition 3.1. For the Poisson input, however, if ξ is exponentially stabilizing with a sufficiently high exponent, then it is also weakly stabilizing for the add-one cost.

Definition 3.7 (Uniform bounded moments).
The functional H satisfies the *uniform bounded moments condition* on $\mathcal{W}$ if

$$\sup_{W\in\mathcal{W}:\mathbf{o}\in W}\ \sup_{m\in[\lambda|W|_d/2,3\lambda|W|_d/2]} \{\mathbb{E}\left[\Delta(U_{m,W})^4\right]\} < \infty,$$

where $U_{m,W}$ denotes the point process consisting of m independent uniform variables in W.

Definition 3.8 (Poisson bounded moments).
The functional H satisfies the *Poisson bounded moments condition* on $\mathcal{W}$ if

$$\sup_{W\in\mathcal{W}:\mathbf{o}\in W} \{\mathbb{E}\left[\Delta(\eta\cap W)^4\right]\} < \infty.$$

The Poisson bounded moments condition is weaker than the uniform bounded moments condition.

Definition 3.9 (Polynomially bounded functional).
The functional H is *polynomially bounded* if there exists a constant β_2 such that for all finite sets $\mathbf{x} \subset \mathbb{R}^d$, we have

$$|H(\mathbf{x})| \leqslant \beta_2(\operatorname{diam}(\mathbf{x}) + \#(\mathbf{x}))^{\beta_2},$$

where $\operatorname{diam}(\mathbf{x})$ is the diameter of $\mathbf{x}$ and $\#(\mathbf{x})$ stands for the cardinality of $\mathbf{x}$.

The Poisson bounded moments condition is weaker than the uniform moments condition in the sense of the following lemma.

Lemma 3.1 (Lemma 4.1 in Penrose and Yukich [2001]).
If H is polynomially bounded and satisfies the uniform bounded moments condition, then H satisfies the Poisson bounded moments condition.

Definition 3.10 (Homogeneous functional).
We say H is *homogeneous of order* γ if for all $\mathbf{x} \in \mathbf{N}$ and $a \in \mathbb{R}$,

$$H(a\mathbf{x}) = a^{\gamma} H(\mathbf{x}).$$

General results

The two following results for Poisson and binomial point processes are shown in Penrose and Yukich [2001] (cf. Theorem 2.1, Corollary 2.1 and Theorem 3.1).

Theorem 3.1 (CLT for functionals of a Poisson point process).
Suppose that H is weakly stabilizing on $\mathcal{W}$ and satisfies the Poisson bounded moments condition on $\mathcal{W}$. Then there exists $\sigma^2 \geqslant 0$ such that as $n \to \infty$,

$$n^{-1}\mathrm{Var}\,(H(\eta_n)) \to \sigma^2$$

and

$$n^{-1/2}(H(\eta_n) - \mathbb{E}\,H(\eta_n)) \xrightarrow{\mathcal{D}} N(0, \sigma^2),$$

where by $N(c, d)$ we denote a Gaussian random variable with mean value c and variance d.

Theorem 3.2 (CLT for functionals of a binomial point process).
Suppose that H is strongly stabilizing, satisfies the uniform bounded moments condition on $\mathcal{W}$, and is polynomially bounded. Then there exist constants $\tau^2 \geqslant 0$ such that as $n \to \infty$,

$$n^{-1}\mathrm{Var}\,(H(\mu_n)) \to \tau^2$$

and

$$n^{-1/2}(H(\mu_n) - \mathbb{E}\,H(\mu_n)) \xrightarrow{\mathcal{D}} N(0, \tau^2).$$

Also, given λ, σ^2 from Theorem 3.1 and τ^2 are independent of the choice of $(W_n)_{n\geqslant 1}$ and $\tau^2 \leqslant \sigma^2$.

If, moreover, $\Delta(\infty)$ is non-degenerate, then $\tau^2 > 0$, and hence also $\sigma^2 > 0$.

Theorem 3.3 (CLT for functionals of increasing sample size in a fixed region).
Suppose H is strongly stabilizing, satisfies uniform bounded moments condition on $\mathcal{W}_0$, is polynomially bounded and homogeneous of order γ. Then with τ^2 from Theorem 3.2 corresponding to $\lambda = 1$, as $n \to \infty$,

$$n^{2\gamma/d-1}\mathrm{Var}\,(H(\mathcal{X}_n)) \to \tau^2$$

and

$$n^{\gamma/d-1/2}(H(\mathcal{X}_n) - \mathbb{E}\,H(\mathcal{X}_n)) \xrightarrow{\mathcal{D}} N(0, \tau^2).$$

The main ingredient in the proof of Theorem 3.1 is a representation of $H(\eta_n) - \mathbb{E}\,H(\eta_n)$ as a sum of suitable martingale differences and then application of the central limit theorem for such object (Theorem 2.3 in McLeish [1974]). Then, by *de-Poissonizing* (see Section 2.5 in Penrose [2003]) the limits of Theorem 3.1 and using coupling of η_n and μ_n, one can prove Theorem 3.2.

Remark. 1. Theorem 3.1 is valid also the under weaker condition of convergence in probability in the definition of the weakly stabilizing functionals (Definition 3.6).

2. Theorem 3.1 and 3.2 were generalized to the setting of marked point processes in Penrose and Yukich [2002].

Example 3.1 (k-th nearest neighbour graph).
Fix $k \in \mathbb{N}$ and let H be the total edge length of the undirected k-th nearest neighbour graph.

Lemma 3.2.
The statistics H is polynomially bounded with $H(\mathbf{x}) \leqslant k \operatorname{diam}(\mathbf{x})\#(\mathbf{x})$ for any $\mathbf{x} \in \mathbf{N}$ and it is strongly stabilizing. If, moreover, $\mathcal{W}$ is regular, then H satisfies the uniform bounded moments condition on $\mathcal{W}$.

Theorem 3.4 (CLT and variance asymptotics for the k-th nearest neighbour graph).
Fix $k \in \mathbb{N}$ and assume that η_n is a unit intensity Poisson process on W_n. For $\mathbf{x} \in \mathbf{N}$, let $H(\mathbf{x})$ denote the total edge length of the undirected k-th nearest neighbour graph on $\mathbf{x}$. Provided $\mathcal{W}$ is regular, there exists $\sigma^2 > 0$ such that as $n \to \infty$,

$$n^{-1}\operatorname{Var}(H(\eta_n)) \to \sigma^2$$

and

$$n^{-1/2}(H(\eta_n) - \mathbb{E}\, H(\eta_n)) \xrightarrow{\mathcal{D}} N(0, \sigma^2).$$

Additionally, there exists $\tau^2 \in (0, \sigma^2]$ such that as $n \to \infty$,

$$n^{-1}\operatorname{Var}(H(\mu_n)) \to \tau^2$$

and

$$n^{-1/2}(H(\mu_n) - \mathbb{E}\, H(\mu_n)) \xrightarrow{\mathcal{D}} N(0, \tau^2).$$

Moreover, if $\mathcal{W}_0$ is regular, then with $n \to \infty$,

$$n^{2/d-1}\operatorname{Var}(H(\mathcal{X}_n)) \to \tau^2$$

and

$$n^{1/d-1/2}(H(\mathcal{X}_n) - \mathbb{E}\, H(\mathcal{X}_n)) \xrightarrow{\mathcal{D}} N(0, \tau^2).$$

We refer to Section 6 in Penrose and Yukich [2001] for more details and the proofs of Lemma 3.2 and Theorem 3.4.

Other functionals of interest of the k-th nearest neighbour graph can be investigated based on Theorems 3.1, 3.2 and 3.3. Those include for instance the number of components (i.e. subgraphs in which there is a path connecting each two vertices and no connection to any other vertex from the rest of the graph) with omitting the condition of the regularity of $\mathcal{W}$ and $\mathcal{W}_0$.

3.2 Moment approach

Another way how to obtain Gaussian limits for functionals of Poisson and binomial input in $\mathbb{R}^d$ is based on the cumulant method described in detail in Chapter 4. This approach, which takes into account stabilization of scores, has the particular benefit of describing the limiting variance. Moreover, it admits the non-homogeneous inputs.

Assumptions

Let $\mathcal{W}$ be the set of all compact, convex sets in $\mathbb{R}^d$ with non-empty interior together with $\mathbb{R}^d$ itself. For $W \in \mathcal{W}$ and κ a probability density with support on W, we denote by $\eta_{\lambda\kappa}$ the Poisson point process on W with intensity measure $\lambda\kappa, \lambda \geqslant 1$, and μ_n the binomial point process on W with points distributed according to κ.

Let $\xi(x, \mathbf{x})$ be a measurable real-valued function defined for all pairs $(x, \mathbf{x})$, where $\mathbf{x}$ is a finite point set in $\mathbb{R}^d$ and $x \in \mathbf{x}$ and we assume it is translation-invariant (so that $\xi(x + y, \mathbf{x} + y) = \xi(x, \mathbf{x})$ for all finite $\mathbf{x} \subset \mathbb{R}^d$ and $y, x \in \mathbb{R}^d$ where $\mathbf{x} + y = \{x + y, x \in \mathbf{x}\}$). The authors in Baryshnikov and Yukich [2005] considered random measures of the form

$$\mathcal{H} := \sum_{x \in \mathbf{x}} \xi(x, \mathbf{x}) \delta_x.$$

Note that the definition of H^ξ in (3.1) corresponds to

$$H(\mathbf{x}) = \mathcal{H}(\mathbb{R}^d).$$

For $\lambda > 0$, we denote $\xi_\lambda(x, \mathbf{x}) := \xi(\lambda^{1/d} x, \lambda^{1/d} \mathbf{x})$ the rescaled version of ξ. Similarly, let $\xi_n(x, \mathbf{x}) := \xi(n^{-1/d} x, n^{-1/d} \mathbf{x})$ for $n \in \mathbb{N}$. We define the corresponding rescaled measures by

$$\mathcal{H}_\lambda := \sum_{x \in \eta_{\lambda\kappa}} \xi_\lambda(x, \eta_{\lambda\kappa}) \delta_x, \quad \lambda \geqslant 1,$$

resp.

$$\mathcal{H}'_n := \sum_{x \in \mu_n} \xi_n(x, \mu_n) \delta_x, \quad n \in \mathbb{N}.$$

We show that $\mathcal{H}_\lambda$, resp. $\mathcal{H}'_n$ converges weakly to a Gaussian field with a covariance functional described in terms of the score function ξ and the choice of κ.

Recall that measures ρ_n converge to a Gaussian field as $n \to \infty$ if their finite-dimensional distributions converge to those of a Gaussian field. By the convergence of finite-dimensional distributions we mean the convergence in distribution of the integrals $\int f \mathrm{d}\rho_n$ to the corresponding normal random variables for all continuous functions f.

For $0 \leqslant a < b < \infty$ we define the set of test functions $\mathcal{F}(a, b)$, i.e. the set of $f : \mathbb{R}^d \to \mathbb{R}^+$ with support in $\mathcal{W}$ and range in $[a, b] \cup \{0\}$.

Definition 3.11 (Stabilizing scores, radius of stabilization).
The score ξ is said to be *stabilizing* if for all $W \in \mathcal{W}$, $0 \leqslant a < b < \infty, \lambda > 0$ and all $x \in \lambda W := \{\lambda w, w \in W\}$, there exists an almost surely finite random

variable $R(x) := R(x, \lambda, a, b, W)$ (a *radius of stabilization* for ξ at x) such that for all $f \in \mathcal{F}(a, b)$ with $\operatorname{supp} f = \lambda W$, and all finite $\mathbf{x} \subset \lambda W \backslash B_{R(x)}(x)$ we have

$$\xi(x, (\eta_f \cap B_{R(x)}(x)) \cup \mathbf{x}) = \xi(x, \eta_f \cap B_{R(x)}(x))$$

and, moreover, $\sup_{x \in \mathbb{R}^d} \mathbb{P}(R(x, \lambda, a, b, W) > t) \to 0$ as $t \to \infty$.

Definition 3.12 (Exponentially stabilizing scores).
We say the score ξ is *exponentially stabilizing* if for all $W \in \mathcal{W}$ and $0 \leqslant a < b < \infty$ the tail probabilities $r(t) := r(t, a, b, \mathcal{W}) := \sup_{x \in \mathbb{R}^d, \lambda > 0} \mathbb{P}(R(x, \lambda, a, b, W) \geqslant t)$ decays exponentially in t.

We say ξ is *polynomially stabilizing* if for all a, b and $W \in \mathcal{W}$ we have

$$\int_0^\infty (r(t))^{1/2} t^{d-1} \mathrm{d}t < \infty.$$

Definition 3.13 (p-moment condition 1).
We say that a score ξ satisfies a *moment condition* 1 of order $p > 0$ with respect to κ if

$$\sup_{\lambda > 0, x \in [0, \lambda^{1/d}]^d W, \mathcal{A} \subset \mathbb{R}^d \text{ finite}} \mathbb{E}\left[|\xi_\lambda(x, \eta_{\lambda\kappa} \cup \mathcal{A})|^p\right] < \infty$$

and for all $\lambda > 0$

$$\sup_{x \in \mathbb{R}^d, \mathcal{A} \subset \mathbb{R}^d \text{ finite}} \mathbb{E}\left[|\xi(x, \eta_\lambda \cup \mathcal{A})|^p\right] < \infty.$$

General results

Theorem 3.5 (Theorem 2.1 in Baryshnikov and Yukich [2005]).
If ξ is polynomially stabilizing and satisfies the p-moment condition 1 for $p = 4$ then

$$\frac{\operatorname{Var} \int_W f \mathrm{d}\mathcal{H}_{\lambda\kappa}}{\lambda} \xrightarrow[\lambda \to \infty]{} \int_W f^2(x) V^\xi(\kappa(x)) \kappa(x) \mathrm{d}x,$$

where

$$V^\xi(t) := \mathbb{E}\, \xi^2(\mathbf{o}, \eta_t) + \int_{\mathbb{R}^d} t(\mathbb{E}\, \xi(\mathbf{o}, \eta_t \cup \{y\}) \cdot \xi(y, \eta_t \cup \{\mathbf{o}\}) - \mathbb{E}\, \xi(\mathbf{o}, \eta_t) \mathbb{E}\, \xi(y, \eta_t')) \mathrm{d}y \tag{3.2}$$

and η_t' denotes an independent copy of η_t.

If, moreover, ξ is exponentially stabilizing and satisfies the moment condition 1 for all $p > 0$, then $\lambda^{-1/2}(\mathcal{H}_{\lambda\kappa} - \mathbb{E}\, \mathcal{H}_{\lambda\kappa})$ converges as $\lambda \to \infty$ to a Gaussian field with covariance kernel $\int_W f_1(x) f_2(x) V^\xi(\kappa(x)) \kappa(x) \mathrm{d}x$.

Theorem 3.5 can be reformulated for the binomial input under the assumptions of strong stability and uniform bounded moments condition similar to those in Definitions 3.5 and 3.7. Recall that those conditions were defined for the functional $H(\mathbf{x}) = \sum_{x \in \mathbf{x}} \xi(x, \mathbf{x})$ for the purposes of binomial input driven by the uniform distribution. For the case of general density κ, we update the definition saying that H satisfies *uniformly bounded moments condition for κ* if

$$\sup_n \sup_{x \in n^{1/d} A} \sup_{m \in [n/2, 3n/2]} \mathbb{E}\left[(D_x H(U_{m,n}))^4\right] < \infty,$$

where $U_{m,n}$ is a point process consisting of m i.i.d. random variables of the form $n^{1/d}U$ on $n^{1/d}W$, where U has density κ. Further, we say that H is *strongly stabilizing for* κ if for all $t > 0$ there exists an almost surely finite random variable S and $\Delta(t)$ such that with probability 1,

$$\Delta(t) = \Delta((\eta_t \cap B_S(\mathbf{o})) \cup \mathcal{A}), \tag{3.3}$$

for all finite $\mathcal{A} \subset \mathbb{R}^d \backslash B_S(\mathbf{o})$. Recall the notation of Definition 3.5.

Theorem 3.6 (Theorem 2.2 in Baryshnikov and Yukich [2005]).
Assume that ξ is polynomially stabilizing, satisfies the p-moment condition 1 for $p = 4$ and H is a strongly stabilizing functional satisfying the uniform bounded moments condition for κ. Then for all continuous functions f with support on W,

$$\frac{\operatorname{Var} \int_W f \mathrm{d}\mathcal{H}'_n}{n} \xrightarrow[n\to\infty]{} \int_W f^2(x) V^\xi(\kappa(x))\kappa(x)\mathrm{d}x - \left(\int_W f(x)\mathbb{E}\,\Delta(\kappa(x))\kappa(x)\mathrm{d}x\right)^2.$$

If ξ is exponentially stabilizing, satisfies the p-moment condition 1 for all $p > 0$ and H is a strongly stabilizing functional for κ, then $n^{-1/2}(\mathcal{H}'_n - \mathbb{E}\,\mathcal{H}'_n)$ converges as $n \to \infty$ to a Gaussian field with covariance kernel

$$\int_W f_1(x) f_2(x) V^\xi(\kappa(x))\kappa(x)\mathrm{d}x \\ - \int_W f_1(x)\Delta(\kappa(x))\kappa(x)\mathrm{d}x \int_W f_2(x)\Delta(\kappa(x))\kappa(x)\mathrm{d}x.$$

Remark.

1. Theorems 3.5 and 3.6 can be further generalized for score functions that are not translation-invariant (see Section 2.3.2 in Baryshnikov and Yukich [2005]).
2. If the distribution of $\Delta(\kappa(U))$, U being a random variable with density κ, is non-degenerate, then the limiting Gaussian field is non-degenerate.
3. Theorems 3.5 and 3.6 generalize Theorems 3.1 and 3.2 by showing the convergence of measures induced by not necessarily homogeneous point sets to Gaussian random field and they identify the limiting variance.
4. The evaluation of the limiting variances can be rather difficult, but simplify under the additional assumption of homogeneity (see Definition 3.10). If ξ is homogeneous of order γ, then

$$V^\xi(t) = V^\xi(1)t^{-2\gamma/d} \quad \text{and} \quad \Delta(t) = \Delta(1)t^{-\gamma/d}.$$

Example 3.2 (k-nearest neighbour graph).
Let k be a fixed positive integer. For a point set $\mathcal{X} \subset \mathbb{R}^d$, we denote by $NG(\mathcal{X})$ the undirected k-nearest neighbour graph induced by $\mathcal{X}$ and by $\mathcal{E}(x, NG(\mathcal{X}))$ the set of edges in $NG(\mathcal{X})$ incident to $x \in \mathcal{X}$.

The application of the preceding results is demonstrated only on the binomial input. Assume $\mu_n := \{X_1, \dots, X_n\}$ is a binomial sample in $\mathbb{R}^d$ driven by density

κ. We define a random measure

$$\mathcal{H}'_{\phi,n} := \sum_{i=1}^{n} \sum_{e \in \mathcal{E}(n^{1/d} X_i, NG(\mu_n))} \phi(|e|) \delta_{X_i},$$

where $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ is some weight function.

Theorem 3.7 (Theorem 3.1 in Baryshnikov and Yukich [2005]). *Suppose that ϕ has a polynomial growth, i.e. there exist $C, a < \infty$ such that $\phi(x) \leqslant C(1 + x^a)$ for all $x \in \mathbb{R}^+$. Moreover, assume that κ is bounded away from infinity and zero on its support* $\operatorname{supp}\kappa = W \in \mathcal{W}$. *Then*

$$\frac{\operatorname{Var} \int_W f \mathrm{d}H'_{\phi,n}}{n} \xrightarrow[n\to\infty]{} \int_W f^2(x) V_\phi(\kappa(x))\kappa(x)\mathrm{d}x - \left(\int_W f(x)\Delta_\phi(\kappa(x))\kappa(x)\mathrm{d}x \right)^2$$

for all continuous functions f with support on W, where V_ϕ and Δ_ϕ are given by (3.2) *and* (3.3) *when* $\xi(x, \mathbf{x}) = \sum_{e \in \mathcal{E}(x, NG(\mathbf{x}))} \phi(|e|)$.

Moreover, as $n \to \infty$, $n^{-1/2}(\mathcal{H}'_{\phi,n} - \mathbb{E}\,\mathcal{H}'_{\phi,n})$ converges to a Gaussian field with covariance kernel

$$\int_W f_1(x) f_2(x) V_\phi(\kappa(x))\kappa(x)\mathrm{d}x \\ - \int_W f_1(x)\Delta_\phi(\kappa(x))\kappa(x)\mathrm{d}x \int_W f_2(x)\Delta_\phi(\kappa(x))\kappa(x)\mathrm{d}x$$

for any f_1, f_2 continuous with support on W.

Remark. If we set $\phi(x) = x/2$ for $x \in \mathbb{R}^+$ then we obtain a central limit theorem for the total edge length of the k-nearest neighbour graph. If $\phi(x)$ is either 0 or 1 depending on whether x is less than some $t \geqslant 0$ or not, then we obtain a central limit theorem for the empirical distribution function of the rescaled lengths of the edges.

3.3 Stabilization in the Malliavin–Stein bounds

Recently in Last et al. [2016], the Malliavin calculus combined with Stein's method of normal approximation (see Chapter 2 for the details) was proved to yield rates of normal approximation for general Poisson functionals. Moreover, the authors used their general results to deduce central limit theorems together with rates of convergence in terms of Kolmogorov and Wasserstein distance for Poisson functionals satisfying a type of stabilization.

Recall that the Malliavin–Stein rates of normal convergence are expressed in terms of moments of the first- and second-order difference operators. Those can be difficult to evaluate, yet if combined with the stabilization property, the bounds remarkably simplify.

First, assume that $(\mathbb{X}, \mathcal{X})$ is a measurable space and η_λ is a Poisson point process with intensity measure $\alpha_\lambda = \lambda\mathbb{Q}$, where $\lambda \geqslant 1$ and $\mathbb{Q}$ is a fixed finite measure on $\mathbb{X}$. Moreover, we let $F_\lambda \in L^0_{\eta_\lambda}$ be a Poisson functional with a representative f_λ (recall the notation at the beginning of Section 2.2).

Theorem 3.8 (Theorem 6.1 in Last et al. [2016]).
Let $F \in \operatorname{dom} D$ *(recall Definition 2.9) with* $\operatorname{Var} F > 0$*. Assume there are constants* $c_1, c_2, p_1, p_2 > 0$ *such that*

$$\mathbb{E}\,|D_x F|^{4+p_1} \leqslant c_1, \quad \text{for } \mathbb{Q}\text{-a.e. } x \in \mathbb{X}, \tag{3.4}$$

$$\mathbb{E}\,|D^2_{x_1,x_2} F|^{4+p_2} \leqslant c_2, \quad \text{for } \mathbb{Q}^2\text{-a.e. } (x_1,x_2) \in \mathbb{X}^2 \tag{3.5}$$

and denote $c := \max\{1, c_1, c_2\}$*. Then*

$$\begin{aligned} &d_W\left(\frac{F - \mathbb{E}\,F}{\sqrt{\operatorname{Var} F}}, Z\right) \\ &\leqslant \frac{5c}{\operatorname{Var} F}\left[\int_{\mathbb{X}}\left(\int_{\mathbb{X}} \mathbb{P}(D^2_{x_1,x_2}F \neq 0)^{p_2/(16+4p_2)}\mathbb{Q}(\mathrm{d}x_2)\right)^2 \mathbb{Q}(\mathrm{d}x_1)\right]^{1/2} \\ &+ \frac{c}{(\operatorname{Var} F)^{3/2}}\int_{\mathbb{X}} \mathbb{P}(D_x F \neq 0)^{(1+p_1)/(4+p_1)}\mathbb{Q}(\mathrm{d}x) \end{aligned}$$

and

$$\begin{aligned} &d_K\left(\frac{F - \mathbb{E}\,F}{\sqrt{\operatorname{Var} F}}, Z\right) \\ &\leqslant c\left\{\frac{5}{\operatorname{Var} F}\left[\int_{\mathbb{X}}\left(\int_{\mathbb{X}} \mathbb{P}(D^2_{x_1,x_2}F \neq 0)^{p_2/(16+4p_2)}\mathbb{Q}(\mathrm{d}x_2)\right)^2 \mathbb{Q}(\mathrm{d}x_1)\right]^{1/2}\right. \\ &+ \frac{\Gamma_F^{1/2}}{\operatorname{Var} F} + \frac{2\Gamma_F}{(\operatorname{Var} F)^{3/2}} + \frac{\Gamma_F^{5/4} + 2\Gamma_F^{3/2}}{(\operatorname{Var} F)^2} \\ &\left.+ \frac{\sqrt{6}+\sqrt{3}}{\operatorname{Var} F}\left[\int_{\mathbb{X}} \mathbb{P}(D^2_{x_1,x_2}F \neq 0)^{p_2/(8+2p_2)}\mathbb{Q}^2(\mathrm{d}(x_1,x_2))\right]^{1/2}\right\}, \end{aligned}$$

where

$$\Gamma_F := \int_{\mathbb{X}} \mathbb{P}(D_x F \neq 0)^{p_1/(8+2p_1)}\mathbb{Q}(\mathrm{d}x)$$

and Z *denotes the standard Gaussian random variable.*

Proof. The proof of Theorem 3.8 consists of estimating the terms $\alpha_{F,1}, \ldots, \alpha_{F,6}$ in Theorems 2.13 and 2.14. To ease the notation, we denote by $v_F : \mathbb{X} \to [0,1]$ and $w_F : \mathbb{X}^2 \to [0,1]$ the functions

$$v(x) := \mathbb{P}(D_x F \neq 0),\; w(x,y) := \mathbb{P}(D^2_{x,y}F \neq 0), \quad x, y \in \mathbb{X}.$$

By Hölder's inequality, (3.4) and (3.5), we have that

$$\begin{aligned} \mathbb{E}\,(D_x F)^4 &\leqslant [v(x)]^{p_1/(4+p_1)}\left[\mathbb{E}\,|D_x F|^{4+p_1}\right]^{4/(4+p_1)} \\ &\leqslant c_1^{4/(4+p_1)}[v(x)]^{p_1/(4+p_1)}, \\ \mathbb{E}\,|D_x F|^3 &\leqslant c_1^{3/(4+p_1)}[v(x)]^{(1+p_1/(4+p_1))} \end{aligned}$$

for $\mathbb{Q}$-a.e. $x \in \mathbb{X}$ and

$$\begin{aligned}\mathbb{E}\,(D^2_{x_1,x_2})^4 &\leqslant [w(x_1,x_2)]^{p_2/(4+p_2)}\left[\mathbb{E}\,|D^2_{x_1,x_2}F|^{4+p_2}\right]^{4/(4+p_2)}\\ &\leqslant c_2^{4/(4+p_2)}[w(x_1,x_2)]^{p_2/(4+p_2)}\end{aligned}$$

for $\mathbb{Q}^2$-a.e. $(x_1,x_2)\in\mathbb{X}^2$. By inserting the latter estimates in $\alpha_{F,1},\ldots,\alpha_{F,6}$ and further application of Hölder's inequality, we obtain that

$$\begin{aligned}
\alpha_{F,1} &\leqslant \frac{4c_1^{1/(4+p_1)}c_2^{1/(4+p_2)}}{\operatorname{Var} F}\left[\int_{\mathbb{X}^3}(w(x_1,x_3)w(x_2,x_3))^{p_2/(16+4p_2)}\,\mathbb{Q}^3(\mathrm{d}(x_1,x_2,x_3))\right]^{1/2},\\
\alpha_{F,2} &\leqslant \frac{c_2^{2/(4+p_2)}}{\operatorname{Var} F}\left[\int_{\mathbb{X}^3}(w(x_1,x_3)w(x_2,x_3))^{p_2/(8+2p_2)}\,\mathbb{Q}(\mathrm{d}(x_1,x_2,x_3))\right]^{1/2},\\
\alpha_{F,3} &\leqslant \frac{c_1^{3/(4+p_1)}}{(\operatorname{Var} F)^{3/2}}\int_{\mathbb{X}}[v(x)]^{(1+p_1)/(4+p_1)}\mathbb{Q}(\mathrm{d}x),\\
\alpha_{F,4} &\leqslant \frac{c_1^{3/(4+p_1)}}{2(\operatorname{Var} F)^2}\left[\mathbb{E}\,(F-\mathbb{E}\,F)^4\right]^{1/4}\int_{\mathbb{X}}[v(x)]^{p_1/(8+2p_1)}\mathbb{Q}(\mathrm{d}x),\\
\alpha_{F,5} &\leqslant \frac{c_1^{2/(4+p_1)}}{\operatorname{Var} F}\left[\int_{\mathbb{X}}[v(x)]^{p_1/(4+p_1)}\mathbb{Q}(\mathrm{d}x)\right]^{1/2},\\
\alpha_{F,6} &\leqslant \frac{\sqrt{6}c_1^{1/(4+p_1)}c_2^{1/(4+p_2)}}{\operatorname{Var} F}\left[\int_{\mathbb{X}^2}[w(x_1,x_2)]^{p_2/(8+2p_2)}\mathbb{Q}^2(\mathrm{d}(x_1,x_2))\right]^{1/2}\\
&\quad+\frac{\sqrt{3}c_2^{2/(4+p_2)}}{\operatorname{Var} F}\left[\int_{\mathbb{X}^2}[w(x_1,x_2)]^{p_2/(4+p_2)}\mathbb{Q}^2(\mathrm{d}(x_1,x_2))\right]^{1/2}.
\end{aligned}$$

Let $G\in L^2_\eta$ be such that $\mathbb{E}\,G=0, \operatorname{Var} G=1$ and $G=g(\eta)$ a.s. Then

$$\mathbb{E}\,G^4=\operatorname{Var} G^2+(\mathbb{E}\,G^2)^2=\operatorname{Var} G^2+1\leqslant\int_{\mathbb{X}}\mathbb{E}\,(D_xG^2)^2\mathbb{Q}(\mathrm{d}x)+1,$$

where we used the Poincaré inequality (see Corollary 2.1). Further,

$$\begin{aligned}D_xG^2 &= g^2(\eta+\delta_x)-g^2(\eta)+2g^2(\eta)-2g^2(\eta)+2g(\eta)g(\eta+\delta_x)-2g(\eta)g(\eta+\delta_x)\\ &=(g(\eta+\delta_x)-g(\eta))^2-2g^2(\eta)+2g(\eta)g(\eta+\delta_x)\\ &=(D_xG)^2+2GD_xG\end{aligned}$$

and

$$((D_xG)^2+2GD_xG)^2\leqslant 2(D_xG)^4+8G^2(D_xG)^2$$

almost surely. Combined with the Cauchy–Schwarz inequality, we arrive at

$$\begin{aligned}\mathbb{E}\,G^4 &\leqslant 8\left[\mathbb{E}\,G^4\right]^{1/2}\int_{\mathbb{X}}\left[\mathbb{E}\,(D_xG)^4\right]^{1/2}\mathbb{Q}(\mathrm{d}x)+2\int_{\mathbb{X}}\mathbb{E}\,(D_xG)^4\mathbb{Q}(\mathrm{d}x)+1\\ &\leqslant \max\left\{16\left[\mathbb{E}\,G^4\right]^{1/2}\int_{\mathbb{X}}\left[\mathbb{E}\,(D_xG)^4\right]^{1/2}\mathbb{Q}(\mathrm{d}x), 4\int_{\mathbb{X}}\mathbb{E}\,(D_xG)^4\mathbb{Q}(\mathrm{d}x)+2\right\}.\end{aligned}$$

Now, we substitute $G=(F-\mathbb{E}\,F)^4/(\operatorname{Var} F)^2$ to see that

$$\alpha_{F,4}\leqslant\frac{c_1^{3/(4+p_1)}}{(\operatorname{Var} F)^{3/2}}\Gamma_F+\frac{c_1^{4/(4+p_1)}}{(\operatorname{Var} F)^2}\Gamma_F^{5/4}+\frac{2c_1^{4/(4+p_1)}}{(\operatorname{Var} F)^2}\Gamma_F^{3/2}.$$

Combining the estimates for $\alpha_{F,1}\dots,\alpha_{F,6}$ concludes the proof.

□

Corollary 3.1 (Proposition 1.3 in Last et al. [2016]).
Let $F_t \in L^2_{\eta_t}$ and assume there exist finite constants $p_1, p_2, c, v > 0$ such that

$$\mathbb{E}\,|D_x F_t|^{4+p_1} \leqslant c, \quad \text{for } \mathbb{Q}\text{-a.e. } x \in \mathbb{X}, \quad t \geqslant 1,$$

$$\mathbb{E}\,|D^2_{x_1,x_2} F_t|^{4+p_2} \leqslant c, \quad \text{for } \mathbb{Q}^2\text{-a.e. } (x_1,x_2) \in \mathbb{X}^2, \quad t \geqslant 1,$$

$$\frac{\operatorname{Var} F_t}{t} > v, \quad t \geqslant 1$$

and

$$m := \sup_{x\in\mathbb{X}, t\geqslant 1} \int_{\mathbb{X}} \mathbb{P}(D^2_{x,y} F_t \neq 0)^{p_2/(16+4p_2)} \alpha_\lambda(\mathrm{d}y) < \infty. \tag{3.6}$$

Then there exists a finite constant C depending uniquely on $c, p_1, p_2, v, \mathbb{Q}(\mathbb{X})$ and m such that

$$\max\left\{ d_W\left(\frac{F_t - \mathbb{E}\,F_t}{\sqrt{\operatorname{Var} F_t}}, Z\right), d_K\left(\frac{F_t - \mathbb{E}\,F_t}{\sqrt{\operatorname{Var} F_t}}, Z\right)\right\} \leqslant C t^{-1/2}.$$

Remark. If we relax the assumptions on the generality of $\mathbb{X}$, so that $\mathbb{X}$ is a compact subset of $\mathbb{R}^d$ and $\mathbb{Q}$ a restriction of the Lebesgue measure $|\cdot|_d$ on $\mathbb{X}$, then the assumption (3.6) in Corollary 3.1 can be replaced the following assumptions concerning stabilization. The Poisson functionals F_t are strongly stabilizing (in the sense of Definition 3.5) if for each $x \in \mathbb{X}$ there exists an almost surely finite random variable $R_t(x, \eta_t)$ (radius of stabilization) such that

$$D_x f_t(\eta_t) = D_x f_t(\eta_t \cap B(x, R_t(x, \eta_t))), \text{ a.s.}$$

Moreover, we assume that

$$\sup_{x\in\mathbb{X}, t\geqslant 1} \int_{\mathbb{X}} t\mathbb{P}(y \in B(x, R_t(x,\eta_t)) \text{ or } R_t(x, \eta_t + \delta_y) \neq R_t(x,\eta_t))^\alpha \mathrm{d}y < \infty$$

for some suitable α.

Now, let $(\mathbb{X}, \mathcal{X})$ be a measurable space equipped with a σ-finite measure $\mathbb{Q}$ and a semi-metric $d : \mathbb{X}\times\mathbb{X} \to [0,\infty)$. By $B^{\mathbb{X}}_r(x)$ we denote the ball of radius $r > 0$ around $x \in \mathbb{X}$ with respect to d. We assume that there are constants $\gamma, \kappa > 0$ such that

$$\limsup_{\epsilon\to 0} \frac{\mathbb{Q}(B^{\mathbb{X}}_{r+\epsilon}(x)) - \mathbb{Q}(B^{\mathbb{X}}_r(x))}{\epsilon} \leqslant \kappa\gamma r^{\gamma-1}. \tag{3.7}$$

Apart from the Euclidean space $\mathbb{R}^d$, the condition holds also for m-dimensional Riemannian manifolds, where $m \leqslant d$. More examples are listed in Lachièze-Rey et al. [2019].

The approach of Last et al. [2016] was further revised in Lachièze-Rey et al. [2019] and extended in some directions. The authors established presumably optimal rates of normal convergence with respect to the Kolmogorov distance for functionals of marked Poisson and binomial point processes.

In order to deal with marked point processes, let us first rewrite Definition 3.1 and 3.2 in an appropriate way. Let $(\mathbb{M}, \mathcal{M}, \mathbb{P}_{\mathbb{M}})$ be a probability space (the mark space) and denote $\hat{\mathbb{X}} := \mathbb{X} \times \mathbb{M}$. We consider score functions $\xi_\lambda, \xi_n : \hat{\mathbb{X}} \times \mathbf{N}(\hat{\mathbb{X}}) \to \mathbb{R}$, $\lambda \geqslant 1, n \in \mathbb{N}$ to be measurable functions defined on pairs $((x, m), \mathbf{x})$, where $\mathbf{x} \in \mathbf{N}(\hat{\mathbb{X}})$ and $(x, m) \in \mathbf{x}$.

For $\lambda \geqslant 1$, we consider η_λ to be a marked Poisson point process on $\mathbb{X}$ with intensity measure $\lambda\mathbb{Q}$. Alternatively, if $\mathbb{Q}$ is a probability measure, we let μ_n to be a marked binomial point process of n point distributed independently according to $\mathbb{Q}$. Then, we denote

$$H_\lambda := \sum_{x \in \eta_\lambda} \xi_\lambda(x, \eta_\lambda),\ \lambda \geqslant 1,$$
$$H'_n := \sum_{x \in \mu_n} \xi_n(x, \mu_n),\ n \in \mathbb{N}.$$

Definition 3.14 (Stabilization scores of marked point processes).
A score function ξ_λ is *stabilizing* if there is a measurable map $R_\lambda : \hat{\mathbb{X}} \times \mathbf{N}(\hat{\mathbb{X}}) \to \mathbb{R}$ (a *radius of stabilization*) such that for all $\hat{x} := (x, m_x) \in \hat{\mathbb{X}}, \hat{\mathbf{x}} \in \mathbf{N}(\hat{\mathbb{X}})$ and $\hat{\mathcal{A}} \subset \hat{\mathbb{X}}$ with $\#(\hat{\mathcal{A}}) \leqslant 7$ we have

$$\xi_\lambda(\hat{x}, \hat{\mathbf{x}} \cup \hat{\mathcal{A}}) = \xi_\lambda(\hat{x}, (\hat{\mathbf{x}} \cup \hat{\mathcal{A}}) \cap \hat{B}(x, R_\lambda(\hat{x}, \hat{\mathbf{x}} \cup \{\hat{x}\}))),$$

where $\hat{B}(y, r) := B_r^{\mathbb{X}}(y) \times \mathbb{M}$ for $y \in \mathbb{X}$ and $r > 0$.

For a given point $x \in \mathbb{X}$ we denote by M_x the corresponding random mark, which is distributed according to $\mathbb{P}_{\mathbb{M}}$ and is independent of everything else. Moreover, for a finite set $\mathcal{A} \subset \mathbb{X}$, we denote by $(\mathcal{A}, M_{\mathcal{A}})$ the set obtained by equipping each point of $\mathcal{A}$ with a random mark distributed according to $\mathbb{P}_{\mathbb{M}}$ independently of everything else.

Definition 3.15 (Exponentially stabilizing scores of marked Poisson and binomial point processes).
We say that $(\xi_\lambda)_{\lambda \geqslant 1}$ (resp. $(\xi_n)_{n \in \mathbb{N}}$) are *exponentially stabilizing* if there are radii of stabilization $(R_\lambda)_{\lambda \geqslant 1}$ (resp. $(R_n)_{n \in \mathbb{N}}$) and constants $c_{stab}, \alpha_{stab} \in (0, \infty)$ such that for $x \in \mathbb{X}, r \geqslant 0$ and $\lambda \geqslant 1$

$$\mathbb{P}(R_\lambda((x, M_x), \eta_\lambda \cup \{(x, M_x)\}) \geqslant r) \leqslant c_{stab} \exp\left(-\frac{1}{c_{stab}}(\lambda^{1/\gamma} r)^{\alpha_{stab}}\right),$$

resp. for $x \in \mathbb{X}, r \geqslant 0$ and $n \geqslant 9$

$$\mathbb{P}(R_n((x, M_x), \mu_{n-8} \cup \{(x, M_x)\}) \geqslant r) \leqslant c_{stab} \exp\left(-\frac{1}{c_{stab}}(n^{1/\gamma} r)^{\alpha_{stab}}\right),$$

where γ is from (3.7).

Definition 3.16 ($(4 + p)$th moment condition).
Given $p \in [0, \infty)$, we say that $(\xi_\lambda)_{\lambda \geqslant 1}$, resp. $(\xi_n)_{n \in \mathbb{N}}$ satisfy $(4 + p)$th moment

condition if there is a constant $C_p \in (0, \infty)$ such that for all $\mathcal{A} \subset \mathbb{X}$ with $\#(\mathcal{A}) \leqslant 7$ we have

$$\sup_{\lambda \in [1,\infty)} \sup_{x \in \mathbb{X}} \mathbb{E}\, |\xi_\lambda((x, M_x), \eta_\lambda \cup \{(x, M_x)\} \cup (\mathcal{A}, M_\mathcal{A}))|^{4+p} \leqslant C_p,$$

resp.

$$\sup_{n \in \mathbb{N}, n \geqslant 9} \sup_{x \in \mathbb{X}} \mathbb{E}\, |\xi_n((x, M_x), \mu_{n-8} \cup \{(x, M_x)\} \cup (\mathcal{A}, M_\mathcal{A}))|^{4+p} \leqslant C_p.$$

Definition 3.17 (Exponentially fast decaying scores).
Let $K \subset \mathbb{X}$ be measurable. Denote by $d(x, K) := \inf_{y \in K} d(x, y)$ the distance between a point $x \in \mathbb{X}$ and the set K. We say that $(\xi_\lambda)_{\lambda \geqslant 1}$, resp. $(\xi_n)_{n \in \mathbb{N}}$ *decay exponentially fast with the distance to* K if there are constants $c_K, \alpha_K \in (0, \infty)$ such that for all $\mathcal{A} \subset \mathbb{X}$ with $\#(\mathcal{A}) \leqslant 7$ we have for $x \in \mathbb{X}$ and $\lambda \geqslant 1$,

$$\mathbb{P}(\xi_\lambda((x, M_x), \eta_\lambda \cup \{(x, M_x)\} \cup (\mathcal{A}, M_\mathcal{A})) \neq 0) \leqslant c_K \exp\left(-c_K^{-1}(\lambda^{1/\gamma} d(x, K))^{\alpha_K}\right),$$

resp. for $x \in \mathbb{X}$ and $n \geqslant 9$,

$$\mathbb{P}(\xi_n((x, M_x), \mu_{n-8} \cup \{(x, M_x)\} \cup (\mathcal{A}, M_\mathcal{A})) \neq 0) \leqslant c_K \exp\left(-c_K^{-1}(n^{1/\gamma} d(x, K))^{\alpha_K}\right).$$

Remark. Definition 3.17 describes scores whose variances exhibit surface area order scaling. When dealing with volume order scaling, one can put $K = \mathbb{X}, c_K = 1$ and choose an arbitrary $\alpha_K \in (0, \infty)$.

Theorem 3.9 (Theorem 2.1 in Lachièze-Rey et al. [2019]).
Denote $\alpha := \min\{\alpha_{stab}, \alpha_K\}$ *and*

$$I_{K,t} := t \int_{\mathbb{X}} \exp\left(-\frac{\min\{c_{stab}, c_K\} p (t^{1/\gamma} d(x, K))^\alpha}{36 \cdot 4^{\alpha+1}}\right) \mathbb{Q}(\mathrm{d}x), \quad t \in \mathbb{R}. \tag{3.8}$$

(a) Assume that the score functions $(\xi_\lambda)_{\lambda \geqslant 1}$ *are exponentially stabilizing, satisfy the* $(4+p)$*th moment condition for some* $p \in (0, 1]$ *and decay exponentially fast with distance to a measurable set* $K \subset \mathbb{X}$. *Then there exists a constant* $C \in (0, \infty)$ *only depending on the constants* γ *in* (3.7), α, c_{stab}, C_p *and* c_K *from Definitions 3.15, 3.16 and 3.17 such that*

$$d_K\left(\frac{H_\lambda - \mathbb{E}\, H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}\right) \leqslant C\left(\frac{\sqrt{I_{K,\lambda}}}{\operatorname{Var} H_\lambda} + \frac{I_{K,\lambda}}{(\operatorname{Var} H_\lambda)^{3/2}} + \frac{I_{K,\lambda}^{5/4} + I_{K,\lambda}^{3/2}}{(\operatorname{Var} H_\lambda)^2}\right), \quad \lambda \geqslant 1.$$

(b) Assume that the score functions $(\xi_n)_{n \in \mathbb{N}}$ *are exponentially stabilizing, satisfy the* $(4+p)$*th moment condition for some* $p \in (0, 1]$ *and decay exponentially fast with distance to a measurable set* $K \subset \mathbb{X}$. *Then there exists a constant* $C' \in (0, \infty)$ *only depending on the constants* $\gamma, \alpha, c_{stab}, c_K$ *and* C_p *such that*

$$d_K\left(\frac{H'_n - \mathbb{E}\, H'_n}{\sqrt{\operatorname{Var} H'_n}}\right) \leqslant C'\left(\frac{\sqrt{I_{K,n}}}{\operatorname{Var} H'_n} + \frac{I_{K,n}}{(\operatorname{Var} H'_n)^{3/2}} + \frac{I_{K,n} + I_{K,n}^{3/2}}{(\operatorname{Var} H'_n)^2}\right), \quad n \geqslant 9.$$

Theorem 3.9 is a consequence of the general result given by Theorem 3.8 and results from Lachièze-Rey and Peccati [2017] giving Malliavin–Stein bounds for functionals of the binomial point process.

Remark. If $K = \mathbb{X}$ in Theorem 3.9, then

$$I_{\mathbb{X},t} = t, \quad t \geqslant 1.$$

Corollary 3.2 (Corollary 2.2 in Lachièze-Rey et al. [2019]).

(a) Let the conditions of Theorem 3.9 (a) hold and assume that there is a constant $C \in (0, \infty)$ such that $\sup_{\lambda \geqslant 1} I_{K,\lambda}/\operatorname{Var} H_\lambda \leqslant C$. Then there is a constant $\tilde{C} \in (0, \infty)$ only depending on C and the constants $\gamma, c_{stab}, \alpha_{stab}, C_p, c_K$ and α_K such that

$$d_K\left(\frac{H_\lambda - \mathbb{E}\, H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) \leqslant \frac{\tilde{C}}{\sqrt{\operatorname{Var} H_\lambda}}, \quad \lambda \geqslant 1.$$

(b) Let the conditions of Theorem 3.9 (b) hold and assume that there is a constant $C' \in (0, \infty)$ such that $\sup_{n \in \mathbb{N}} I_{K,n}/\operatorname{Var} H'_n \leqslant C'$. Then there is a constant $\tilde{C}' \in (0, \infty)$ only depending on C' and the constants $\gamma, c_{stab}, \alpha_{stab}, C_p, c_K$ and α_K such that

$$d_K\left(\frac{H'_n - \mathbb{E}\, H'_n}{\sqrt{\operatorname{Var} H'_n}}, Z\right) \leqslant \frac{\tilde{C}'}{\sqrt{\operatorname{Var} H'_n}}, \quad n \in \mathbb{N}.$$

Remark. Let $Y_1, \ldots, Y_n$ be i.i.d. random variables such that $\mathbb{E}\,|Y_1|^3 < \infty$. We have for $S_n = \sum_{i=1}^n Y_i$ the Berry–Esseen theorem saying that

$$d_K\left(\frac{S_n - \mathbb{E}\, S_n}{\sqrt{\operatorname{Var} S_n}}, Z\right) \leqslant \frac{C\mathbb{E}\,|Y_1 - \mathbb{E}\, Y_1|^3}{\operatorname{Var} Y_1} \frac{1}{\sqrt{\operatorname{Var} S_n}}, \; n \in \mathbb{N},$$

for some $C \in (0, \infty)$. By considering special choices of $Y_1, \ldots, Y_n$, it can be shown that the rate $1/\sqrt{\operatorname{Var} S_n}$ is optimal. One can deduce that the rates $1/\sqrt{\operatorname{Var} H_\lambda}$, resp. $1/\sqrt{\operatorname{Var} H'_n}$ occurring in Corollary 3.2 are as well presumably optimal.

Another simplification of Theorem 3.9 is applicable if we take $\mathbb{X} \subset \mathbb{R}^d$ yielding rates of convergence for functionals with variances being proportional to λ or n, resp. Given an unbounded set $I \subset (0, \infty)$, we say $(a_i)_{i \in I}$ is *proportional* to $(b_i)_{i \in I}$ if $a_i = O(b_i)$ and vice versa.

Theorem 3.10 (Theorem 2.3 in Lachièze-Rey et al. [2019]).
Let $\mathbb{X} \subset \mathbb{R}^d$ be full-dimensional, let $\mathbb{Q}$ have a bounded density with respect to the Lebesgue measure and suppose the conditions of Theorem 3.9 hold with $\gamma = d$. Moreover, let K be a full-dimensional subset of $\mathbb{X}$, whose boundary ∂K satisfies

$$\limsup_{r \to 0} \frac{|K_r|_d}{2r} < \infty,$$

where $K_r := \{y \in \mathbb{R}^d : d(y, \partial K) \leqslant r\}$ denotes the r-parallel set of K. If $\operatorname{Var} H_\lambda$ is proportional to λ, resp. $\operatorname{Var} H'_n$ is proportional to n, then there is a constant $c \in (0, \infty)$ such that for $\lambda \geqslant 1$, resp. $n \geqslant 9$,

$$d_K\left(\frac{H_\lambda - \mathbb{E}\, H_\lambda}{\sqrt{\operatorname{Var} H_\lambda}}, Z\right) \leqslant \frac{c}{\sqrt{\lambda}}, \quad \textit{resp.} \quad d_K\left(\frac{H'_n - \mathbb{E}\, H'_n}{\sqrt{\operatorname{Var} H'_n}}, Z\right) \leqslant \frac{c}{\sqrt{n}}.$$

Remark. The rates of convergence for Poisson input in Theorem 3.10 improve upon the rates given by results in Barbour and Xia [2006] and Penrose and Yukich [2005], which contain extra logarithmic factors.

Example 3.3 (k-th nearest neighbour graph).
We let η_λ be a homogeneous Poisson point process of intensity λ in a compact convex observation window $W \subset \mathbb{R}^d$ with interior points. For a fixed $k \in \mathbb{N}$, we consider the following functional of η_λ

$$H_\lambda^\alpha = \frac{1}{2} \sum_{(x,y)\in\eta^2_{\lambda,\neq}} \mathbf{1}_{\{x \text{ is a } k\text{-nearest neighbour of } y \text{ of vice versa}\}} \|x - y\|^\alpha.$$

By taking $\alpha = 0$, we obtain the number of edges, while for $\alpha = 1$, the total edge length. The first central limit theorem including rates of convergence for the k-th nearest neighbour graph was shown in Avram and Bertsimas [1993] with order $(\log \lambda)^{1+3/4}\lambda^{-1/4}$. Later, it was improved in Penrose and Yukich [2005] with rate $(\log \lambda)^{3d}\lambda^{-1/2}$. Finally, the logarithmic factor was removed in Last et al. [2016].

Theorem 3.11 (Theorem 7.1 in Last et al. [2016]).
There is a constant C_α only depending on k, W and α such that

$$d_K\left(\frac{H_\lambda^\alpha - \mathbb{E}\, H_\lambda^\alpha}{\sqrt{\operatorname{Var} H_\lambda^\alpha}}\right) \leqslant C_\alpha \lambda^{-1/2}, \quad \lambda \geqslant 1.$$

See Section 5.1 in Lachièze-Rey et al. [2019] for further generalizations of this result concerning the underlying space and the binomial input.

Conclusion

The asymptotic analysis based on the add-one cost method and the moment approach yield variance asymptotic, which is not addressed in the method based on Malliavin–Stein bounds. The latter method on the other hand, yields error bound providing useful information about the rate of convergence and without requiring higher-order moment calculations.

The advantage of the add-one cost method over the moment approach and the method based on the Malliavin–Stein bounds is that the add-one cost method does not require bounds on the tail of the radius of stabilization (i.e., on the range of the local effect of an inserted point). It requires only that this radius be almost surely finite. Therefore, it can be applicable to some examples such as those concerned with the minimal spanning tree. On the other hand, only the homogeneous point process input on the Euclidean space is considered.

3.4 Stabilization of weighted Poisson–Voronoi tessellations

A lot of attention was given to the stabilization of the random tessellations in the literature since they play an important role in computational geometry. Many algorithms for solving some geometric problems are based on them. For example, the Delaunay triangulation is important in finding the minimum spanning tree because it is a subgraph of the Delaunay triangulation. The property of being defined locally in the setting of Poisson–Voronoi tessellation and the corresponding Delaunay triangulation is well known since Avram and Bertsimas [1993], where the authors used dependency graphs technique. We shall be interested in developing the limit theory for unbiased and consistent estimators of statistics of a typical cell in a generalized weighted Voronoi tessellation using the stabilization method.

The estimators are constructed by observing the tessellation within a bounded window. Unbiased estimators are constructed by considering only those cells which lie within the bounded window. This technique, known as minus-sampling, has a long history going back to Miles [1974] as well as Horvitz and Thompson; see Baddeley [1999] for details.

The authors in Flimmel et al. [2020] used stabilization methods described in the previous sections to develop expectation and variance asymptotics, as well as central limit theorems, for unbiased and asymptotically consistent estimators of geometric statistics of a typical cell. Under mild conditions on the weights of the cells, they established variance asymptotics, weak consistency and the asymptotic normality of the estimators as the observation window tends to the whole space. The rest of this section consists of the results stated in the paper.

Assumptions

Let μ_m be a unit intensity stationary point process in $\mathbb{R}^d$ where each point of the process carry an independent mark in $\mathbb{M} := [0, a]$ for some fixed constant $a < \infty$. Let $\mathcal{B}(\mathbb{M})$ be the Borel σ-field on $\mathbb{M}$ and let $\mathbb{Q}_{\mathbb{M}}$ be the mark distribution (recall the definitions of Section 1.3). The elements of $\mathbb{R}^d \times \mathbb{M}$ will be denoted by $\hat{x} := (x, m_x)$.

For a weight function $\rho : \mathbb{R}^d \times (\mathbb{R}^d \times \mathbb{M}) \to \mathbb{R}$ we consider the notion of the *typical cell* of the weighted Voronoi tessellation defined by the weight ρ (see Example 1.22). By the typical cell $K^\rho_{\mathbf{o}} := K^\rho_{\mathbf{o}}(\mu_m)$ we understand the cell generated by the typical point of μ_m. This can be formally introduced by considering the Palm probability $\mathbb{P}^0$ (recall Definition 1.14) which corresponds to $\mathbb{P}$ conditional on the event that μ_m has a point at the origin. Let Q^ρ denote the distribution of the typical cell. The expectation with respect to $\mathbb{P}^0$ is denoted by $\mathbb{E}^0$.

Remark. In the case of Laguerre or Johnson–Mehl tessellations the typical cell $K^\rho_{\mathbf{o}}$ could satisfy $K^\rho_{\mathbf{o}} = \varnothing$. This is different from the definition of the typical cell described in e.g. Section 10.4 in Schneider and Weil [2008], where the typical cell is meant to be the typical non-empty cell. For a Voronoi tessellation both approaches coincide. For weighted Voronoi tessellations in general, $K^\rho_{\mathbf{o}}$ is distributed as a mixture of the typical non-empty cell and the empty cell with mixture weights $1-p_\varnothing$ and $p_\varnothing$, where $p_\varnothing$ is the probability that the cell generated by the typical point is empty.

Notation. Denote by $\mathcal{F}^d$ the space of all closed subsets of $\mathbb{R}^d$ equipped with the Borel σ-field $\mathcal{B}(\mathcal{F}^d)$ generated by the open sets from the Fell topology, see Definition 2.1.1 in Schneider and Weil [2008].

Moreover, let $\mathbb{P}^{\hat{x}}$, respectively $\mathbb{P}^{\hat{x},\hat{y}}$, denote the Palm probability measures conditioned on $\mathbb{P}$ having an additional marked point $\hat{x}$, respectively, two additional marked points $\hat{x}$ and $\hat{y}$. In particular, $\mathbb{P}^0(\cdot) = \int_{\mathbb{M}} \mathbb{P}^{(\mathbf{o},m)}(\cdot)\, \mathbb{Q}_{\mathbb{M}}(\mathrm{d}m)$. By $\mathbb{E}^{\hat{x}}$, respectively $\mathbb{E}^{\hat{x},\hat{y}}$, we denote expectation with respect to $\mathbb{P}^{\hat{x}}$, respectively, $\mathbb{P}^{\hat{x},\hat{y}}$.

Now, we introduce for $z \in \mathbb{R}^d$ and $\hat{x} \in \mu_m$, the shifted cell

$$C^\rho_z(\hat{x}, \mu_m) := C^\rho(\hat{x}, \mu_m) + (z - x).$$

Specially, $C^\rho(\hat{x}, \mu_m) = x + C^\rho_{\mathbf{o}}(\hat{x}, \mu_m)$. Note that $K^\rho_{\mathbf{o}} = C^\rho_{\mathbf{o}}((\mathbf{o}, M_{\mathbf{o}}), \mu_m)$ $\mathbb{P}^0$-almost surely, where $M_{\mathbf{o}}$ is the typical mark distributed according to $\mathbb{Q}_{\mathbb{M}}$.

Let $h : (\mathcal{F}^d, \mathcal{B}(\mathcal{F}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ describe a geometric characteristic of elements of $\mathcal{F}^d$ (e.g. diameter, volume) such that $h(\varnothing) = 0$ and it is invariant with respect to shifts, specially for all $x \in \mathbb{R}^d$ and $m_x \in \mathbb{M}$

$$h(C^\rho((x, m_x), \mu_m)) = h(x + C^\rho_{\mathbf{o}}((x, m_x), \mu_m)) = h(C^\rho_{\mathbf{o}}((x, m_x), \mu_m)).$$

We have two goals:

(i) use minus-sampling to construct unbiased estimators of

$$\mathbb{E}^0 h(K^\rho_{\mathbf{o}}) = \int h(K)\, Q^\rho(\mathrm{d}K)$$

(ii) establish variance asymptotics and asymptotic normality of such estimators. As a by-product, we also establish the limit theory for geometric statistics of Laguerre and Johnson–Mehl tessellations, adding to the results of Penrose [2007b] and Penrose and Yukich [2001] which are confined to Voronoi tessellations.

Put $W_\lambda := [-\frac{\lambda^{1/d}}{2}, \frac{\lambda^{1/d}}{2}]^d$ and $\hat{W}_\lambda := W_\lambda \times \mathbb{M}$, $\lambda > 0$. Given h and a tessellation defined by the weight ρ, we define for all $\lambda > 0$

$$H^\rho_\lambda(\mu_m \cap \hat{W}_\lambda) := \sum_{\hat{x} \in \mu_m \cap \hat{W}_\lambda} \frac{h(C^\rho(\hat{x}, \mu_m))}{|W_\lambda \ominus C^\rho(\hat{x}, \mu_m)|_d} \mathbf{1}\{C^\rho(\hat{x}, \mu_m) \subset W_\lambda\}.$$

Here, for sets A and B, $A \ominus B := \{x \in \mathbb{R}^d : B + x \subset A\}$ denotes the *erosion* of A by B. The statistic $H^\rho_\lambda(\mu_m \cap \hat{W}_\lambda)$ disregards cells contained in the window W_λ that are generated by the points outside W_λ. Such cells do not exist in the Voronoi case but they could appear for weighted cells. Therefore, we also consider

$$H^\rho_\lambda(\mu_m) := \sum_{\hat{x} \in \mu_m} \frac{h(C^\rho(\hat{x}, \mu_m))}{|W_\lambda \ominus C^\rho(\hat{x}, \mu_m)|_d} \mathbf{1}\{C^\rho(\hat{x}, \mu_m) \subset W_\lambda\}.$$

Unfortunately, controlling the moments of $H^\rho_\lambda(\mu_m \cap \hat{W}_\lambda)$ is problematic since $|W_\lambda \ominus C^\rho(\hat{x}, \mu_m)|_d$ may become arbitrarily small. It will therefore be convenient to consider versions of $H^\rho_\lambda(\mu_m \cap \hat{W}_\lambda)$ and $H^\rho_\lambda(\mu_m)$ given by

$$\hat{H}^\rho_\lambda(\mu_m \cap \hat{W}_\lambda) := \sum_{\hat{x} \in \mu_m \cap \hat{W}_\lambda} \frac{h(C^\rho(\hat{x}, \mu_m))\, \mathbf{1}\{C^\rho(\hat{x}, \mu_m) \subset W_\lambda\}}{|W_\lambda \ominus C^\rho(\hat{x}, \mu_m)|_d} \mathbf{1}\left\{|W_\lambda \ominus C^\rho(\hat{x}, \mu_m)|_d \geqslant \frac{\lambda}{2}\right\}$$

and

$$\hat{H}^{\rho}_{\lambda}(\mu_m) := \sum_{\hat{x}\in\mu_m} \frac{h(C^{\rho}(\hat{x},\mu_m))\,\mathbf{1}\{C^{\rho}(\hat{x},\mu_m)\subset W_\lambda\}}{|W_\lambda \ominus C^{\rho}(\hat{x},\mu_m)|_d}\mathbf{1}\left\{|W_\lambda \ominus C^{\rho}(\hat{x},\mu_m)|_d \geqslant \frac{\lambda}{2}\right\}.$$

Note that $H^{\rho}_{\lambda}(\mathbb{P}\cap\hat{W}_\lambda)$, $\hat{H}^{\rho}_{\lambda}(\mathbb{P})$ and $\hat{H}^{\rho}_{\lambda}(\mathbb{P}\cap\hat{W}_\lambda)$ are not unbiased. Under the assumptions of Theorem 3.12, one instead has

$$\mathbb{E}\,H^{\rho}_{\lambda}(\mu_m\cap\hat{W}_\lambda) = \mathbb{E}^0\left(h(K^{\rho}_{\mathbf{o}})\frac{|W_\lambda\cap(W_\lambda\ominus K^{\rho}_{\mathbf{o}})|_d}{|W_\lambda\ominus K^{\rho}_{\mathbf{o}}|_d}\right),$$

$$\mathbb{E}\,\hat{H}^{\rho}_{\lambda}(\mu_m\cap\hat{W}_\lambda) = \mathbb{E}^0\left(h(K^{\rho}_{\mathbf{o}})\frac{|W_\lambda\cap(W_\lambda\ominus K^{\rho}_{\mathbf{o}})|_d}{|W_\lambda\ominus K^{\rho}_{\mathbf{o}}|_d}\mathbf{1}\{|W_\lambda\ominus K^{\rho}_{\mathbf{o}}|_d \geqslant \frac{\lambda}{2}\}\right),$$

and

$$\mathbb{E}\,\hat{H}^{\rho}_{\lambda}(\mu_m) = \mathbb{E}^0\left(h(K^{\rho}_{\mathbf{o}})\mathbf{1}\{|W_\lambda\ominus K^{\rho}_{\mathbf{o}}|_d \geqslant \frac{\lambda}{2}\}\right).$$

The general form of the bias is given by Theorem 2.1 of Baddeley [1999].

We need some additional terminology. For every weight ρ and geometric statistic h we define the score $\xi^{\rho}:\hat{\mathbb{R}^d}\times\mathbf{N}\to\mathbb{R}$ by

$$\xi^{\rho}(\hat{x},\mathcal{A}) := h(C^{\rho}(\hat{x},\mathcal{A}))\mathbf{1}\{C^{\rho}(\hat{x},\mathcal{A}) \text{ is bounded}\}, \quad \hat{x}\in\hat{\mathbb{R}^d},\ \mathcal{A}\in\mathbf{N}. \tag{3.9}$$

We use this representation to explicitly link our statistics with the stabilizing statistics in the literature Baryshnikov and Yukich [2005], Błaszczyszyn et al. [2019], Lachièze-Rey et al. [2019], Penrose [2007b,a], Penrose and Yukich [2001, 2003]. Translation invariance for h implies

$$\xi^{\rho}(\hat{x},\mathcal{A}) = \xi^{\rho}((x,m_x),\mathcal{A}) = \xi^{\rho}((\mathbf{o},m_x),\mathcal{A}-x),$$

for every $\hat{x}\in\hat{\mathbb{R}^d}$, $\hat{x}:=(x,m_x)$ and $\mathcal{A}\in\mathbf{N}$, where $\mathcal{A}-x := \{(a-x,m_a):(a,m_a)\in\mathcal{A}\}$. If $C^{\rho}(\hat{x},\mu_m)$ is empty we have $\xi^{\rho}(\hat{x},\mu_m) = h(\varnothing) = 0$. Write $C^{\rho}(\hat{x},\mu_m) := C^{\rho}(\hat{x},\mu_m\cup\{\hat{x}\})$ for $\hat{x}\notin\mu_m$ and so $\xi^{\rho}(\hat{x},\mu_m) := \xi^{\rho}(\hat{x},\mu_m\cup\{\hat{x}\})$ for $\hat{x}\notin\mu_m$.

Definition 3.18 (p-moment condition 2).
The score ξ^{ρ} is said to satisfy a *p-moment condition* 2, $p\in[1,\infty)$, if

$$\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}^d}} \mathbb{E}^{\hat{x},\hat{y}}|\xi^{\rho}(\hat{x},\mu_m)|^p < \infty. \tag{3.10}$$

In contrast to Definitions 3.13 and 3.16, here we assume the addition of just two marked points into the process μ_m.

Definition 3.19 (Diameters with exponentially decaying tails).
We say that the cells of the tessellation defined by ρ and generated by μ_m have *diameters with exponentially decaying tails* if there is a constant $c_{diam}\in(0,\infty)$ such that for all $\hat{x}:=(x,m_x)\in\mu_m$ there exists an almost surely finite random variable $D_{\hat{x}}$ such that $C^{\rho}(\hat{x},\mu_m)\subset B_{D_{\hat{x}}}(x)$ and

$$\mathbb{P}^{\hat{x}}(D_{\hat{x}}\geqslant t)\leqslant c_{diam}\exp\left(-\frac{1}{c_{diam}}t^d\right), \quad t\geqslant 0. \tag{3.11}$$

Next, we make use of the concept of (exponentially) stabilizing scores of marked point processes introduced in Section 3.3 (see Definitions 3.14 and 3.15). Translated to our situation, we say that ξ^ρ is *stabilizing* with respect to μ_m if for all $\hat{x} := (x, m_x) \in \mu_m$ there exists an almost surely finite random variable $R_{\hat{x}} := R_{\hat{x}}(\mu_m)$ (*radius of stabilization*), such that

$$\xi^\rho(\hat{x}, (\mu_m \cup \mathcal{A}) \cap \hat{B}_{R_{\hat{x}}}(x)) = \xi^\rho(\hat{x}, \mu_m \cup \mathcal{A}) \tag{3.12}$$

for all $\mathcal{A}$ with $\#(\mathcal{A}) \leqslant 7$. We say that ξ^ρ is *exponentially stabilizing* with respect to μ_m if there are constants $c_{stab}, \alpha \in (0, \infty)$ such that

$$\mathbb{P}^{\hat{x}}(R_{\hat{x}} \geqslant t) \leqslant c_{stab} \exp\left(-\frac{1}{c_{stab}} t^\alpha\right), \quad t \geqslant 0.$$

In other words, ξ^ρ is stabilizing with respect to μ_m if there is $R_{\hat{x}}$ such that the cell $C^\rho(\hat{x}, \mu_m)$ is not affected by changes in point configurations outside $\hat{B}_{R_{\hat{x}}}(x)$.

Main results

Our first main result has a proof which is short and illustrative. The result holds if W_λ is replaced by any observation window. It is a special case of a more general result given by Theorem 2.1 in Baddeley [1999] and formulated for stationary germ-grain models and general sampling rules.

Theorem 3.12 (Theorem 2.1 in Flimmel et al. [2020]).
Let μ_m be a stationary marked point process with unit intensity and mark distribution $\mathbb{Q}_\mathbb{M}$. Let $h : \mathcal{F}^d \to \mathbb{R}$ be translation-invariant as above. Then for all $\lambda > 0$, the statistic $H^\rho_\lambda(\mu_m)$ is an unbiased estimator of $\mathbb{E}^0 h(K^\rho_\mathbf{o})$.

Proof of Theorem 3.12. We have

$$\begin{aligned}
\mathbb{E}\, H^\rho_\lambda(\mu_m) &= \mathbb{E} \sum_{\hat{x} \in \mu_m} \frac{h(C^\rho(\hat{x}, \mu_m))}{|W_\lambda \ominus C^\rho(\hat{x}, \mu_m)|_d} \mathbf{1}\{C^\rho(\hat{x}, \mu_m) \subset W_\lambda\} \\
&= \mathbb{E} \sum_{\hat{x} \in \mu_m} \frac{h(C^\rho_\mathbf{o}(\hat{x}, \mu_m))}{|W_\lambda \ominus C^\rho_\mathbf{o}(\hat{x}, \mu_m)|_d} \mathbf{1}\{x + C^\rho_\mathbf{o}(\hat{x}, \mu_m) \subset W_\lambda\} \\
&= \int_{\mathbb{R}^d} \mathbb{E}^0 \left(\frac{h(K^\rho_\mathbf{o})}{|W_\lambda \ominus K^\rho_\mathbf{o}|_d} \mathbf{1}\{x + K^\rho_\mathbf{o} \subset W_\lambda\} \right) \mathrm{d}x \\
&= \mathbb{E}^0 \int_{\mathbb{R}^d} \left(\frac{h(K^\rho_\mathbf{o})}{|W_\lambda \ominus K^\rho_\mathbf{o}|_d} \mathbf{1}\{x \in W_\lambda \ominus K^\rho_\mathbf{o}\} \right) \mathrm{d}x \\
&= \mathbb{E}^0 h(K^\rho_\mathbf{o}),
\end{aligned}$$

where the second equality uses the translation invariance of h and translation invariance of erosions, the third uses the refined Campbell theorem for stationary marked point processes (a version of Theorem 1.8 for marked point processes, see e.g. Theorem 3.5.3 in Schneider and Weil [2008]), while the fourth uses Fubini's theorem. Hence, we have shown the unbiasedness of $H^\rho_\lambda(\mu_m)$. □

By η_s, $s \in (0, \infty)$, we denote a homogeneous marked Poisson point process with values in $\mathbb{R}^d \times \mathbb{M}$ and such that the unmarked process on $\mathbb{R}^d$ has intensity

s. We write η for η_1. Our next results establish the limit theory for the above estimators.

Theorem 3.13 (Theorem 2.2 in Flimmel et al. [2020]).
Let $M_{\mathbf{o}}$ be a random mark distributed according to $\mathbb{Q}_{\mathbb{M}}$.

(i) If ξ^ρ satisfies the p-moment condition 2 (3.10) for some $p \in (1, \infty)$ and if the cell $C^\rho((\mathbf{o}, M_{\mathbf{o}}), \eta)$ has a diameter with an exponentially decaying tail, then $H^\rho_\lambda(\eta \cap \hat{W}_\lambda)$, $\hat{H}^\rho_\lambda(\eta)$ and $\hat{H}^\rho_\lambda(\eta \cap \hat{W}_\lambda)$ are asymptotically unbiased estimators of $\mathbb{E}^0 h(K^\rho_{\mathbf{o}}(\eta))$.

(ii) Under the conditions of (i) and assuming that ξ^ρ stabilizes with respect to η as at (3.12), the statistics $H^\rho_\lambda(\eta)$, $H^\rho_\lambda(\eta \cap \hat{W}_\lambda)$, $\hat{H}^\rho_\lambda(\eta)$ and $\hat{H}^\rho_\lambda(\eta \cap \hat{W}_\lambda)$ are consistent estimators of $\mathbb{E}^0 h(K^\rho_{\mathbf{o}}(\eta))$.

Given the score ξ^ρ at (3.9), put

$$\sigma^2(\xi^\rho) := \mathbb{E}\,(\xi^\rho(\mathbf{o}_M, \eta))^2 + \int_{\mathbb{R}^d} [\mathbb{E}\,\xi^\rho(\mathbf{o}_M, \eta \cup \{x_M\})\xi^\rho(x_M, \eta \cup \{\mathbf{o}_M\}) - \mathbb{E}\,\xi^\rho(\mathbf{o}_M, \eta)\,\mathbb{E}\,\xi^\rho(x_M, \eta)]\,\mathrm{d}x, \tag{3.13}$$

where $\mathbf{o}_M := (\mathbf{o}, M_{\mathbf{o}})$, $x_M := (x, M_x)$, and $M_{\mathbf{o}}$ and M_x are independent random marks distributed according to $\mathbb{Q}_{\mathbb{M}}$. Note that $\mathbb{E}^0 h(K^\rho_{\mathbf{o}}(\eta)) = \mathbb{E}\,h(C^\rho(\mathbf{o}_M, \eta)) = \mathbb{E}\,\xi^\rho(\mathbf{o}_M, \eta)$ by the Slivnyak–Mecke theorem (Theorem 1.9). Here we use that, given the Poisson process η, the Palm distribution corresponds to the usual distribution with a point inserted at the origin.

Theorem 3.14 (Theorem 2.3 in Flimmel et al. [2020]).
Let h be translation-invariant and assume that ξ^ρ is exponentially stabilizing with respect to η.

(i) If ξ^ρ satisfies the p-moment condition 2 (3.10) for some $p \in (2, \infty)$, then

$$\lim_{\lambda\to\infty} \lambda \mathrm{Var}\, \hat{H}^\rho_\lambda(\eta \cap \hat{W}_\lambda) = \lim_{\lambda\to\infty} \lambda \mathrm{Var}\, \hat{H}^\rho_\lambda(\eta) = \sigma^2(\xi^\rho) \in [0, \infty). \tag{3.14}$$

(ii) If $\sigma^2(\xi^\rho) \in (0, \infty)$ and if the p-moment condition 2 (3.10) holds for some $p \in (4, \infty)$, then

$$\sqrt{\lambda}\left(H^\rho_\lambda(\eta \cap \hat{W}_\lambda) - \mathbb{E}\,H^\rho_\lambda(\eta \cap \hat{W}_\lambda)\right) \xrightarrow[\lambda\to\infty]{\mathcal{D}} N(0, \sigma^2(\xi^\rho))$$

and

$$\sqrt{\lambda}\left(H^\rho_\lambda(\eta) - \mathbb{E}^0 h(K^\rho_{\mathbf{o}}(\eta))\right) \xrightarrow[\lambda\to\infty]{\mathcal{D}} N(0, \sigma^2(\xi^\rho)),$$

where $N(0, \sigma^2(\xi^\rho))$ denotes a mean zero Gaussian random variable with variance $\sigma^2(\xi^\rho)$.

Remarks. (i) The assumption $\sigma^2(\xi^\rho) \in (0, \infty)$ is often satisfied by scores of interest, as seen in the upcoming applications. According to Theorem 2.1 in Penrose and Yukich [2001], where it has been shown that whenever we have

$$\frac{\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda}(\xi^\rho(\hat{x}, \eta) - \mathbb{E}\,\xi^\rho(\hat{x}, \eta))}{\sqrt{\mathrm{Var}\,\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda}\xi^\rho(\hat{x}, \eta)}} \xrightarrow{\mathcal{D}} N(0, \sigma^2(\xi^\rho)),$$

then necessarily $\sigma^2(\xi^\rho) \in (0, \infty)$ provided (a) there is a random variable $S < \infty$ and a random variable $\Delta^\rho(\infty)$ such that for all finite $\mathcal{A} \subset \hat{B}_S(\mathbf{o})^c$ we have

$$\Delta^\rho(\infty) = \sum_{\hat{x} \in (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}} \xi^\rho(\hat{x}, (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}) - \sum_{\hat{x} \in (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A}} \xi^\rho(\hat{x}, (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A}),$$

and (b) $\Delta^\rho(\infty)$ is non-degenerate, that is to say it is not almost surely constant. We will use this fact in showing positivity of $\sigma^2(\xi^\rho)$ in the applications which follow.
(ii) Theorems 3.13 and 3.14 hold for translation-invariant statistics h of Poisson–Voronoi cells regardless of the mark distribution because ξ^{ρ_1} stabilizes exponentially fast and diameters of Voronoi cells have exponentially decaying tails as shown in Penrose [2007a], Penrose and Yukich [2001]. In Proposition 3.1 we establish that the cells of the Laguerre and the Johnson–Mehl tessellations also have diameters with exponentially decaying tails and that $\xi^{\rho_i}, i = 2, 3$ are exponentially stabilizing with respect to η.

Applications. We provide some applications of our main results. Our first result gives the limit theory for an unbiased estimator of the distribution function of the volume of a typical cell in a weighted Poisson–Voronoi tessellation.

Theorem 3.15 (Theorem 2.4 in Flimmel et al. [2020]).
(i) For all $i = 1, 2, 3$ and $t \in (0, \infty)$ the statistic

$$\sum_{\hat{x} \in \eta} \frac{\mathbf{1}\{|C^{\rho_i}(\hat{x}, \eta)|_d \leqslant t\}}{|W_\lambda \ominus C^{\rho_i}(\hat{x}, \eta)|_d} \mathbf{1}\{C^{\rho_i}(\hat{x}, \eta) \subset W_\lambda\}$$

is an unbiased estimator of $\mathbb{P}^0(|K_{\mathbf{o}}^{\rho_i}(\eta)|_d \leqslant t)$.
(ii) For all $i = 1, 2, 3$ and $t \in (0, \infty)$ we have that

$$\sqrt{\lambda} \left(\sum_{\hat{x} \in \eta} \frac{\mathbf{1}\{|C^{\rho_i}(\hat{x}, \eta)|_d \leqslant t\}}{|W_\lambda \ominus C^{\rho_i}(\hat{x}, \eta)|_d} \mathbf{1}\{C^{\rho_i}(\hat{x}, \eta) \subset W_\lambda\} - \mathbb{P}^0(|K_{\mathbf{o}}^{\rho_i}(\eta)|_d \leqslant t) \right) \qquad (3.15)$$

tends to $N(0, \sigma^2(\varphi^{\rho_i}))$ in distribution as $\lambda \to \infty$, where

$$\varphi^{\rho_i}(\hat{x}, \eta) := \mathbf{1}\{|C^{\rho_i}(\hat{x}, \eta)|_d \leqslant t\}$$

and where $\sigma^2(\varphi^{\rho_i}) \in (0, \infty)$ is given by (3.13).

Our next result gives the limit theory for an unbiased estimator of the $(d-1)$-dimensional Hausdorff measure $\mathcal{H}^{d-1}$ of the boundary of a typical cell in a weighted Poisson–Voronoi tessellation.

Theorem 3.16 (Theorem 2.5 in Flimmel et al. [2020]).
(i) For all $i = 1, 2, 3$ we have that

$$\sum_{\hat{x} \in \eta} \frac{\mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x}, \eta))}{|W_\lambda \ominus C^{\rho_i}(\hat{x}, \eta)|_d} \mathbf{1}\{C^{\rho_i}(\hat{x}, \eta) \subset W_\lambda\}$$

is an unbiased estimator of $\mathbb{E}^0\mathcal{H}^{d-1}(\partial K_{\mathbf{o}}^{\rho_i}(\eta))$.
(ii) For all $i = 1, 2, 3$ *we have that*

$$\sqrt{\lambda}\left(\sum_{\hat{x}\in\eta}\frac{\mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta))}{|W_\lambda \ominus C^{\rho_i}(\hat{x},\eta)|_d}\mathbf{1}\{C^{\rho_i}(\hat{x},\eta)\subset W_\lambda\} - \mathbb{E}^0\mathcal{H}^{d-1}(\partial K_{\mathbf{o}}^{\rho_i}(\eta))\right)$$

tends to $N(0, \sigma^2(\xi^{\rho_i}))$ *in distribution as* $\lambda \to \infty$*, where*

$$\xi^{\rho_i}(\hat{x},\eta) := \mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta))\mathbf{1}\{C^{\rho_i}(\hat{x},\eta) \text{ is bounded}\}$$

and where $\sigma^2(\xi^{\rho_i}) \in (0,\infty)$ *is given by* (3.13).

There are naturally other applications of the general theorems. By choosing h appropriately, one could for example use the general results to deduce the limit theory for an unbiased estimator of the distribution function of either the surface area, inradius, or circumradius of a typical cell in a weighted Poisson–Voronoi tessellation.

Stabilization of statistics of weighted Poisson–Voronoi tessellations

We establish that

(i) the cells in the Voronoi, Laguerre and Johnson–Mehl tessellations generated by Poisson input have diameters with exponentially decaying tails (see Definition 3.19) and

(ii) the scores ξ^{ρ_i}, $i = 1, 2, 3$, as defined at (3.9) are exponentially stabilizing (see Definition 3.15). These two conditions arise in the statements of Theorems 3.13 and 3.14.

Conditions (i) and (ii) have already been established in the case of the Poisson–Voronoi tessellation (ρ_1) in Penrose [2007a] and Penrose and Yukich [2001]. The Voronoi cell is a special example of both the Laguerre and the Johnson–Mehl cell when putting $\mathbb{M} = \{0\}$ (or any constant). Thus it will be enough to show that these two conditions hold for the Laguerre (ρ_2) and the Johnson–Mehl (ρ_3) tessellations.

By definition we have

$$C^\rho(\hat{x},\mu_m) = \bigcap_{\hat{z}\in\mu_m\setminus\{\hat{x}\}} \mathbb{H}^\rho_{\hat{z}}(\hat{x}),$$

where $\mathbb{H}^\rho_{\hat{z}}(\hat{x}) := \{y \in \mathbb{R}^d : \rho(y,\hat{x}) \leqslant \rho(y,\hat{z})\}$. Note that $\mathbb{H}^\rho_\cdot(\cdot)$ is a closed half-space in the context of the Voronoi and Laguerre tessellations, whereas it has a hyperbolic boundary for the Johnson–Mehl tessellation. Tessellations generated by μ_m are stationary and are examples of stationary particle processes (see Section 1.4).

Proposition 3.1 (Proposition 3.1 in Flimmel et al. [2020]). *The cells of the tessellation defined by* $\rho_i, i = 1, 2, 3$, *and generated by Poisson input* η *have diameters with exponentially decaying tails as at* (3.11).

Proof. We need to prove (3.11) for all $\hat{x} \in \eta$. Without loss of generality, we may assume that $\hat{x}$ is the origin $\hat{\mathbf{o}} := (\mathbf{o}, m_{\mathbf{o}})$ and we denote $D := D_{\hat{\mathbf{o}}}$.

Let $\mathcal{K}_j$, $j = 1, \dots, J$, be a collection of convex cones in $\mathbb{R}^d$ with $\cup_{j=1}^J \mathcal{K}_j = \mathbb{R}^d$ and $\langle x, y\rangle \geqslant 3\|x\|\|y\|/4$ for any x and y from the same cone $\mathcal{K}_j$. Each cone has an apex at the origin $\mathbf{o}$. Denote $\hat{\mathcal{K}}_j := \mathcal{K}_j \times \mathbb{M}$. We take $(x_j, m_j) \in \eta \cap \hat{\mathcal{K}}_j \cap \hat{B}_{2\mu}(\mathbf{o})^c$ so that x_j is closer to $\mathbf{o}$ than any other point from $\eta \cap \hat{\mathcal{K}}_j \cap \hat{B}_{2\mu}(\mathbf{o})^c$. This condition means that the balls $B_{m_{\mathbf{o}}}(\mathbf{o})$ and $B_{m_j}(x_j)$ do not overlap. Then

$$C^{\rho_i}(\hat{\mathbf{o}}, \eta) \subset \bigcap_{j=1}^J \mathbb{H}^{\rho_i}_{(x_j, m_j)}(\hat{\mathbf{o}}), \quad i = 1, 2, 3.$$

Thus, it is sufficient to find D such that for all $i = 1, 2, 3$ and $j = 1, \dots, J$ we have $\mathbb{H}^{\rho_i}_{(x_j, m_j)}(\hat{\mathbf{o}}) \cap \mathcal{K}_j \subset B_D(\mathbf{o})$, to obtain that $C^{\rho_i}(\hat{\mathbf{o}}, \eta) \subset B_D(\mathbf{o})$. Consider $y \in \mathbb{H}^{\rho_i}_{(x_j, m_j)}(\hat{\mathbf{o}}) \cap \mathcal{K}_j$. Then $\rho_i(y, \hat{\mathbf{o}}) \leqslant \rho_i(y, (x_j, m_j))$ and $\langle y, x_j\rangle \geqslant 3\|x_j\|\|y\|/4$. For the Laguerre cell (i.e. the case $i = 2$), the first condition necessarily means that $\|y\|^2 - m_{\mathbf{o}}^2 \leqslant \|y - x_j\|^2 - m_j^2 = \|y\|^2 + \|x_j\|^2 - 2\langle y, x_j\rangle - m_j^2$. Thus,

$$2\langle y, x_j\rangle \leqslant \|x_j\|^2 + m_{\mathbf{o}}^2 - m_j^2 \leqslant \|x_j\|^2 + \mu^2 < \frac{3}{2}\|x_j\|^2$$

and so $\|y\| < \|x_j\|$. For the Johnson–Mehl cell ($i = 3$) we have

$$\|y - x_j\| \geqslant \|y\| - m_{\mathbf{o}} + m_j \geqslant \|y\| - \mu,$$

which for $\|y\| > \mu$ gives

$$2\langle y, x_j\rangle \leqslant 2\mu\|y\| - \mu^2 + \|x_j\|^2.$$

Hence, using the assumptions $\langle x_j, y\rangle \geqslant 3\|x_j\|\|y\|/4$ and $\|x_j\| > 2\mu$,

$$\|y\| \leqslant \frac{2(\|x_j\|^2 - \mu^2)}{3\|x_j\| - 4\mu} < \frac{2\|x_j\|^2}{\|x_j\|} = 2\|x_j\|.$$

Consequently, for either the Laguerre or Johnson–Mehl cells, we can take

$$D = 2 \max_{j=1,\dots,J} \|x_j\|. \tag{3.16}$$

Then, for $t \in (4\mu, \infty)$ we have

$$\begin{aligned}\mathbb{P}^{\hat{\mathbf{o}}}(D \geqslant t) &\leqslant \sum_{j=1}^J \mathbb{P}(2\|x_j\| \geqslant t) = \sum_{j=1}^J \mathbb{P}(\eta \cap (\hat{B}_{t/2}(\mathbf{o}) \backslash \hat{B}_{2\mu}(\mathbf{o})) \cap \hat{\mathcal{K}}_j = \varnothing) \\ &= \sum_{j=1}^J \exp(-|(B_{t/2}(\mathbf{o}) \backslash B_{2\mu}(\mathbf{o})) \cap \mathcal{K}_j|_d) \leqslant c_{diam} \exp\left(-\frac{1}{c_{diam}} t^d\right)\end{aligned}$$

for some $c_{diam} := c_{diam}(d, \mu) \in (0, \infty)$ depending on d and μ. This shows Proposition 3.1 for $i = 2, 3$ and hence for $i = 1$ as well.

□

Proposition 3.2 (Proposition 3.2 in Flimmel et al. [2020]).
For all $i = 1, 2, 3$ the score ξ^{ρ_i} defined at (3.9) is exponentially stabilizing with respect to η.

Proof. We will prove (3.12) when $\hat{x}$ is the origin and we denote $R := R_{\hat{\mathbf{o}}}$. For simplicity of exposition, we prove (3.12) when $\mathcal{A}$ is the empty set, as the arguments do not change otherwise. In other words if $\mathcal{A}$ is not empty, then the resulting radius of stabilization will not be larger, as seen by the following arguments. By (3.9), it is enough to show that there is an almost surely finite random variable R such that

$$C^{\rho_i}(\hat{\mathbf{o}}, \eta \cap \hat{B}_R(\mathbf{o})) = C^{\rho_i}(\hat{\mathbf{o}}, (\eta \cap \hat{B}_R(\mathbf{o})) \cup \{(z, m_z)\}),$$

almost surely whenever $\|z\| \in (R, \infty)$. To see this we put $R := 2D + \mu$, where D is at (3.16). Given $\hat{z} := (z, m_z)$, with $\|z\| \in (R, \infty)$, we assert that

$$B_D(\mathbf{o}) \subset \mathbb{H}^{\rho_i}_{\hat{z}}(\hat{\mathbf{o}}).$$

To prove this, we take any point $y \in B_D(\mathbf{o})$ and show that

$$\rho_i(y, \hat{\mathbf{o}}) \leqslant \rho_i(y, \hat{z}), \quad i = 1, 2, 3. \tag{3.17}$$

Note that $y \in B_D(\mathbf{o})$ implies $\|y - z\| \in (D + \mu, \infty)$. The proof of (3.17) is shown for the Laguerre and Johnson–Mehl cases individually. First, assume that $C^{\rho_2}(\hat{\mathbf{o}}, \eta)$ is the cell in the Laguerre tessellation. Then

$$\rho_2(y, \hat{\mathbf{o}}) = \|y\|^2 - m_{\mathbf{o}}^2 \leqslant D^2 < (D+\mu)^2 - \mu^2 < \|y-z\|^2 - \mu^2 \leqslant \|y-z\|^2 - m_z^2 = \rho_2(y, \hat{z}),$$

showing that $y \in \mathbb{H}^{\rho_2}_{\hat{z}}(\hat{\mathbf{o}})$. For the Johnson–Mehl case,

$$\rho_3(y, \hat{\mathbf{o}}) = \|y\| - m_{\mathbf{o}} \leqslant D = (D + \mu) - \mu < \|y - z\| - \mu \leqslant \|y - z\| - m_z = \rho_3(y, \hat{z}),$$

thus again $y \in \mathbb{H}^{\rho_3}_{\hat{z}}(\hat{\mathbf{o}})$, which shows our assertion.

The radius D at (3.16) has a tail decaying exponentially fast, showing that R also has the same property. Consequently, for all $i = 1, 2, 3$, the score ξ^{ρ_i} is exponentially stabilizing with respect to η. □

Remark. (i) The assertion $C^{\rho_i}(\hat{\mathbf{o}}, \mu_m) \subset B_D(\mathbf{o})$ holds for a larger class of marked point processes. We only need that the unmarked point process has at least one point in each cone $\mathcal{K}_j \cap B_{2\mu}(\mathbf{o})^c$, $j = 1, \dots, J$, with Palm probability $\mathbb{P}^0$ equal to 1. Consequently, scores $\xi^{\rho_i}, i = 1, 2, 3$, are stabilizing with respect to such marked point processes.

(ii) Proposition 3.2 implies that the limit theory developed in McGivney and Yukich [1999], Penrose and Yukich [2001, 2003] for the total edge length and other stabilizing functionals of the Poisson–Voronoi tessellation extends to Poisson tessellation models with weighted Voronoi cells. Thus Proposition 3.2 provides expectation and variance asymptotics, as well as normal convergence, for such functionals of the Poisson tessellation.

(iii) Aside from weighted Poisson–Voronoi tessellations, Propositions 3.1 and 3.2 hold also for the Poisson–Delaunay triangulation. On the other hand, Proposition 3.1 holds for Poisson-line tessellation, but Proposition 3.2 does not.

Proofs of the main results

Preliminary lemmas. In this section, we omit in the notation the dependence on the weight ρ that defines the tessellation. For simplicity, we write

$$H_\lambda(\eta \cap \hat{W}_\lambda) := H^\rho_\lambda(\eta \cap \hat{W}_\lambda), \quad H_\lambda(\eta) := H^\rho_\lambda(\eta),$$

as well as

$$\hat{H}_\lambda(\eta \cap \hat{W}_\lambda) := \hat{H}^\rho_\lambda(\eta \cap \hat{W}_\lambda), \quad \hat{H}_\lambda(\eta) := \hat{H}^\rho_\lambda(\eta).$$

Let us start with some useful first order results.

Lemma 3.3 (Lemma 4.1 in Flimmel et al. [2020]).
Under the assumptions of Theorem 3.13(ii), we have

$$\lim_{\lambda\to\infty} \lambda\, \mathbb{E}\left| H_\lambda(\eta \cap \hat{W}_\lambda) - \hat{H}_\lambda(\eta \cap \hat{W}_\lambda)\right| = 0.$$

Proof. We denote by $\hat{\mathbb{Q}}$ the product of the Lebesgue measure on $\mathbb{R}^d$ and $\mathbb{Q}_{\mathbb{M}}$. By the refined Campbell theorem and stationarity,

$$\begin{aligned}
&\mathbb{E}\left| H_\lambda(\eta \cap \hat{W}_\lambda) - \hat{H}_\lambda(\eta \cap \hat{W}_\lambda)\right| \\
&\leqslant \mathbb{E} \sum_{\hat{x}\in\eta\cap\hat{W}_\lambda} \frac{|h(C(\hat{x},\eta))|}{|W_\lambda \ominus C(\hat{x},\eta)|_d} \mathbf{1}\{C(\hat{x},\eta)\subset W_\lambda\}\, \mathbf{1}\{|W_\lambda \ominus C(\hat{x},\eta)|_d < \frac{\lambda}{2}\} \\
&= \int_{\hat{W}_\lambda} \mathbb{E}^{\hat{x}} \frac{|h(C(\hat{x},\eta))|}{|W_\lambda \ominus C(\hat{x},\eta)|_d} \mathbf{1}\{C(\hat{x},\eta)\subset W_\lambda\} \\
&\qquad\qquad \cdot \mathbf{1}\{|W_\lambda \ominus C(\hat{x},\eta)|_d < \frac{\lambda}{2}\}\hat{\mathbb{Q}}(\mathrm{d}\hat{x}) \\
&= \int_{W_\lambda}\int_{\mathbb{M}} \mathbb{E}^{(\mathbf{o},m)}\Bigg(\frac{|h(C((\mathbf{o},m),\eta))|}{|W_\lambda \ominus C((\mathbf{o},m),\eta)|_d} \mathbf{1}\{x\in W_\lambda \ominus C((\mathbf{o},m),\eta)\} \\
&\qquad\qquad \cdot \mathbf{1}\{|W_\lambda \ominus C((\mathbf{o},m),\eta)|_d < \frac{\lambda}{2}\}\Bigg)\mathbb{Q}_{\mathbb{M}}(\mathrm{d}m)\,\mathrm{d}x.
\end{aligned}$$

Changing the order of integration we get

$$\begin{aligned}
&\mathbb{E}\left| H_\lambda(\eta \cap \hat{W}_\lambda) - \hat{H}_\lambda(\eta \cap \hat{W}_\lambda)\right| \\
&\leqslant \int_{\mathbb{M}} \mathbb{E}^{\mathbf{o}_m}\Bigg(|h(C(\mathbf{o}_m,\eta))| \mathbf{1}\{|W_\lambda \ominus C(\mathbf{o}_m,\eta)|_d < \frac{\lambda}{2}\} \\
&\qquad\qquad \cdot \int_{W_\lambda} \frac{\mathbf{1}\{x\in W_\lambda \ominus C(\mathbf{o}_m,\eta)\}}{|W_\lambda \ominus C(\mathbf{o}_m,\eta)|_d}\,\mathrm{d}x\Bigg)\mathbb{Q}_{\mathbb{M}}(\mathrm{d}m),
\end{aligned} \tag{3.18}$$

where $\mathbf{o}_m := (\mathbf{o}, m)$. The inner integral over W_λ is bounded by one, showing that for all $p \in (1, \infty)$ we have

$$\mathbb{E}\left|H_\lambda(\eta \cap \hat{W}_\lambda) - \hat{H}_\lambda(\eta \cap \hat{W}_\lambda)\right|$$
$$\leqslant \int_{\mathbb{M}} \mathbb{E}^{\mathbf{o}_m}\left(|h(C(\mathbf{o}_m, \eta))|\, \mathbf{1}\{|W_\lambda \ominus C(\mathbf{o}_m, \eta)|_d < \frac{\lambda}{2}\}\right) \mathbb{Q}_{\mathbb{M}}(\mathrm{d}m)$$
$$\leqslant \int_{\mathbb{M}} (\mathbb{E}^{\mathbf{o}_m}|h(C(\mathbf{o}_m, \eta))|^p)^{\frac{1}{p}}\, \mathbb{P}^{\mathbf{o}_m}\left(|W_\lambda \ominus C(\mathbf{o}_m, \eta)|_d < \frac{\lambda}{2}\right)^{\frac{p-1}{p}} \mathbb{Q}_{\mathbb{M}}(\mathrm{d}m).$$

The random variable D at (3.16) satisfies $C(\hat{\mathbf{o}}, \eta) \subset B_D(\mathbf{o})$ almost surely. Thus,

$$\mathbb{P}^{\hat{\mathbf{o}}}\left(|W_\lambda \ominus C(\hat{\mathbf{o}}, \eta)|_d < \frac{\lambda}{2}\right) \leqslant \mathbb{P}^{\hat{\mathbf{o}}}\left(|W_\lambda \ominus B_D(\mathbf{o})|_d < \frac{\lambda}{2}\right).$$

The volume of the erosion on the right-hand side equals $(\lambda^{1/d} - 2D)_+^d$. By conditioning on $Y := \mathbf{1}\{\lambda^{1/d} \geqslant 2D\}$, we obtain

$$\begin{aligned}\mathbb{P}^{\hat{\mathbf{o}}}\left((\lambda^{1/d} - 2D)_+^d < \frac{\lambda}{2}\right) &= \mathbb{P}^{\hat{\mathbf{o}}}\left((\lambda^{1/d} - 2D)_+^d < \frac{\lambda}{2} \mid Y = 1\right) \mathbb{P}^{\hat{\mathbf{o}}}(Y = 1) \\ &\quad + \mathbb{P}^{\hat{\mathbf{o}}}\left((\lambda^{1/d} - 2D)_+^d < \frac{\lambda}{2} \mid Y = 0\right) \mathbb{P}^{\hat{\mathbf{o}}}(Y = 0) \\ &\leqslant \mathbb{P}^{\hat{\mathbf{o}}}\left((\lambda^{1/d} - 2D)^d < \frac{\lambda}{2}\right) + \mathbb{P}^{\hat{\mathbf{o}}}(\lambda^{1/d} < 2D) \\ &\leqslant 2\mathbb{P}^{\hat{\mathbf{o}}}(D > e(\lambda)),\end{aligned}$$

where $e(\lambda) := (\lambda^{1/d} - (\lambda/2)^{1/d})/2$. Finally, recalling that D has exponentially decaying tails as at (3.11), we obtain

$$\mathbb{P}^{\hat{\mathbf{o}}}\left(|W_\lambda \ominus C(\hat{\mathbf{o}}, \eta)|_d < \frac{\lambda}{2}\right) \leqslant 2\, c_{diam} \exp\left(-\frac{1}{c_{diam}}\, e(\lambda)^d\right).$$

Using this bound we have

$$\lambda\, \mathbb{E}\left|H_\lambda(\eta \cap \hat{W}_\lambda) - \hat{H}_\lambda(\eta \cap \hat{W}_\lambda)\right|$$
$$\leqslant \lambda \int_{\mathbb{M}} (\mathbb{E}^{\mathbf{o}_m}|h(C(\mathbf{o}_m, \eta))|^p)^{\frac{1}{p}} \left(2\, c_{diam} \exp\left(-\frac{1}{c_{diam}}\, e(\lambda)^d\right)\right)^{\frac{p-1}{p}} \mathbb{Q}_{\mathbb{M}}(\mathrm{d}m).$$

Now ξ satisfies the p-moment condition 2 (3.10) for $p \in (1, \infty)$ and thus the first factor is bounded by a constant uniformly over all $m \in \mathbb{M}$. Lemma 3.3 follows.

□

Lemma 3.4 (Lemma 4.2 in Flimmel et al. [2020]).
Under the assumptions of Theorem 3.13(i), we have

$$\lim_{\lambda \to \infty} \lambda\, \mathbb{E}\left|H_\lambda(\eta) - \hat{H}_\lambda(\eta)\right| = 0.$$

Proof. We follow the proof of Lemma 3.3. In (3.18), we integrate over $\mathbb{R}^d$ instead of over W_λ, yielding a value of one for the inner integral. Now follow the proof of Lemma 3.3 verbatim.

□

Lemma 3.5 (Lemma 4.3 in Flimmel et al. [2020]).
Under the assumptions of Theorem 3.13(i), we have

$$\lim_{\lambda\to\infty} \mathbb{E}\left|\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta)\right| = 0.$$

Proof. Write

$$\hat{\nu}_\lambda(\hat{x},\eta) := \frac{h(C(\hat{x},\eta))\,\mathbf{1}\{C(\hat{x},\eta)\subset W_\lambda\}}{|W_\lambda\ominus C(\hat{x},\eta)|_d} \times \mathbf{1}\{|W_\lambda\ominus C(\hat{x},\eta)|_d \geqslant \frac{\lambda}{2}\}\,\mathbf{1}\{D_{\hat{x}}\geqslant d(x,W_\lambda)\}, \tag{3.19}$$

where $D_{\hat{x}}$ is the radius of the ball centered at x and containing $C(\hat{x},\eta)$ and where $D_{\hat{x}}$ is equal in distribution to D, with D at (3.16). Here $d(x,W_\lambda)$ denotes the Euclidean distance between x and W_λ. We observe that

$$\mathbb{E}\left|\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta)\right| \leqslant \mathbb{E}\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda^c} |\hat{\nu}_\lambda(\hat{x},\eta)|\,.$$

From now on, we use the notation c to denote a universal positive constant whose value may change from line to line. By the Hölder inequality, the p-moment condition 2 on ξ, and the assumption that $C(\hat{x},\eta)$ has an exponentially decaying tail, we have $\mathbb{E}\,|\hat{\nu}_\lambda(\hat{x},\eta)| \leqslant (c/\lambda)\exp\left(-\frac{1}{c}d(x,W_\lambda)^d\right)$. Thus

$$\mathbb{E}\left|\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta)\right| \leqslant \frac{c}{\lambda}\int_{W_\lambda^c} \exp\left(-\frac{1}{c}\,d(x,W_\lambda)^d\right)\mathrm{d}x.$$

Let $W_{\lambda,\varepsilon}$ be the set of points in W_λ^c at distance ε from W_λ. The co-area formula implies

$$\mathbb{E}\left|\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta)\right| \leqslant \frac{c}{\lambda}\int_0^\infty\int_{W_{\lambda,\varepsilon}} \exp\left(-\frac{1}{c}\,\varepsilon^d\right)\mathcal{H}^{d-1}(\mathrm{d}y)\,\mathrm{d}\varepsilon.$$

Since $\mathcal{H}^{d-1}(W_{\lambda,\varepsilon}) \leqslant c\,(\lambda^{1/d}(1+\varepsilon))^{d-1}$, we get

$$\mathbb{E}\left|\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta)\right| = O(\lambda^{-1/d}).$$

□

Proof of Theorem 3.13.
(i) The asymptotic unbiasedness of $H_\lambda(\eta\cap\hat{W}_\lambda)$, $\hat{H}_\lambda(\eta\cap\hat{W}_\lambda)$ and $\hat{H}_\lambda(\eta)$ is a consequence of Lemmas 3.3, 3.4 and 3.5. For example, concerning $H_\lambda(\eta\cap\hat{W}_\lambda)$, one may write

$$\begin{aligned}|\mathbb{E}\,H_\lambda(\eta\cap\hat{W}_\lambda) - \mathbb{E}^0 h(K_{\mathbf{o}}(\eta))| &\leqslant \mathbb{E}\,|H_\lambda(\eta\cap\hat{W}_\lambda) - H_\lambda(\eta)|\\ &\leqslant \mathbb{E}\,|H_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta\cap\hat{W}_\lambda)| + \mathbb{E}\,|\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) - \hat{H}_\lambda(\eta)|\\ &\quad + \mathbb{E}\,|\hat{H}_\lambda(\eta) - H_\lambda(\eta)|,\end{aligned}$$

which in view of Lemmas 3.3, 3.4 and 3.5 goes to zero as $\lambda \to \infty$. This gives the asymptotic unbiasedness of $H_\lambda(\eta \cap \hat{W}_\lambda)$. One may similarly show the asymptotic unbiasedness for $\hat{H}_\lambda(\eta \cap \hat{W}_\lambda)$ and $\hat{H}_\lambda(\eta)$.

(ii) To show consistency, we introduce $T_\lambda(\eta \cap \hat{W}_\lambda) = \lambda^{-1} \sum_{\hat{x} \in \eta \cap \hat{W}_\lambda} \xi(\hat{x}, \eta)$. By assumption, ξ stabilizes and satisfies the p-moment condition for $p \in (1, \infty)$. Thus, using Theorem 2.1 of Penrose and Yukich [2003], we get that $T_\lambda(\eta \cap \hat{W}_\lambda)$ is a consistent estimator of $\mathbb{E}^0 h(K_\mathbf{o}(\eta))$. To prove the consistency of the estimators in Theorem 3.13(iii), it is enough to show for one of them that it has the same L_1 limit as $T_\lambda(\eta \cap \hat{W}_\lambda)$. We choose $\hat{H}_\lambda(\eta \cap \hat{W}_\lambda)$ and write

$$
\begin{aligned}
&\mathbb{E}\left|\hat{H}_\lambda(\eta \cap \hat{W}_\lambda) - T_\lambda(\eta \cap \hat{W}_\lambda)\right| \\
&= \mathbb{E}\left|\lambda^{-1} \sum_{\hat{x} \in \eta \cap \hat{W}_\lambda} \xi(\hat{x}, \eta)\left(\frac{\lambda \mathbf{1}\{C(\hat{x}, \eta) \subset W_\lambda\} \mathbf{1}\{|W_\lambda \ominus C(\hat{x}, \eta)|_d \geqslant \frac{\lambda}{2}\}}{|W_\lambda \ominus C(\hat{x}, \eta)|_d} - 1\right)\right| \\
&\leqslant \lambda^{-1}\mathbb{E} \sum_{\hat{x} \in \eta \cap \hat{W}_\lambda} |\xi(\hat{x}, \eta)| \left|\frac{\lambda \mathbf{1}\{C(\hat{x}, \eta) \subset W_\lambda\} \mathbf{1}\{|W_\lambda \ominus C(\hat{x}, \eta)|_d \geqslant \frac{\lambda}{2}\}}{|W_\lambda \ominus C(\hat{x}, \eta)|_d} - 1\right| \\
&\leqslant \int_{W_\lambda} \lambda^{-1}\mathbb{E}^0 |h(K_\mathbf{o}(\eta))| \left|\frac{\lambda \mathbf{1}\{x + K_\mathbf{o}(\eta) \subset W_\lambda\} \mathbf{1}\{|W_\lambda \ominus K_\mathbf{o}(\eta)|_d \geqslant \frac{\lambda}{2}\}}{|W_\lambda \ominus K_\mathbf{o}(\eta)|_d} - 1\right| \mathrm{d}x \\
&= \int_{[-\frac{1}{2}, \frac{1}{2}]^d} \mathbb{E}^0\big(|h(K_\mathbf{o}(\eta))| Y_\lambda(u)\big)\, \mathrm{d}u,
\end{aligned}
$$

where we substituted $\lambda^{1/d}u$ for x in the last equality and defined random variables

$$
Y_\lambda(u) := \left|\frac{\lambda \mathbf{1}\{\lambda^{1/d}u + K_\mathbf{o}(\eta) \subset W_\lambda\} \mathbf{1}\{|W_\lambda \ominus K_\mathbf{o}(\eta)|_d \geqslant \frac{\lambda}{2}\}}{|W_\lambda \ominus K_\mathbf{o}(\eta)|_d} - 1\right|.
$$

We show that $Y_\lambda(u)$ converges to zero in $\mathbb{P}^0$ probability for any $u \in (-1/2, 1/2)^d$. Write the inclusion $K_\mathbf{o}(\eta) \subset B_D(\mathbf{o})$, where D has exponentially decaying tails by assumption. We conclude that both $\lambda/|W_\lambda \ominus K_\mathbf{o}(\eta)|_d$ and $\mathbf{1}\{|W_\lambda \ominus K_\mathbf{o}(\eta)|_d \geqslant \lambda/2\}$ tend to one in $\mathbb{P}^0$ probability. To prove the convergence of $Y_\lambda(u)$ to zero in $\mathbb{P}^0$ probability, it remains to show that $\mathbf{1}\{\lambda^{1/d}u + K_\mathbf{o}(\eta) \subset W_\lambda\}$ converges to one in $\mathbb{P}^0$ probability. Equivalently, we show that the $\mathbb{P}^0$ probability of the event $\{\lambda^{1/d}u + K_\mathbf{o}(\eta) \subset W_\lambda\}$ goes to 1. Let $u \in (-1/2, 1/2)^d$ be fixed. Then

$$
\begin{aligned}
&\mathbb{P}^0(\lambda^{1/d}u \in W_\lambda \ominus K_\mathbf{o}(\eta)) \geqslant \mathbb{P}^0(\lambda^{1/d}u \in W_\lambda \ominus B_D(\mathbf{o})) \\
&= \mathbb{P}^0\left(u \in \left[-\frac{1}{2} + \frac{D}{\lambda^{1/d}}, \frac{1}{2} - \frac{D}{\lambda^{1/d}}\right]^d\right) \\
&= \mathbb{P}^0\left(u \in \left[-\frac{1}{2} + \frac{D}{\lambda^{1/d}}, \frac{1}{2} - \frac{D}{\lambda^{1/d}}\right]^d \mid D \leqslant \log\lambda\right) \mathbb{P}^0(D \leqslant \log\lambda) \\
&\quad + \mathbb{P}^0\left(u \in \left[-\frac{1}{2} + \frac{D}{\lambda^{1/d}}, \frac{1}{2} - \frac{D}{\lambda^{1/d}}\right]^d \mid D > \log\lambda\right) \mathbb{P}^0(D > \log\lambda)
\end{aligned}
$$

$$\geqslant \mathbf{1}\left\{u \in \left[-\frac{1}{2}+\frac{\log\lambda}{\lambda^{1/d}}, \frac{1}{2}-\frac{\log\lambda}{\lambda^{1/d}}\right]^d\right\} \mathbb{P}^0(D \leqslant \log\lambda)$$
$$+ \mathbb{P}^0\left(u \in \left[-\frac{1}{2}+\frac{D}{\lambda^{1/d}}, \frac{1}{2}-\frac{D}{\lambda^{1/d}}\right]^d \mid D > \log\lambda\right) \mathbb{P}^0(D > \log\lambda).$$

Again, D has exponentially decaying tails, so the lower bound converges to $\mathbb{P}^0(u \in (-1/2,1/2)^d) = 1$, showing that $Y_\lambda(u)$ goes to zero in probability as $\lambda \to \infty$. We proved that $Y_\lambda(u)$ converge to zero in $\mathbb{P}^0$ probability, but they are also uniformly bounded by one, hence it follows from the moment condition 2 on ξ that $h(K_{\mathbf{o}}(\eta))Y_\lambda(u)$ goes to zero in L^1. Finally, by the dominated convergence theorem, we get

$$\lim_{\lambda\to\infty} \mathbb{E}\left|\hat{H}_\lambda(\eta \cap \hat{W}_\lambda) - T_\lambda(\eta \cap \hat{W}_\lambda)\right| = 0.$$

Thus $\hat{H}_\lambda(\eta \cap \hat{W}_\lambda)$ converges to $\mathbb{E}^0 h(K_{\mathbf{o}}(\eta))$ in L^1 and also in probability. The consistency of the remaining estimators in Theorem 3.13 follows from Lemmas 3.3, 3.4 and 3.5. This completes the proof of Theorem 3.13. □

Proof of Theorem 3.14 (i). We prove the variance asymptotics (3.14). The proof is split into two lemmas (Lemma 3.7 and Lemma 3.8). We first show an auxiliary result used in the proofs of both lemmas. Then we prove the variance asymptotics for $\hat{H}_\lambda(\eta\cap\hat{W}_\lambda)$. This is easier, since, after scaling by λ, the scores are bounded by $2|\xi(\hat{x},\eta)|$ and thus, by assumption, satisfy a p-moment condition 2 for some $p \in (2,\infty)$. Finally, we conclude the proof by showing that the asymptotic variance of $\hat{H}_\lambda(\eta)$ is the same as the asymptotic variance of $\hat{H}_\lambda(\eta \cap \hat{W}_\lambda)$.

Lemma 3.6 (Lemma 4.4 in Flimmel et al. [2020]).
Let $\varphi : \hat{\mathbb{R}}^d \times \mathbf{N} \to \mathbb{R}$ be an exponentially stabilizing function with respect to η and which satisfies the p-moment condition 2 for some $p \in (2,\infty)$. Then there exists a constant $c \in (0,\infty)$ such that for all $\hat{x}, \hat{y} \in \hat{\mathbb{R}}^d$

$$|\mathbb{E}\,\varphi(\hat{x},\eta\cup\{\hat{y}\})\varphi(\hat{y},\eta\cup\{\hat{x}\}) - \mathbb{E}\,\varphi(\hat{x},\eta)\,\mathbb{E}\,\varphi(\hat{y},\eta)|$$
$$\leqslant c\left(\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}}^d} \mathbb{E}\,|\varphi(\hat{x},\eta\cup\{\hat{y}\})|^p\right)^{\frac{2}{p}} \exp\left(-\frac{1}{c}\,\|x-y\|^\alpha\right), \tag{3.20}$$

where $\varphi(\hat{x},\eta) := \varphi(\hat{x},\eta\cup\{\hat{x}\})$ if $\hat{x} \notin \eta$.

Proof. We follow the proof of Lemma 5.2 in Baryshnikov and Yukich [2005] and show that the constant $A_{1,1}$ there involves the moment $(\mathbb{E}\,|\varphi(\hat{x},\eta\cup\{\hat{y}\})|^p)^{\frac{2}{p}}$. Put $R := \max(R_{\hat{x}}, R_{\hat{y}})$, where $R_{\hat{x}}, R_{\hat{y}}$ are the radii of stabilization as in Proposition 3.2 for $\hat{x}$ and $\hat{y}$, respectively. Furthermore, put $r := \|x-y\|/3$ and define the event $E := \{R \leqslant r\}$. Hölder's inequality gives

$$|\mathbb{E}\,\varphi(\hat{x},\eta\cup\{\hat{y}\})\varphi(\hat{y},\eta\cup\{\hat{x}\}) - \mathbb{E}\,\varphi(\hat{x},\eta\cup\{\hat{y}\})\varphi(\hat{y},\eta\cup\{\hat{x}\})\mathbf{1}\{E\}|$$
$$\leqslant c\left(\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}}^d} \mathbb{E}\,|\varphi(\hat{x},\eta\cup\{\hat{y}\})|^p\right)^{\frac{2}{p}} \mathbb{P}(E^c)^{\frac{p-2}{p}}. \tag{3.21}$$

Notice that

$$
\begin{aligned}
&\mathbb{E}\,\varphi(\hat{x},\eta\cup\{\hat{y}\})\varphi(\hat{y},\eta\cup\{\hat{x}\})\mathbf{1}\{E\}\\
&\quad=\mathbb{E}\,\varphi(\hat{x},(\eta\cup\{\hat{y}\})\cap\hat{B}_r(\hat{x}))\varphi(\hat{y},(\eta\cup\{\hat{x}\})\cap\hat{B}_r(\hat{x}))\mathbf{1}\{E\}\\
&\quad=\mathbb{E}\,\varphi(\hat{x},(\eta\cup\{\hat{y}\})\cap\hat{B}_r(\hat{x}))\varphi(\hat{y},(\eta\cup\{\hat{x}\})\cap\hat{B}_r(\hat{x}))(1-\mathbf{1}\{E^c\}).
\end{aligned}
$$

A second application of Hölder's inequality gives

$$
\begin{aligned}
&|\mathbb{E}\,\varphi(\hat{x},\eta\cup\{\hat{y}\})\varphi(\hat{y},\eta\cup\{\hat{x}\})\mathbf{1}\{E\}\\
&\qquad\qquad-\mathbb{E}\,\varphi(\hat{x},(\eta\cup\{\hat{y}\})\cap\hat{B}_r(\hat{x}))\varphi(\hat{y},(\eta\cup\{\hat{x}\})\cap\hat{B}_r(\hat{y}))|\\
&\quad\leqslant c\left(\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}}^d}\mathbb{E}\,|\varphi(\hat{x},\eta\cup\{\hat{y}\})|^p\right)^{\frac{2}{p}}\mathbb{P}(E^c)^{\frac{p-2}{p}}.
\end{aligned}
\tag{3.22}
$$

Combining (3.21) and (3.22) and using independence of $\varphi(\hat{x},(\eta\cup\{\hat{y}\})\cap\hat{B}_r(\hat{x}))$ and $\varphi(\hat{y},(\eta\cup\{\hat{x}\})\cap\hat{B}_r(\hat{y}))$, we have

$$
\begin{aligned}
&|\mathbb{E}\,\varphi(\hat{x},\eta\cup\{\hat{y}\})\varphi(\hat{y},\eta\cup\{\hat{x}\})\\
&\qquad\qquad-\mathbb{E}\,\varphi(\hat{x},(\eta\cup\{\hat{y}\})\cap\hat{B}_r(\hat{x}))\mathbb{E}\,\varphi(\hat{y},(\eta\cup\{\hat{x}\})\cap\hat{B}_r(\hat{y}))|\\
&\quad\leqslant c\left(\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}}^d}\mathbb{E}\,|\varphi(\hat{x},\eta\cup\{\hat{y}\})|^p\right)^{\frac{2}{p}}\mathbb{P}(E^c)^{\frac{p-2}{p}}.
\end{aligned}
\tag{3.23}
$$

Likewise we may show

$$
\begin{aligned}
&|\mathbb{E}\,\varphi(\hat{x},\eta)\mathbb{E}\,\varphi(\hat{y},\eta)-\mathbb{E}\,\varphi(\hat{x},\eta\cap\hat{B}_r(\hat{x}))\mathbb{E}\,\varphi(\hat{y},\eta\cap\hat{B}_r(\hat{y}))|\\
&\quad\leqslant c\left(\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}}^d}\mathbb{E}\,|\varphi(\hat{x},\eta\cup\{\hat{y}\})|^p\right)^{\frac{2}{p}}\mathbb{P}(E^c)^{\frac{p-2}{p}}.
\end{aligned}
\tag{3.24}
$$

Combining (3.23) and (3.24) and using that $\mathbb{P}(E^c)$ decreases exponentially in $\|x-y\|^\alpha$, we thus obtain (3.20).

□

Lemma 3.7 (Lemma 4.5 in Flimmel et al. [2020]). *If ξ is exponentially stabilizing with respect to η then*

$$\lim_{\lambda\to\infty}\lambda\mathrm{Var}\,\hat{H}_\lambda(\eta\cap\hat{W}_\lambda)=\sigma^2(\xi),$$

where $\sigma^2(\xi)$ is at (3.13).

Proof. Put for all $\hat{x}\in\hat{\mathbb{R}}^d$ and any marked point process μ_m,

$$\zeta_\lambda(\hat{x},\mu_m):=\frac{\lambda\,\xi(\hat{x},\mu_m)}{|W_\lambda\ominus C(\hat{x},\mu_m)|_d}\mathbf{1}\{|W_\lambda\ominus C(\hat{x},\mu_m)|_d\geqslant\frac{\lambda}{2}\}$$

and

$$\nu_\lambda(\hat{x},\mu_m):=\zeta_\lambda(\hat{x},\mu_m)\,\mathbf{1}\{C(\hat{x},\mu_m)\subset W_\lambda\}.$$

Note that ζ_λ is translation-invariant whereas ν_λ is not translation-invariant. Then $\lambda\,\hat{H}_\lambda(\eta \cap \hat{W}_\lambda) = \sum_{\hat{x}\in\eta\cap\hat{W}_\lambda} \nu_\lambda(\hat{x},\eta)$.

Recall that $\hat{\mathbb{Q}}$ is the product measure of Lebesgue measure on $\mathbb{R}^d$ and $\mathbb{Q}_{\mathbb{M}}$. By the Slivnyak–Mecke theorem (Theorem 1.9) we have

$$\begin{aligned}
\lambda \mathrm{Var}\,\hat{H}_\lambda(\eta\cap\hat{W}_\lambda) &= \lambda^{-1}\mathbb{E}\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda}\nu_\lambda^2(\hat{x},\eta) + \lambda^{-1}\mathbb{E}\sum_{\hat{x},\hat{y}\in\eta\cap\hat{W}_\lambda;\hat{x}\neq\hat{y}}\nu_\lambda(\hat{x},\eta)\nu_\lambda(\hat{y},\eta)\\
&\qquad -\lambda^{-1}\left(\mathbb{E}\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda}\nu_\lambda(\hat{x},\eta)\right)^2\\
&= \lambda^{-1}\int_{\hat{W}_\lambda}\mathbb{E}\,\nu_\lambda^2(\hat{x},\eta)\,\hat{\mathbb{Q}}(\mathrm{d}\hat{x})\\
&+ \lambda^{-1}\int_{\hat{W}_\lambda}\int_{\hat{W}_\lambda}[\mathbb{E}\,\nu_\lambda(\hat{x},\eta\cup\{\hat{y}\})\nu_\lambda(\hat{y},\eta\cup\{\hat{x}\}) - \mathbb{E}\,\nu_\lambda(\hat{x},\eta)\,\mathbb{E}\,\nu_\lambda(\hat{y},\eta)]\\
&\qquad\qquad \times\hat{\mathbb{Q}}(\mathrm{d}\hat{y})\,\hat{\mathbb{Q}}(\mathrm{d}\hat{x})\\
&=: I_1(\lambda)+I_2(\lambda).
\end{aligned}$$

Here we use the convention that $\nu_\lambda(\hat{x},\mu_m) := \nu_\lambda(\hat{x},\mu_m\cup\{\hat{x}\})$ if $\hat{x}\notin\mu_m$.

Using stationarity and the transformation $u := \lambda^{1/d}x$ we rewrite $I_1(\lambda)$ as

$$I_1(\lambda) = \lambda^{-1}\int_{W_\lambda}\int_{\mathbb{M}}\mathbb{E}\,Z_\lambda^2(\mathbf{o}_m,\eta,x)\,\mathbb{Q}_{\mathbb{M}}(\mathrm{d}m)\,\mathrm{d}x = \int_{W_1}\mathbb{E}\,Z_\lambda^2(\mathbf{o}_M,\eta,\lambda^{1/d}u)\,\mathrm{d}u,$$

where $Z_\lambda((z,m_z),\mu_m,x) := \zeta_\lambda((z,m_z),\mu_m)\,\mathbf{1}\{C((z,m_z),\mu_m)\subset W_\lambda - x\}$. Similarly, by translation invariance of ζ_λ, we have

$$\begin{aligned}
I_2(\lambda) &= \lambda^{-1}\int_{W_\lambda}\int_{W_\lambda - x}\int_{\mathbb{M}}\int_{\mathbb{M}}[\mathbb{E}\,Z_\lambda(\mathbf{o}_{m_1},\eta\cup\{z_{m_2}\},x)\,Z_\lambda(z_{m_2},\eta\cup\{\mathbf{o}_{m_1}\},x)\\
&\qquad - \mathbb{E}\,Z_\lambda(\mathbf{o}_{m_1},\eta,x)\,\mathbb{E}\,Z_\lambda(z_{m_2},\eta,x)]\,\mathbb{Q}_{\mathbb{M}}(\mathrm{d}m_1)\,\mathbb{Q}_{\mathbb{M}}(\mathrm{d}m_2)\,\mathrm{d}z\,\mathrm{d}x\\
&= \int_{W_1}\int_{W_\lambda-\lambda^{1/d}u}[\mathbb{E}\,Z_\lambda(\mathbf{o}_M,\eta\cup\{z_M\},\lambda^{1/d}u)\,Z_\lambda(z_M,\eta\cup\{\mathbf{o}_M\},\lambda^{1/d}u)\\
&\qquad - \mathbb{E}\,Z_\lambda(\mathbf{o}_M,\eta,\lambda^{1/d}u)\,\mathbb{E}\,Z_\lambda(z_M,\eta,\lambda^{1/d}u)]\,\mathrm{d}z\,\mathrm{d}u,
\end{aligned}$$

where $\mathbf{o}_{m_1} := (\mathbf{o},m_1)$, $z_{m_2} := (z,m_2)$, $\mathbf{o}_M := (\mathbf{o},M_{\mathbf{o}})$, $z_M := (z,M_z)$ and $M_{\mathbf{o}}$, M_z are random marks distributed according to $\mathbb{Q}_{\mathbb{M}}$.

Since $|\zeta_\lambda(\hat{x},\eta)| \leqslant 2|\xi(\hat{x},\eta)|$, ζ_λ satisfies a p-moment condition 2, $p\in(2,\infty)$. Recall that $|W_\lambda \ominus C(\hat{x},\eta)|_d/\lambda$ tends in probability to 1 and notice that $W_\lambda - \lambda^{1/d}u$ for $u\in(-1/2,1/2)^d$ increases to $\mathbb{R}^d$ as $\lambda\to\infty$. Thus, as $\lambda\to\infty$, we have for any $\hat{\mathbf{o}} := (\mathbf{o},m_{\mathbf{o}})$, $\hat{z} := (z,m_z)\in\hat{\mathbb{R}^d}$ and $u\in(-1/2,1/2)^d$,

$$\mathbb{E}\,Z_\lambda(\hat{\mathbf{o}},\eta,\lambda^{1/d}u)\to\mathbb{E}\,\xi(\hat{\mathbf{o}},\eta), \tag{3.25}$$

$$\mathbb{E}\,Z_\lambda^2(\hat{\mathbf{o}},\eta,\lambda^{1/d}u)\to\mathbb{E}\,\xi^2(\hat{\mathbf{o}},\eta), \tag{3.26}$$

$$\mathbb{E}\,Z_\lambda(\hat{\mathbf{o}},\eta\cup\{\hat{z}\},\lambda^{1/d}u)Z_\lambda(\hat{z},\eta\cup\{\hat{\mathbf{o}}\},\lambda^{1/d}u)\to\mathbb{E}\,\xi(\hat{\mathbf{o}},\eta\cup\{\hat{z}\})\xi(\hat{z},\eta\cup\{\hat{\mathbf{o}}\}). \tag{3.27}$$

These ingredients are already enough to establish variance asymptotics for $\hat{H}_\lambda(\eta\cap\hat{W}_\lambda)$. Indeed, $I_1(\lambda)$ converges to $\mathbb{E}\,\xi^2(\mathbf{o}_M,\eta)$ by (3.26). Concerning $I_2(\lambda)$, for each $u \in (-1/2, 1/2)^d$ we have

$$\lim_{\lambda\to\infty} \int_{W_\lambda - \lambda^{1/d}u} [\mathbb{E}\, Z_\lambda(\mathbf{o}_M, \eta \cup \{z_M\}, \lambda^{1/d}u) Z_\lambda(z_M, \eta \cup \{\mathbf{o}_M\}, \lambda^{1/d}u) \\ - \mathbb{E}\, Z_\lambda(\mathbf{o}_M, \eta, \lambda^{1/d}u)\, \mathbb{E}\, Z_\lambda(z_M, \eta, \lambda^{1/d}u)]\, \mathrm{d}z \\ = \int_{\mathbb{R}^d} [\mathbb{E}\, \xi(\mathbf{o}_M, \eta \cup \{z_M\}) \xi(z_M, \eta \cup \{\mathbf{o}_M\}) - \mathbb{E}\, \xi(\mathbf{o}_M, \eta) \mathbb{E}\, \xi(z_M, \eta)]\, \mathrm{d}z.$$

Here we use that for any $x \in \mathbb{R}^d$, the function $Z_\lambda(\cdot, \cdot, x) : \hat{\mathbb{R}^d} \times \mathbf{N} \to \mathbb{R}$ is exponentially stabilizing with respect to η and satisfies the p-moment condition 2 for some $p \in (2, \infty)$. Thus, from Lemma 3.6, the integrand is dominated by an exponentially decaying function of $\|z\|^\alpha$. Applying the dominated convergence theorem, together with (3.25) and (3.27), we obtain the desired variance asymptotics since $|W_1|_d = 1$.

□

The next lemma completes the proof of Theorem 3.14 (i).

Lemma 3.8 (Lemma 4.6 in Flimmel et al. [2020]).
If ξ is exponentially stabilizing with respect to η then

$$\lim_{\lambda\to\infty} \lambda \mathrm{Var}\, \hat{H}_\lambda(\eta) = \lim_{\lambda\to\infty} \lambda \mathrm{Var}\, \hat{H}_\lambda(\eta \cap \hat{W}_\lambda) = \sigma^2(\xi).$$

Proof. Write

$$\lambda\, \hat{H}_\lambda(\eta) = \sum_{\hat{x} \in \eta \cap \hat{W}_\lambda} \nu_\lambda(\hat{x}, \eta) + \sum_{\hat{x} \in \eta \cap \hat{W}_\lambda^c} \nu_\lambda(\hat{x}, \eta).$$

Now

$$\lambda \mathrm{Var}\, \hat{H}_\lambda(\eta) = \lambda^{-1} \mathrm{Var} \left(\sum_{\hat{x} \in \eta \cap \hat{W}_\lambda} \nu_\lambda(\hat{x}, \eta) \right) + \lambda^{-1} \mathrm{Var} \left(\sum_{\hat{x} \in \eta \cap \hat{W}_\lambda^c} \nu_\lambda(\hat{x}, \eta) \right) \\ + 2\lambda^{-1} \mathrm{Cov} \left(\sum_{\hat{x} \in \eta \cap \hat{W}_\lambda} \nu_\lambda(\hat{x}, \eta), \sum_{\hat{x} \in \eta \cap \hat{W}_\lambda^c} \nu_\lambda(\hat{x}, \eta) \right).$$

It suffices to show $\mathrm{Var}\left(\sum_{\hat{x} \in \eta \cap \hat{W}_\lambda^c} \nu_\lambda(\hat{x}, \eta)\right) = O(\lambda^{(d-1)/d})$, for then the Cauchy–Schwarz inequality shows that the covariance term in the above expression is negligible compared to λ.

Now we show $\mathrm{Var}\left(\sum_{\hat{x} \in \eta \cap \hat{W}_\lambda^c} \nu_\lambda(\hat{x}, \eta)\right) = O(\lambda^{(d-1)/d})$ as follows. Note that $\hat{H}_\lambda(\eta) = \sum_{\hat{x} \in \eta} \hat{\nu}_\lambda(\hat{x}, \eta)$, where $\hat{\nu}_\lambda(\hat{x}, \eta)$ is at (3.19). By the Slivnyak–Mecke theorem we have

$$\begin{aligned}
\lambda \mathrm{Var}\left(\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda^c} \nu_\lambda(\hat{x},\eta)\right) &= \lambda^{-1}\mathbb{E}\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda^c} \hat{\nu}_\lambda^2(\hat{x},\eta)\\
&\quad + \lambda^{-1}\mathbb{E}\sum_{\hat{x},\hat{y}\in\eta\cap\hat{W}_\lambda^c;\hat{x}\neq\hat{y}} \hat{\nu}_\lambda(\hat{x},\eta)\hat{\nu}_\lambda(\hat{y},\eta) - \lambda^{-1}\left(\mathbb{E}\sum_{\hat{x}\in\eta\cap\hat{W}_\lambda^c}\hat{\nu}_\lambda(\hat{x},\eta)\right)^2\\
&= \lambda^{-1}\int_{\hat{W}_\lambda^c}\mathbb{E}\,\hat{\nu}_\lambda^2(\hat{x},\eta)\,\hat{\mathbb{Q}}(\mathrm{d}\hat{x})\\
&\quad + \lambda^{-1}\int_{\hat{W}_\lambda^c}\int_{\hat{W}_\lambda^c}[\mathbb{E}\,\hat{\nu}_\lambda(\hat{x},\eta\cup\{\hat{y}\})\hat{\nu}_\lambda(\hat{y},\eta\cup\{\hat{x}\}) - \mathbb{E}\,\hat{\nu}_\lambda(\hat{x},\eta)\,\mathbb{E}\,\hat{\nu}_\lambda(\hat{y},\eta)]\\
&\qquad\qquad \times\hat{\mathbb{Q}}(\mathrm{d}\hat{x})\,\hat{\mathbb{Q}}(\mathrm{d}\hat{y})\\
&=: I_1^*(\lambda) + I_2^*(\lambda).
\end{aligned}$$

By the Hölder inequality, the moment condition 2 on ξ and the assumed exponential decay of the tail of the diameter of $C(\hat{x},\eta)$, we have $\mathbb{E}\,\hat{\nu}_\lambda(\hat{x},\eta)^p \leqslant c\exp\left(-\frac{1}{c}\,d(x,W_\lambda)^d\right)$ for some positive constant c. Then, similarly as in Lemma 3.5, we may use the co-area formula to obtain $I_1^*(\lambda) = O(\lambda^{-1/d})$.

To bound $I_2^*(\lambda)$ we appeal to Lemma 3.6. Notice that $|\hat{\nu}_\lambda(\hat{x},\eta)| \leqslant 2|\xi(\hat{x},\eta)|$. Since $\hat{\nu}_\lambda, \lambda \geqslant 1$, are exponentially stabilizing with respect to η and satisfy the p-moment condition 2 for $p\in(2,\infty)$, then by Lemma 3.6

$$\begin{aligned}
&|\mathbb{E}\,\hat{\nu}_\lambda(\hat{x},\eta\cup\{\hat{y}\})\hat{\nu}_\lambda(\hat{y},\eta\cup\{\hat{x}\}) - \mathbb{E}\,\hat{\nu}_\lambda(\hat{x},\eta)\,\mathbb{E}\,\hat{\nu}_\lambda(\hat{y},\eta)|\\
&\qquad\leqslant c\left(\sup_{\hat{x},\hat{y}\in\hat{\mathbb{R}}^d}\mathbb{E}\,|\hat{\nu}_\lambda(\hat{x},\eta\cup\{\hat{y}\})|^p\right)^{\frac{2}{p}}\exp\left(-\frac{1}{c}\,\|x-y\|^\alpha\right).
\end{aligned}$$

Using this estimate we compute

$$\begin{aligned}
I_2^*(\lambda) &\leqslant \lambda^{-1}\int_{\hat{W}_\lambda^c}\int_{W_\lambda^c} c\,(\mathbb{E}\,|\hat{\nu}_\lambda(\hat{x},\eta)|^p)^{\frac{2}{p}}\exp\left(-\frac{1}{c}\,\|x-y\|^\alpha\right)\,\mathrm{d}y\,\hat{\mathbb{Q}}(\mathrm{d}\hat{x})\\
&\leqslant c\,\lambda^{-1}\int_{\hat{W}_\lambda^c}(\mathbb{E}\,|\hat{\nu}_\lambda(\hat{x},\eta)|^p)^{\frac{2}{p}}\int_{\mathbb{R}^d}\exp\left(-\frac{1}{c}\,\|x-y\|^\alpha\right)\,\mathrm{d}y\,\hat{\mathbb{Q}}(\mathrm{d}\hat{x})\\
&\leqslant c\,\lambda^{-1}\int_{W_\lambda^c}\exp\left(-\frac{1}{c}\,d(x,W_\lambda)^d\right)\,\mathrm{d}x\int_{\mathbb{R}^d}\exp\left(-\frac{1}{c}\,\|y\|^\alpha\right)\,\mathrm{d}y.
\end{aligned}$$

Since $\int_{\mathbb{R}^d}\exp(-\|y\|^\alpha/c)\,\mathrm{d}y < \infty$, we obtain

$$I_2^*(\lambda) \leqslant c\,\lambda^{-1}\int_{W_\lambda^c}\exp\left(-\frac{1}{c}\,d(x,W_\lambda)^d\right)\,\mathrm{d}x.$$

Arguing as we did for $I_1^*(\lambda)$ we obtain $I_2^*(\lambda) = O(\lambda^{-1/d})$.

□

Proof of Theorem 3.14 (ii). Now we prove the central limit theorems for $H_\lambda(\eta\cap\hat{W}_\lambda)$ and $H_\lambda(\eta)$. Let us first introduce some notation. Define for any stationary marked point process μ_m on $\hat{\mathbb{R}}^d$,

$$\xi_\lambda(\hat{x}, \mu_m) := \frac{\lambda\, \xi(\lambda^{1/d}\hat{x}, \lambda^{1/d}\mu_m)}{|W_\lambda \ominus C(\lambda^{1/d}\hat{x}, \lambda^{1/d}\mu_m)|_d} \mathbf{1}\{C(\lambda^{1/d}\hat{x}, \lambda^{1/d}\mu_m) \subset W_\lambda\},$$

$$\hat{\xi}_\lambda(\hat{x}, \mu_m) := \xi_\lambda(\hat{x}, \mu_m)\, \mathbf{1}\left\{|W_\lambda \ominus C(\lambda^{1/d}\hat{x}, \lambda^{1/d}\mu_m)|_d \geqslant \frac{\lambda}{2}\right\},$$

where $\lambda^{1/d}\hat{x} := (\lambda^{1/d}x, m_x)$ and $\lambda^{1/d}\mu_m := \{\lambda^{1/d}\hat{x} : \hat{x} \in \mu_m\}$.

Put

$$S_\lambda(\eta_\lambda \cap \hat{W}_1) := \sum_{\hat{x} \in \eta_\lambda \cap \hat{W}_1} \xi_\lambda(\hat{x}, \eta_\lambda), \quad \hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1) := \sum_{\hat{x} \in \eta_\lambda \cap \hat{W}_1} \hat{\xi}_\lambda(\hat{x}, \eta_\lambda),$$

as well as

$$S_\lambda(\eta_\lambda) := \sum_{\hat{x} \in \eta_\lambda} \xi_\lambda(\hat{x}, \eta_\lambda), \quad \hat{S}_\lambda(\eta_\lambda) := \sum_{\hat{x} \in \eta_\lambda} \hat{\xi}_\lambda(\hat{x}, \eta_\lambda).$$

Notice that

$$S_\lambda(\eta_\lambda \cap \hat{W}_1) \overset{\mathcal{D}}{=} \lambda\, H_\lambda(\eta \cap \hat{W}_\lambda), \quad S_\lambda(\eta_\lambda) \overset{\mathcal{D}}{=} \lambda\, H_\lambda(\eta)$$

and

$$\hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1) \overset{\mathcal{D}}{=} \lambda\, \hat{H}_\lambda(\eta \cap \hat{W}_\lambda) \quad \text{and} \quad \hat{S}_\lambda(\eta_\lambda) \overset{\mathcal{D}}{=} \lambda\, \hat{H}_\lambda(\eta)$$

due to the distributional identity $\lambda^{1/d}\eta_\lambda \overset{\mathcal{D}}{=} \eta_1$. The reason for expressing the statistic $\lambda\, H_\lambda(\eta \cap \hat{W}_\lambda)$ in terms of the scores $\xi_\lambda(\hat{x}, \eta_\lambda)$ is that it puts us in a better position to apply the normal approximation results of Theorem 3.10 to the sums $S_\lambda(\eta_\lambda \cap \hat{W}_1)$.

In particular, we use the previous result to establish a central limit theorem for $\hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1)$. We may put $\mathbb{X}$ to be $\mathbb{R}^d$ and we let $\mathbb{Q}$ be Lebesgue measure on $\mathbb{R}^d$ so that η_λ has intensity measure $\lambda\mathbb{Q}$, and we put $K = W_1$. We may write $\hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1) = \sum_{\hat{x} \in \eta_\lambda \cap \hat{W}_1} \hat{\xi}_\lambda(\hat{x}, \eta_\lambda)\, \mathbf{1}\{x \in W_1\}$. Note that $\hat{\xi}_\lambda(\hat{x}, \eta_\lambda)\mathbf{1}\{x \in W_1\}$, $\hat{x} \in \hat{\mathbb{X}}$, are exponentially stabilizing with respect to the input η_λ, they satisfy the p-moment condition 2 for some $p \in (4, \infty)$, they vanish for $x \in W_1^c$, and they (trivially) decay exponentially fast with respect to the distance to K. (see Definition 3.17), Since the distance to K is zero for $x \in K$ this condition is trivially satisfied. This makes $I_{K,\lambda} = \Theta(\lambda)$ where $I_{K,\lambda}$ is defined at (3.8). Thus all conditions of Theorem 3.10 are fulfilled and we deduce a central limit theorem for $\hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1)$ and hence for $\hat{H}_\lambda(\eta \cap \hat{W}_\lambda)$.

We may also apply Theorem 3.10 to show a central limit theorem for $\hat{S}_\lambda(\eta_\lambda)$. For $x \in W_1^c$ we find the radius D_x such that $C(\lambda^{1/d}\hat{x}, \lambda^{1/d}\eta_\lambda) \subset B_{D_x}(\lambda^{1/d}x)$. Then the score $\hat{\xi}_\lambda(\hat{x}, \eta_\lambda)$ vanishes if $D_x > d(\lambda^{1/d}x, W_\lambda)$. As in Proposition 3.1, D_x has exponentially decaying tails and thus $\hat{\xi}_\lambda$ decays exponentially fast with respect to the distance to K.

Let $d_K(X, Y)$ denote the Kolmogorov distance between random variables X and Y. Applying Theorem 3.10 we obtain

$$d_K\left(\frac{\hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1) - \mathbb{E}\,\hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1)}{\sqrt{\operatorname{Var} \hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1)}}, N(0,1)\right) \leqslant \frac{c}{\sqrt{\operatorname{Var} \hat{S}_\lambda(\eta_\lambda \cap \hat{W}_1)}}$$

and

$$d_K\left(\frac{\hat{S}_\lambda(\eta_\lambda)-\mathbb{E}\,\hat{S}_\lambda(\eta_\lambda)}{\sqrt{\operatorname{Var}\hat{S}_\lambda(\eta_\lambda)}}, N(0,1)\right)\leqslant \frac{c}{\sqrt{\operatorname{Var}\hat{S}_\lambda(\eta_\lambda)}}.$$

Combining this with (3.14) and using $\operatorname{Var}\hat{S}_\lambda(\eta_\lambda\cap\hat{W}_1)\geqslant c\,\lambda$, we obtain as $\lambda\to\infty$

$$\frac{\hat{S}_\lambda(\eta_\lambda\cap\hat{W}_1)-\mathbb{E}\,\hat{S}_\lambda(\eta_\lambda\cap\hat{W}_1)}{\sqrt{\lambda}}\xrightarrow{\mathcal{D}} N(0,\sigma^2(\xi))$$

and

$$\frac{\hat{S}_\lambda(\eta_\lambda)-\mathbb{E}\,\hat{S}_\lambda(\eta_\lambda)}{\sqrt{\lambda}}\xrightarrow{\mathcal{D}} N(0,\sigma^2(\xi)).$$

To show that

$$\frac{S_\lambda(\eta_\lambda\cap\hat{W}_1)-\mathbb{E}\,S_\lambda(\eta_\lambda\cap\hat{W}_1)}{\sqrt{\lambda}}\xrightarrow{\mathcal{D}} N(0,\sigma^2(\xi)), \tag{3.28}$$

as $\lambda\to\infty$, it suffices to show $\lim_{\lambda\to\infty}\mathbb{E}\,|S_\lambda(\eta_\lambda\cap\hat{W}_1)-\hat{S}_\lambda(\eta_\lambda\cap\hat{W}_1)|=0$. Since $\mathbb{E}\,|S_\lambda(\eta_\lambda\cap\hat{W}_1)-\hat{S}_\lambda(\eta_\lambda\cap\hat{W}_1)|=\lambda\,\mathbb{E}\,|H_\lambda(\eta_\lambda\cap\hat{W}_\lambda)-\hat{H}_\lambda(\eta_\lambda\cap\hat{W}_\lambda)|$, we may use Lemma 3.3 to prove (3.28). Likewise, to obtain the central limit theorem for $S_\lambda(\eta_\lambda)$, it suffices to show $\lim_{\lambda\to\infty}\mathbb{E}\,|S_\lambda(\eta_\lambda)-\hat{S}_\lambda(\eta_\lambda)|=0$, which is a consequence of Lemma 3.4. Hence we deduce from the central limit theorem for $\hat{S}_\lambda(\eta_\lambda)$ that as $\lambda\to\infty$

$$\frac{S_\lambda(\eta_\lambda)-\mathbb{E}\,S_\lambda(\eta_\lambda)}{\sqrt{\lambda}}\overset{\mathcal{D}}{=}\sqrt{\lambda}\left(H_\lambda(\eta)-\mathbb{E}^0 h(K_{\mathbf{o}}(\eta))\right)\xrightarrow{\mathcal{D}} N(0,\sigma^2(\xi)).$$

This completes the proof of Theorem 3.14 (ii). □

Proofs of Theorems 3.15 and 3.16

Before giving the proof of Theorem 3.15 we recall from Proposition 3.2 that translation-invariant cell characteristics ξ^{ρ_i} are exponentially stabilizing with respect to Poisson input η. This allows us to apply Theorem 3.14 to cell characteristics of tessellations defined by ρ_i, $i=1,2,3$. For example, we can take $h(\cdot)$ to be either the volume or surface area of a cell or the radius of the circumscribed or inscribed ball.

Proof of Theorem 3.15. (i) The assertion of unbiasedness follows from Theorem 3.12. (ii) To prove the asymptotic normality, we write

$$h(C^{\rho_i}(\hat{x},\eta)):=\mathbf{1}\{|C^{\rho_i}(\hat{x},\eta)|_d\leqslant t\}=:\varphi^{\rho_i}(\hat{x},\eta).$$

To deduce (3.15) from Theorem 3.14(ii) we need only verify the p-moment condition 2 for $p\in(4,\infty)$ and the positivity of $\sigma^2(\varphi^{\rho_i})$. The moment condition holds for all $p\in[1,\infty)$ since φ is bounded by 1. To verify the positivity of $\sigma^2(\varphi^{\rho_i})$, we recall Remark (i) following Theorem 3.14. More precisely we may use Theorem 2.1 of Penrose and Yukich [2001] and show that there is an almost surely finite random variable S and a non-degenerate random variable $\Delta^{\rho_i}(\infty)$ such that for all finite $\mathcal{A} \subset \hat{B}_S(\mathbf{o})^c$ we have

$$\Delta^{\rho_i}(\infty) = \sum_{\hat{x} \in (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}} \mathbf{1}\{|C^{\rho_i}(\hat{x}, (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\})|_d \leqslant t\} - \sum_{\hat{x} \in (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A}} \mathbf{1}\{|C^{\rho_i}(\hat{x}, (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A})|_d \leqslant t\}.$$

We first explain the argument for the Voronoi case and then indicate how to extend it to treat the Laguerre and Johnson–Mehl tessellations.

Let $t \in (0, \infty)$ be arbitrary but fixed. Let N be the smallest integer of even parity that is larger than $4\sqrt{d}$. The choice of this value will be explained later in the proof. For $L > 0$ we consider a collection of N^d cubes $Q_{L,1}, \dots, Q_{L,N^d}$ centered around $x_i, i = 1, \dots, N^d$, such that

(i) $Q_{L,i}$ has side length $\frac{L}{N}$, and

(ii) $\cup\{Q_{L,i}, i = 1, \dots, N^d\} = [-\frac{L}{2}, \frac{L}{2}]^d$.

Put $\varepsilon_L := L/100N$ and $\hat{Q}_{L,i} := Q_{L,i} \times \mathbb{M}$. Define the event

$$E_{L,N} := \left\{|\eta \cap \hat{Q}_{L,i} \cap \hat{B}_{\varepsilon_L}(x_i)| = 1,\ |\eta \cap \hat{Q}_{L,i} \cap \hat{B}^c_{\varepsilon_L}(x_i)| = 0, \forall i = 1, \dots, N^d\right\}.$$

Elementary properties of the Poisson point process show that $\mathbb{P}(E_{L,N}) > 0$ for all L and N.

On $E_{L,N}$ the faces of the tessellation restricted to $[\frac{-L}{2}, \frac{L}{2}]^d$ nearly coincide with the union of the boundaries of $Q_{L,i}, i = 1, \dots, N^d$ and the cell generated by $\hat{x} \in \eta \cap [\frac{-L}{2} + \frac{L}{N}, \frac{L}{2} - \frac{L}{N}]^d$ is determined only by $\eta \cap (\cup\{Q_{L,j}, j \in I(\hat{x})\})$, where $j \in I(\hat{x})$ if and only if $\hat{x} \in \hat{Q}_{L,j}$ or $\hat{Q}_{L,j} \cap \hat{Q}_{L,i} \neq \emptyset$ for i such that $\hat{x} \in \hat{Q}_{L,i}$. Thus inserting a point at the origin will not affect the cells far from the origin. More precisely, the cells around the points outside $\hat{R}_{L,N} := [-\frac{2L}{N}, \frac{2L}{N}]^d \times \mathbb{M}$ are not affected by inserting a point at the origin. For $S_L := L/2$ we have $\hat{R}_{L,N} \subset \hat{B}_{S_L}(\mathbf{o})$ due to our choice of the value N. Therefore,

$$C^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}) = C^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A})$$

for any finite $\mathcal{A} \subset \hat{B}_{S_L}(\mathbf{o})^c$ and $\hat{x} \in (\eta \cap (\hat{B}_{S_L}(\mathbf{o}) \backslash \hat{R}_{L,N})) \cup \mathcal{A}$. Consequently, on $E_{L,N}$,

$$\Delta^{\rho_1}(\infty) = \sum_{\hat{x} \in (\eta \cap \hat{R}_{L,N}) \cup \{\mathbf{o}_M\}} \mathbf{1}\{|C^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\})|_d \leqslant t\} - \sum_{\hat{x} \in \eta \cap \hat{R}_{L,N}} \mathbf{1}\{|C^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A})|_d \leqslant t\}.$$

Figure 3.1 illustrates the difference appearing in $\Delta^{\rho_1}(\infty)$ on $E_{L,N}$ for $d = 2$. The cells generated by the points outside the square $[-\frac{2L}{N}, \frac{2L}{N}]^2$ are identical for both point configurations whereas the cells generated by the points inside the square may differ.

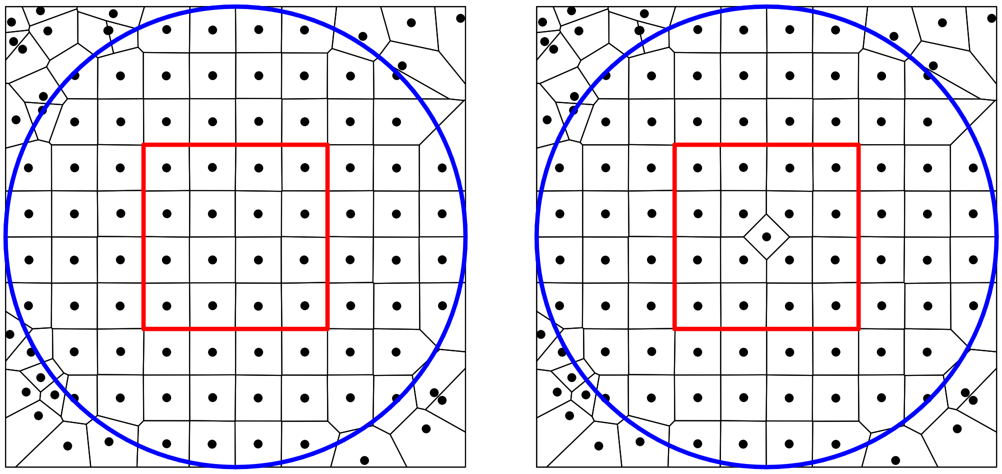


Figure 3.1: Voronoi tessellations in $[-\frac{L}{2}, \frac{L}{2}]^2$ generated by $(\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A}$ (left) and $(\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}$ (right). The ball $B_{S_L}(\mathbf{o})$ shown in blue encloses the red square $[-\frac{2L}{N}, \frac{2L}{N}]^2$ where here N has a value of 10.

On the event $E_{L,N}$, the cell generated by $\hat{x} \in (\eta \cap \hat{R}_{L,N}) \cup \{\mathbf{o}_M\}$ is contained in $\cup\{Q_{L,j}, j \in I(\hat{x})\}$ and thus

$$\sup_{\hat{x} \in (\eta \cap \hat{R}_{L,N}) \cup \{\mathbf{o}_M\}} |C^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A})|_d \leqslant \left(\frac{3L}{N}\right)^d .$$

If $L \in (0, Nt^{1/d}/3)$, then all cell volumes in $\hat{R}_{L,N}$ are at most t; thus $\Delta^{\rho_1}(\infty) = 1$ on the event $E_{L_1,N}$ with $L_1 := \frac{1}{6}Nt^{1/d}$. Similarly,

$$\inf_{\hat{x} \in (\eta \cap \hat{R}_{L,N}) \cup \{\mathbf{o}_M\}} |C^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_{S_L}(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\})|_d \geqslant \left(\frac{L}{3N}\right)^d .$$

If $L \in (3Nt^{1/d}, \infty)$, then all the cell volumes in $\hat{R}_{L,N}$ exceed t and thus $\Delta^{\rho_1}(\infty) = 0$ on the event $E_{L_2,N}$ with $L_2 := 6Nt^{1/d}$. Taking $S := S_{L_1}\mathbf{1}\{E_{L_1,N}\} + S_{L_2}\mathbf{1}\{E_{L_2,N}\}$, we have found two disjoint events $E_{L_1,N}$ and $E_{L_2,N}$, each having positive probability, such that $\Delta^{\rho_1}(\infty)$ takes different values on these events, and thus it is non-degenerate. Hence, $\sigma^2(\varphi^{\rho_1}) > 0$ and we can apply Theorem 3.14(ii).

To prove the positivity of $\sigma^2(\varphi^{\rho_2})$ and $\sigma^2(\varphi^{\rho_3})$ we shall consider a subset of $E_{L,N}$. Assume there exists a parameter $\mu^* \in [0, \mu]$ and a small interval $I_\alpha(\mu^*) \subset [0, \mu]$ for some $\alpha \geqslant 0$ such that $\mathbb{Q}_{\mathbb{M}}(I_\alpha(\mu^*)) > 0$. Define $\hat{E}_{L,N}$ to be the intersection of $E_{L,N}$ and the event $F_{L,N,\alpha}$ that the Poisson points in $[-L/2, L/2]^d$ have marks in $I_\alpha(\mu^*)$. If α is small enough, then the Laguerre and Johnson–Mehl cells nearly coincide with the Voronoi cells on the event $\hat{E}_{L,N}$. Consideration of the events $\hat{E}_{L_1,N}$ and $\hat{E}_{L_2,N}$ shows that $\Delta^{\rho_2}(\infty)$ and $\Delta^{\rho_3}(\infty)$ are non-degenerate, implying that $\sigma^2(\varphi^{\rho_2}) > 0$ and $\sigma^2(\varphi^{\rho_3}) > 0$. Thus Theorem 3.15 holds for the Laguerre and Johnson–Mehl tessellations. □

Remark. In the same way, one can establish that Theorem 3.15 holds for any h taking the form

$$h(K) = \mathbf{1}\{g(K) \leqslant t\} \quad \text{or} \quad h(K) = \mathbf{1}\{g(K) > t\}$$

for $t \in (0,\infty)$ fixed and $g : (\mathcal{F}^d, \mathcal{B}(\mathcal{F}^d)) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$, a homogeneous function of order q, i.e., $g(\alpha K) = \alpha^q g(K)$ for all $K \in \mathbf{F}^d$ and $\alpha \in (0,\infty)$. Examples of the function g include (a) $g(K) := \mathcal{H}^{d-1}(\partial K)$, (b) $g(K) := \mathrm{diam}(K)$, (c) $g(K) :=$ radius of the circumscribed ball of K, and (d) $g(K) :=$ radius of the circumscribed ball of K.

Proof of Theorem 3.16. The unbiasedness is again a consequence of Theorem 3.12. To prove the asymptotic normality, we need to check the p-moment condition 2 for

$$\xi^{\rho_i}(\hat{x},\eta) := \mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta))\mathbf{1}\{C^{\rho_i}(\hat{x},\eta) \text{ is bounded}\}$$

and the positivity of $\sigma^2(\xi^{\rho_i}), i = 1,2,3$.

First we verify the moment condition 2 with $p = 5$. Given any $\hat{x}, \hat{y} \in \hat{\mathbb{R}}^d$, we assert that $\mathbb{E}^{\hat{x},\hat{y}}\mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta))^5 = \mathbb{E}\,\mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta \cup \{\hat{y}\}))^5 \leqslant c < \infty$ for some constant c that does not depend on $\hat{x}$ and $\hat{y}$. From Proposition 3.2 there is a random variable $R_{\hat{x}}$ such that

$$C^{\rho_i}(\hat{x},\eta \cup \{\hat{y}\}) = \bigcap_{\hat{z} \in (\eta \cup \{\hat{y}\} \setminus \{\hat{x}\}) \cap \hat{B}_{R_{\hat{x}}}(x)} \mathbb{H}_{\hat{z}}(\hat{x}).$$

As in Proposition 3.1 we find $D_{\hat{x}}$ such that $C^{\rho_i}(\hat{x},\eta \cup \{\hat{y}\}) \subset B_{D_{\hat{x}}}(\hat{x})$. Then

$$\begin{aligned}\mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta \cup \{\hat{y}\})) &\leqslant \sum_{\hat{z} \in (\eta \cup \{\hat{y}\} \setminus \{\hat{x}\}) \cap \hat{B}_{R_{\hat{x}}}(x)} \mathcal{H}^{d-1}(\partial \mathbb{H}_{\hat{z}}(\hat{x}) \cap B_{D_{\hat{x}}}(\hat{x})) \\ &\leqslant c_{i,d} D_{\hat{x}}^{d-1} \eta(\hat{B}_{R_{\hat{x}}}(x))\end{aligned}$$

for some constant $c_{i,d}$ that depends only on i and d. Using the Cauchy–Schwarz inequality we get

$$\mathbb{E}\,\mathcal{H}^{d-1}(\partial C^{\rho_i}(\hat{x},\eta \cup \{\hat{y}\}))^5 \leqslant c_{i,d}^5 (\mathbb{E}\,D_{\hat{x}}^{10(d-1)})^{1/2} (\mathbb{E}\,\eta(\hat{B}_{R_{\hat{x}}}(x))^{10})^{1/2}.$$

By the property of the Poisson distribution we have

$$\mathbb{E}\,\eta(\hat{B}_{R_{\hat{x}}}(x))^{10} = \mathbb{E}\,(\mathbb{E}\,(\eta(\hat{B}_{R_{\hat{x}}}(x))^{10} \mid R_{\hat{x}})) = \mathbb{E}\,P(|B_{R_{\hat{x}}}(x)|_d),$$

where $P(\cdot)$ is a polynomial of degree 10. Both $D_{\hat{x}}$ and $R_{\hat{x}}$ have exponentially decaying tails and the decay is not depending on $\hat{x}$. Therefore, $(\mathbb{E}\,D_{\hat{x}}^{10(d-1)})^{1/2}(\mathbb{E}\,\eta(\hat{B}_{R_{\hat{x}}}(x))^{10})^{1/2}$ is bounded and the moment condition 2 is satisfied with $p = 5$.

The positivity of the asymptotic variance can be shown similarly as in the proof of Theorem 3.15. We will show it only for the Voronoi case, as the Laguerre and Johnson–Mehl tessellations can be treated similarly. We will again find a random variable S and a $\Delta^{\rho_1}(\infty)$ such that for all finite $\mathcal{A} \subset \hat{B}_S(\mathbf{o})^c$ we have

$$\begin{aligned}\Delta^{\rho_1}(\infty) = &\sum_{\hat{x} \in (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}} \xi^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A} \cup \{\mathbf{o}_M\}) \\ &- \sum_{\hat{x} \in (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A}} \xi^{\rho_1}(\hat{x}, (\eta \cap \hat{B}_S(\mathbf{o})) \cup \mathcal{A})\end{aligned}$$

and moreover $\Delta^{\rho_1}(\infty)$ assumes different values on two events having positive probability and is thus non-degenerate. By Theorem 2.1 of Penrose and Yukich [2001], this is enough to show the positivity of $\sigma^2(\xi^{\rho_1})$.

Let $L > 0$ and let $N \in \mathbb{N}$ have odd parity. Abusing notation, we construct a collection of N^d cubes $Q_{L,1}, \ldots, Q_{L,N^d}$ centered around $x_i \in \mathbb{R}^d, i = 1, \ldots, N^d$ such that

(i) $Q_{L,i}$ has side length $\frac{L}{N}$, and

(ii) $\cup\{Q_{L,i}, i = 1, \ldots, N^d\} = [-\frac{L}{2}, \frac{L}{2}]^d$.

There is a unique index $i_0 \in \{1, \ldots, N^d\}$ such that $x_{i_0} = \mathbf{o}$. We define $\varepsilon_L, \hat{Q}_{L,i}$ and the event $E_{L,N}$ as in the proof of Theorem 3.15. Note that under $E_{L,N}$

$$\inf_{(x,m_x)\in\eta\cap\hat{Q}_{L,i_0}} \|x\| \leqslant \varepsilon_L.$$

Hence, on the event $E_{L,N}$, the insertion of the origin into the point configuration creates a new face of the tessellation whose surface area is bounded below by $c_{min}(L/N)^{d-1}$ and bounded above by $c_{max}(L/N)^{d-1}$. Thus

$$c_{min}\left(\frac{L}{N}\right)^{d-1} + O\left(\varepsilon_L\left(\frac{L}{N}\right)^{d-2}\right) \leqslant \Delta^{\rho_1}(\infty) \leqslant c_{max}\left(\frac{L}{N}\right)^{d-1} - O\left(\varepsilon_L\left(\frac{L}{N}\right)^{d-2}\right),$$

where $O(\varepsilon_L \left(\frac{L}{N}\right)^{d-2})$ is the change in the combined surface areas of the already existing faces after inserting the origin. Events $E_{L_1,N}, E_{L_2,N}, L_1 < L_2$, both occur with positive probability for any L_1, L_2. Similarly as in the proof of Theorem 3.15 we can find N, S, L_1 and L_2 ($L_2 - L_1$ large enough) such that the value of $\Delta^{\rho_1}(\infty)$ differs on each event. Thus $\sigma^2(\xi^{\rho_1})$ is strictly positive.

To show that $\sigma^2(\xi^{\rho_2})$ and $\sigma^2(\xi^{\rho_3})$ are strictly positive we argue as follows. The Laguerre and Johnson–Mehl tessellations are close to the Voronoi tessellation on the event $F_{L,N,\alpha}$, for α small. Arguing as we did in the proof of Theorem 3.15 and considering the event $\hat{E}_{L,N}$ given in the proof of that theorem, we may conclude that $\sigma^2(\xi^{\rho_2}) > 0$ and $\sigma^2(\xi^{\rho_3}) > 0$. □

4. Method of cumulants

4.1 Cumulants of random variables

In Section 1.1, we defined factorial cumulant measures of a point process μ implicitly as the measures occurring in the Taylor series expansion of the logarithm of the probability generating functional. Alternatively, it was defined using the relation (1.4) from the factorial moment measures of μ. This relation is an analogy to the relation between moments and cumulants (also known as semi-invariants) of a real-valued random variable X. We recall here the definition of cumulant, and we present several of its properties. We refer to Brillinger [1975], Shiryayev [1984] and Gnedenko and Kolmogorov [1954] for further detailed probabilistic aspects of this topic.

Denote by $\mu_k := \mathbb{E}\, X^k$ the k-th moment of a random variable X and assume it is finite. Provided that it has a Taylor expansion about the origin, the *moment generating function* equals

$$M_X(r) := \mathbb{E}\, e^{rX} = \sum_{k=0}^{\infty} \mu_k \frac{r^k}{k!}. \tag{4.1}$$

The k-th moment of X is then the k-th derivative of M_X at the origin. The logarithm of the moment generating function is called the *cumulant generating function* since its Taylor expansion about the origin

$$K_X(r) := \log M_X(r) = \sum_{k=0}^{\infty} \kappa_k \frac{r^k}{k!} \tag{4.2}$$

contains the cumulants $\kappa_k := \mathbf{Cum}_k(X)$ as the coefficients. Evidently $\mu_0 = 1$ implies $\kappa_0 = 0$. By extracting coefficients from the expansion, one can further see that $\kappa_1 = \mu_1$, κ_2 is the variance and $\kappa_3 = \mathbb{E}\,(X-\mu_1)^3$. Higher-order cumulants are, however, not the same as centered moments. Explicit relations between higher-order cumulants and moments may be established by formal manipulations of the series (4.1) and (4.2) (see e.g. Corollary 3.1 in Peccati and Taqqu [2008]). If $\mathbb{E}\,|X|^k < \infty$, we have

$$\kappa_k = \sum_{l=1}^{k} (-1)^{l-1}(l-1)! \sum_{k_1+\cdots+k_l=k;\, k_i \geqslant 1} \prod_{i=1}^{l} \mu_{k_i} \tag{4.3}$$

and inversely,

$$\mu_k = \sum_{l=1}^{k} \sum_{k_1+\cdots+k_l=k,\, k_i \geqslant 1} \prod_{i=1}^{l} \kappa_{k_i}.$$

Hence, the existence of the moment μ_k implies the existence of all cumulants up to the order k.

While the moments of a random variable have a simpler interpretation than the cumulants, the cumulants are often mathematically easier to handle. In particular, the advantage of working with cumulant generating function over the ordinary moment generating function is the additivity for two independent random variables X, Y, that is

$$\begin{aligned} K_{X+Y}(r) &= \log \mathbb{E}\left[e^{r(X+Y)}\right] \\ &= \log \mathbb{E}\left[e^{rX}\right]\mathbb{E}\left[e^{rY}\right] \\ &= \log \mathbb{E}\left[e^{rX}\right] + \log \mathbb{E}\left[e^{rY}\right] \\ &= K_X(r) + K_Y(r). \end{aligned}$$

Hence, each cumulant of a sum of independent random variables is the sum of the corresponding cumulants. Similarly, for any constant $c \in \mathbb{R}$,

$$K_{X+c}(r) = K_X(r) + rc. \tag{4.4}$$

Since the k-th cumulant of X is the k-th derivative of K_X at the origin, we get immediately that the first cumulant of the random variable $X + c$ equals the first cumulant of X plus c and the higher-order cumulants are shift-invariant.

Moreover, for any $n \in \mathbb{N}$ and any constant $c \in \mathbb{R}$, we have

$$\mathbf{Cum}_n(cX) = c^n \mathbf{Cum}_n(X). \tag{4.5}$$

This property can be seen directly from (4.3).

Example 4.1 (Cumulants of a Gaussian random variable).
The cumulant generating function of a Gaussian random variable $Z \sim N(\mu, \sigma^2)$ is

$$K_Z(r) = r\mu + r^2\sigma^2/2.$$

Apart from what has been said generally for cumulants of the first and the second order, all cumulants of order three and higher are zero.

It follows from a classical result of Marcinkiewicz (see Marcinkiewicz [1939] or p. 213 in Lukacs [1970]) that if all but a finite number of cumulants of a random variable are non-zero then the random variable must either have a Gaussian distribution or be a constant. Either way we have $\kappa_k = 0$ for all $k \geqslant 3$.

By the moment convergence theorem formulated in terms of cumulants (see e.g. Saulis and Statulevičius [1991]), the convergence of the cumulants of the third and higher orders to zero is equivalent to the convergence in distribution to a Gaussian random variable (assuming the degenerated case is also considered as Gaussian). The latter statement forms the key idea of what is known under the *method of cumulants*.

4.2 Cumulant expansion technique

To our knowledge, there is no unifying approach in the literature for the usage of the method of cumulants in proving asymptotic properties of random geometric objects. Usually, it appears in combination with one of the previously described methods. We have already mentioned in Chapter 3 the moment approach in stabilization introduced in Baryshnikov and Yukich [2005] using the method of cumulants described in the previous section. Alternatively, it can be used directly on a case by case basis for a given functional of a given random geometric object.

Some ideas shall be demonstrated on the planar cylinder process Ξ described in Example 1.15. The content of this and the following section is mainly based on papers Flimmel and Heinrich [2021+] and Heinrich and Spiess [2013].

Assume that $K \subset \mathbb{R}^2$ is a fixed compact set that is star-shaped w.r.t. the origin $\mathbf{o}$ being an inner point of K. The aim of this and the subsequent section is to study the limit behaviour of the random variable $|\Xi \cap \rho K|_2$ as $\rho \to \infty$. The set K is chosen so that $\rho K \nearrow \mathbb{R}^2$. First, we shall link the cumulants of $|\Xi \cap \rho K|_2$ with the characteristics of the unmarked point process Ψ, namely with its factorial moment measures and factorial cumulant measures.

We assume that the probability space $[\Omega, \mathcal{F}, \mathbb{P}]$ on which the underlying marked point process $\Psi^P_{F,G}$ (recall Example 1.15) is defined can be chosen in such a way that the mapping $(x, \omega) \mapsto \mathbf{1}_{\Xi(\omega)}(x) \in \{0,1\}$ for $(x,\omega) \in \mathbb{R}^2 \times \Omega$ is measurable w.r.t. the product σ-field $\mathcal{B}(\mathbb{R}^2) \otimes \mathcal{F}$, see Appendix in Heinrich [2005]. This enables us to apply Fubini's theorem to the random field of indicator variables $\{\mathbf{1}_\Xi(x), x \in \mathbb{R}^2\}$ and implies that the *k-th moment function*

$$p^{(k)}_\Xi(x_1,\dots,x_k) := \mathbb{E}\left(\prod_{j=1}^k \mathbf{1}_\Xi(x_j)\right) = \mathbb{P}(x_1 \in \Xi, \dots, x_k \in \Xi), \quad x_1,\dots,x_k \in \mathbb{R}^2, \tag{4.6}$$

are $\mathcal{B}(\mathbb{R}^{2k})$-measurable for any $k \in \mathbb{N}$.

By $\mathcal{C}^{(2)}$, we denoted the family of non-empty compact sets in $\mathbb{R}^2$. The *Choquet functional* of Ξ is defined by

$$T_\Xi(X) := \mathbb{P}(\Xi \cap X \neq \emptyset), \quad X \in \mathcal{C}^{(2)}. \tag{4.7}$$

In particular, the k-th order moment functions $p^{(k)}_{\Xi^c}$ of the $0-1$-random field $\xi(x) := \mathbf{1}_{\Xi^c}(x)$ can be expressed by (4.6) and (4.7) for any $k \geqslant 1$:

$$\begin{aligned} p^{(k)}_{\Xi^c}(x_1,\dots,x_k) &= \mathbb{E}\left(\prod_{i=1}^k \xi(x_i)\right) \\ &= \mathbb{P}(\{x_1,\dots,x_k\} \cap \Xi = \emptyset) \\ &= 1 - T_\Xi(\{x_1,\dots,x_k\}). \end{aligned}$$

The following lemma connects the Choquet functional of the random set Ξ with the probability generating functional G_Ψ (recall Definition 1.11) of the unmarked point process Ψ.

Lemma 4.1 (Lemma 1 in Flimmel and Heinrich [2021+]).
For any $X \in \mathcal{C}^{(2)}$, we have

$$T_\Xi(X) = 1 - G_\Psi\big[1 - \mathbb{P}\big((\cdot) \in [-R_0, R_0] \oplus \langle v(\Phi_0), X\rangle\big)\big], \tag{4.8}$$

where $\langle v(\Phi_0), X\rangle := \bigcup_{x \in X} \langle v(\Phi_0), x\rangle$ with the scalar product $\langle \cdot, \cdot \rangle$ in $\mathbb{R}^2$ and $v(\Phi_0) = (\cos\Phi_0, \sin\Phi_0)$ is the normal vector.

Proof. To prove formula (4.8), we need the orthogonal matrix

$$\mathcal{O}(\varphi) = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}, \tag{4.9}$$

which represents an anti-clockwise rotation by the angle $\varphi \in [0, \pi)$ so that $\mathcal{O}(-\varphi)v(\varphi) = (1,0)^T$ and $\mathcal{O}(\varphi)(1,0)^T = v(\varphi)$. Note that

$$\mathcal{O}(-\varphi) = \mathcal{O}^T(\varphi) = \mathcal{O}^{-1}(\varphi).$$

By the definition of the probability generating functional of Ξ and the independence assumption in Definition 1.33, we obtain

$$\begin{aligned}
T_\Xi(X) &= 1 - \mathbb{P}\big(\Xi \cap X = \varnothing\big) \\
&= 1 - \mathbb{P}\left(\bigcap_{i:P_i \in \Psi} \{(g(P_i, \Phi_i) \oplus b(\mathbf{o}, R_i)) \cap X = \varnothing\}\right) \\
&= 1 - \mathbb{E} \prod_{i:P_i \in \Psi} \mathbf{1}\big(\{(g(P_i, \Phi_i) \oplus b(\mathbf{o}, R_i)) \cap X = \varnothing\}\big) \\
&= 1 - \int_{\mathbf{N}} \mathbb{E}\left[\prod_{i:p_i \in \psi} (1 - \mathbb{P}((g(p_i, \Phi_i) \oplus b(\mathbf{o}, R_i)) \cap X = \varnothing)) | \Psi = \psi = \{p_i\}\right] P_\Psi(\mathrm{d}\psi) \\
&= 1 - \int_{\mathbf{N}} \prod_{i:p_i \in \psi} \big(1 - \mathbb{P}\big((g(p_i, \Phi_i) \oplus b(\mathbf{o}, R_i)) \cap X = \varnothing\big)\big) P_\Psi(\mathrm{d}\psi) \\
&= 1 - \int_{\mathbf{N}} \prod_{i:p_i \in \psi} \big(1 - \mathbb{P}\big(p_i \in [-R_0, R_0] \oplus \langle v(\Phi_0), X\rangle\big)\big) P_\Psi(\mathrm{d}\psi). \qquad (4.10)
\end{aligned}$$

Obviously, (4.10) coincides with (4.8). The relation (4.10) is seen as follows:

$$\begin{aligned}
\Big\{(g(p, \Phi_0) \oplus b(\mathbf{o}, R_0)) \cap X \neq \varnothing\Big\} &= \Big\{pv(\Phi_0) \in \big(-g(0, \Phi_0) \oplus b(\mathbf{o}, R_0)\big) \oplus X\Big\} \\
&= \Big\{pv(\Phi_0) \in \big(g(0, \Phi_0) \oplus b(\mathbf{o}, R_0)\big) \oplus X\Big\} \\
&= \Big\{p\mathcal{O}(-\Phi_0)v(\Phi_0) \in \big(g(0,0) \oplus b(\mathbf{o}, R_0)\big) \oplus \mathcal{O}(-\Phi_0)X\Big\} \\
&= \Big\{p(1,0)^T \in \big(g(0,0) \oplus b(\mathbf{o}, R_0)\big) \oplus \mathcal{O}(-\Phi_0)X\Big\} \\
&= \Big\{p \in [-R_0, R_0] \oplus \langle v(\Phi_0), X\rangle\Big\}.
\end{aligned}$$

Hence, the proof of (4.8) is complete.

□

Corollary 4.1 (Corollary 1 in Flimmel and Heinrich [2021+]).
For $X = \{x_1, \dots, x_k\}$ with distinct points $x_1 \dots, x_k \in \mathbb{R}^2$, we get the relation

$$T_{\Xi^c}(\{x_1, \dots, x_k\}) = G_\Psi\left[1 - \mathbb{P}\left((\cdot) \in \bigcup_{i=1}^k ([-R_0, R_0] \oplus \langle v(\Phi_0), x_i\rangle)\right)\right].$$

Example 4.2.
For a stationary Poisson process η with intensity λ, we have by (1.7) that

$$\begin{aligned} T_\Xi(X) &= 1 - \exp\left\{ -\lambda \int_{\mathbb{R}} \mathbb{P}\left((g(p,\Phi_0) \oplus b(\mathbf{o}, R_0)) \cap X \neq \varnothing \right) \mathrm{d}p \right\} \\ &= 1 - \exp\left\{ -\lambda \mathbb{E} \left| [-R_0, R_0] \oplus \langle v(\Phi_0), X \rangle \right|_1 \right\} \\ &= 1 - \exp\left\{ -\lambda \int_0^\infty \int_0^\pi \left| [-r, r] \oplus \langle v(\varphi), X \rangle \right|_1 \mathrm{d}G(\varphi) \mathrm{d}F(r) \right\}. \end{aligned}$$

In the special case $X = \{x_1, \dots, x_k\}$ such that $x_i = (x_i^{(1)}, x_i^{(2)})^T$, it follows from Corollary 4.1 that

$$\begin{aligned} T_\Xi(\{x_1, \dots, x_k\}) &= 1 - \exp\left\{ -\lambda \mathbb{E} \left| \bigcup_{i=1}^k ([-R_0, R_0] + x_i^{(1)} \cos \Phi_0 + x_i^{(2)} \sin \Phi_0) \right|_1 \right\} \\ &= 1 - \exp\left\{ -\lambda \int_0^\pi \int_0^\infty \left| \bigcup_{i=1}^k ([-r, r] + x_i^{(1)} \cos \varphi + x_i^{(2)} \sin \varphi) \right|_1 \mathrm{d}F(r) \mathrm{d}G(\varphi) \right\}. \end{aligned}$$

Next, we use Lemma 4.1 to link the cumulants of $|\Xi \cap \rho K|_2$ directly with the probability generating function of Ψ and hence by Theorem 1.4 and Theorem 1.6 with its factorial moment and factorial cumulant measures, respectively. Using the k-th moment function (4.6), one can express the k-th moment of $|\Xi \cap \rho K|_2$ as

$$\mathbb{E}\, |\Xi \cap \rho K|_2^k = \mathbb{E} \int_{(\rho K)^k} \prod_{i=1}^k \mathbf{1}_\Xi(x_i) \mathrm{d}(x_1, \dots, x_k) = \int_{(\rho K)^k} p_\Xi^{(k)}(x_1, \dots, x_k) \mathrm{d}(x_1, \dots, x_k).$$

Next, we define the *k-th cumulant function* by

$$c_\Xi^{(k)}(x_1, \dots, x_k) := \sum_{\ell=1}^k (-1)^{\ell-1} (\ell-1)! \sum_{\substack{K_1 \cup \dots \cup K_\ell \\ = \{1, \dots, k\}}} \prod_{j=1}^\ell p_\Xi^{(\# K_j)}(x_i : i \in K_j), \qquad (4.11)$$

for $x_1, \dots, x_k \in \mathbb{R}^2$. Note that the cumulant functions of the second and higher orders satisfy the identity $c_\Xi^{(k)}(x_1, \dots, x_k) = (-1)^k c_{\Xi^c}^{(k)}(x_1, \dots, x_k)$ and moreover,

$$\begin{aligned} \mathbf{Cum}_k(|\Xi \cap \rho K|_2) : &= \int_{(\rho K)^k} c_\Xi^{(k)}(x_1, \dots, x_k) \mathrm{d}(x_1, \dots, x_k) \\ &= (-1)^k \int_{(\rho K)^k} c_{\Xi^c}^{(k)}(x_1, \dots, x_k) \mathrm{d}(x_1, \dots, x_k). \end{aligned} \qquad (4.12)$$

Then, by applying Fubini's theorem together with Corollary 4.1, we find that

$$\begin{aligned} \mathbb{E}\, |\Xi \cap \rho K|_2 &= \mathbb{E} \int_{\mathbb{R}^2} \mathbf{1}_\Xi(x) \mathbf{1}_{\rho K}(x) \mathrm{d}x = \int_{\rho K} p_\Xi^{(1)}(x) \mathrm{d}x = \rho^2 \int_K T_\Xi(\{\rho x\}) \mathrm{d}x \\ &= \rho^2 \int_K \left(1 - G_\Psi\left[1 - \mathbb{P}\left((\cdot) \in [-R_0, R_0] \oplus \rho \langle v(\Phi_0), x \rangle\right)\right]\right) \mathrm{d}x. \end{aligned} \qquad (4.13)$$

Moreover, it is easily seen that

$$p_{\Xi}^{(2)}(x_1,x_2) - p_{\Xi}^{(1)}(x_1)p_{\Xi}^{(1)}(x_2) = p_{\Xi^c}^{(2)}(x_1,x_2) - p_{\Xi^c}^{(1)}(x_1)p_{\Xi^c}^{(1)}(x_2). \tag{4.14}$$

So, for the variance (the second cumulant), we get from (4.14) and Corollary 4.1 that

$$\begin{aligned}\operatorname{Var}\big(|\Xi\cap\rho K|_2\big) &= \int\limits_{K\rho}\int\limits_{K\rho}\big(p_{\Xi^c}^{(2)}(x_1,x_2) - p_{\Xi^c}^{(1)}(x_1)p_{\Xi^c}^{(1)}(x_2)\big)\mathrm{d}x_1\mathrm{d}x_2\\ &= \int\limits_{K\rho}\int\limits_{K\rho}\Big(G_\Psi\big[1-\mathbb{P}\big((\cdot)\in\bigcup_{i=1}^{2}([-R_0,R_0]\oplus\langle v(\Phi_0),x_i\rangle)\big)\big]\\ &\qquad -\prod_{i=1}^{2}G_\Psi\big[1-\mathbb{P}\big((\cdot)\in[-R_0,R_0]\oplus\langle v(\Phi_0),x_i\rangle\big)\big]\Big)\mathrm{d}x_1\mathrm{d}x_2.\end{aligned} \tag{4.15}$$

Notation. Formula (4.15) can be generalized to higher-order cumulants. Before that, to simplify the notation, we define for $k\geqslant 2$ points $x_1,\dots,x_k\in\mathbb{R}^2$, $\Xi_0=[-R_0,R_0]$ and $v(\varphi)=(\cos\varphi,\sin\varphi)^T$ the functions

$$w_{x_1,\dots,x_k}^{\cup}(p) := \mathbb{P}\Big(p\in\bigcup_{i=1}^{k}(\Xi_0+\langle v(\Phi_0),x_i\rangle)\Big)$$

and

$$w_{x_1,\dots,x_k}^{\cap}(p) := \mathbb{P}\Big(p\in\bigcap_{i=1}^{k}(\Xi_0+\langle v(\Phi_0),x_i\rangle)\Big).$$

For $k=1$ we put $w_x(p):=w_x^{\cup}(p)=w_x^{\cap}(p)$. Obviously, $w_{x_1,x_2}^{\cup}(p)=w_{x_1}(p)+w_{x_2}(p)-w_{x_1,x_2}^{\cap}(p)$.

By Lemma 4.1, the k-th order moment functions $p_{\Xi^c}^{(k)}$ of the $0-1$-random field $\xi(x):=\mathbf{1}_{\Xi^c}(x)$ can be expressed for any $k\geqslant 1$ by:

$$p_{\Xi^c}^{(k)}(x_1,\dots,x_k) = \mathbb{E}\Big(\prod_{j=1}^{k}\xi(x_j)\Big) = \mathbb{P}(\{x_1,\dots,x_k\}\cap\Xi=\varnothing) = G_\Psi\big[1-w_{x_1,\dots,x_k}^{\cup}(\cdot)\big].$$

Then, together with (4.11) and (4.12), we conclude that

$$\begin{aligned}&\mathbf{Cum}_k(|\Xi\cap\rho K|_2)\\ &= (-1)^k\int\limits_{(\rho K)^k}\sum_{l=1}^{k}(-1)^{l-1}(l-1)!\sum_{\substack{K_1\cup\dots\cup K_\ell\\=\{1,\dots,k\}}}\prod_{j=1}^{l}p_{\Xi^c}^{(\#K_j)}(x_s;s\in K_j)\mathrm{d}(x_1,\dots,x_k)\\ &= (-1)^k\sum_{l=1}^{k}(-1)^{l-1}(l-1)!\sum_{\substack{K_1\cup\dots\cup K_\ell\\=\{1,\dots,k\}}}\int\limits_{(\rho K)^k}\prod_{j=1}^{l}G_\Psi(1-w_{(x_s;s\in K_j)}^{\cup}(\cdot))\mathrm{d}(x_1,\dots,x_k).\end{aligned} \tag{4.16}$$

In order to treat the moments and cumulants of $|\Xi \cap \rho K|_2$, the following relations are useful. Let $a_1, a_2, \ldots$ be real numbers in $[0,1]$. Then we have

$$1 - \prod_{i=1}^{n}(1-a_i) = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k!} \sum_{1 \leqslant i_1,\ldots,i_k \leqslant n}^{\neq} a_{i_1} \cdot \ldots \cdot a_{i_k} \quad \text{for } n \geqslant 1.$$

Moreover, for any odd number $m < n$ (provided $n \geqslant 2$), the so-called *Bonferroni inequalities* (see e.g. Galambos and Simonelli [1996]) hold:

$$\sum_{k=1}^{m+1} \frac{(-1)^{k-1}}{k!} \sum_{1 \leqslant i_1,\ldots,i_k \leqslant n}^{\neq} a_{i_1} \cdot \ldots \cdot a_{i_k} \leqslant 1 - \prod_{i=1}^{n}(1-a_i) \leqslant \sum_{k=1}^{m} \frac{(-1)^{k-1}}{k!} \sum_{1 \leqslant i_1,\ldots,i_k \leqslant n}^{\neq} a_{i_1} \cdot \ldots \cdot a_{i_k}. \tag{4.17}$$

As a consequence of (2.17) and (4.17) and the definition of the factorial moment measures $\alpha^{[k]}(\cdot)$ of $\Psi \sim P$, we get the following series expansion

$$\begin{aligned} \mathbb{E}\,|\Xi \cap \rho K|_2 = & \int_{\rho K} \big(1 - G_\Psi\big[1 - w_x(\cdot)\big]\big)\mathrm{d}x = \rho^2 \int_K \big(1 - G_\Psi\big[1 - w_{\rho x}(\cdot)\big]\big)\mathrm{d}x \\ = & \rho^2 \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \int_K \int_{\mathbb{R}^k} \prod_{j=1}^{k} w_{\rho x}(p_j)\alpha^{[k]}(\mathrm{d}(p_1,\ldots,p_k))\mathrm{d}x, \end{aligned} \tag{4.18}$$

provided that the infinite sum on the right-hand side converges. From (4.17), we obtain immediately the estimates

$$\begin{aligned} \Bigg| 1 - G_\Psi[1 - w_{\rho x}(\cdot)] - & \sum_{k=1}^{m-1} \frac{(-1)^{k-1}}{k!} \int_{\mathbb{R}^k} \prod_{j=1}^{k} w_{\rho x}(p_j)\alpha^{[k]}(\mathrm{d}(p_1,\ldots,p_k)) \Bigg| \\ & \leqslant \frac{1}{m!} \int_{\mathbb{R}^m} \prod_{j=1}^{m} w_{\rho x}(p_j)\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m)) \end{aligned} \tag{4.19}$$

for any $m \geqslant 1$. It is easily seen that the right-hand side of (4.18) is convergent if and only if

$$\frac{1}{m!} \int_{\mathbb{R}^m} \prod_{j=1}^{m} w_{\rho x}(p_j)\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m)) \xrightarrow[m\to\infty]{} 0. \tag{4.20}$$

At this place, we specify the point processes of interest. We define a stronger version of Brillinger-mixing property (recall Definition 1.28). If $\gamma_{red}^{[k]}(\cdot)$ possesses a Lebesgue density $c_{red}^{(k)}(\cdot)$ on $\mathbb{R}^{k-1}$ (called the *k-th reduced cumulant density*), we define the canonical L_q-norm $\|c_{red}^{(k)}\|_q := \big(\int_{\mathbb{R}^{k-1}} |c_{red}^{(k)}(x)|^q \mathrm{d}x\big)^{1/q}$ for $k \geqslant 2$ and the modified L_q^*-norm $\|c_{red}^{(k)}\|_q^* := \int_{\mathbb{R}^1} \big(\int_{\mathbb{R}^{k-2}} |c_{red}^{(k)}(x,p)|^q \mathrm{d}x\big)^{1/q}\mathrm{d}p$ for $k \geqslant 3$, where $1 \leqslant q < \infty$.

Definition 4.1 (Strongly (L_q-, L_q^*-)Brillinger-mixing point process). A stationary point process Ψ on $[\mathbb{R}^1, \mathcal{B}(\mathbb{R}^1)]$ with intensity $\lambda = \mathbb{E}\,\Psi([0,1]) > 0$ satisfying $\mathbb{E}\,\Psi^k([0,1]) < \infty$ for all $k \geqslant 2$, is said to be strongly Brillinger-mixing (strongly L_q-Brillinger-mixing, resp. strongly L_q^*-Brillinger-mixing for some $q \geqslant 1$) if there are constants $b > 0$ and $a \geqslant b^{-1}$ such that $\|\gamma_{red}^{[k]}\|_{TV} \leqslant ab^k k!$ (if $c_{red}^{(k)}(\cdot)$ exists such that $\|c_{red}^{(2)}\|_1 < \infty$ and $\|c_{red}^{(k)}\|_q \leqslant a_q(b_q)^k k!$ for $k \geqslant 2$ with constants $a_q, b_q > 0$ resp. $\|c_{red}^{(k)}\|_q^* \leqslant a_q^*(b_q^*)^k k!$ for $k \geqslant 3$ with constants $a_q^*, b_q^* > 0$).

Remark. For formal reason we put $\|\gamma_{red}^{[1]}\|_{TV} := 1$ and $\|c_{red}^{(1)}\|_q := 1$ so that $a \geqslant b^{-1}$ makes $\|\gamma_{red}^{[1]}\|_{TV} := 1 \leqslant ab$. Further, note that the existence and integrability of $c_{red}^{(k)}(\cdot)$ imply that $\|c_{red}^{(2)}\|_1 = \|\gamma_{red}^{[2]}\|_{TV}$ and $\|c_{red}^{(k)}\|_1 = \|c_{red}^{(k)}\|_1^* = \|\gamma_{red}^{[k]}\|_{TV}$ for all $k \geqslant 3$.

In Heinrich and Pawlas [2013] and Heinrich [2021+], the relations between (strong) Brillinger-mixing and classical mixing conditions are studied. Strong Brillinger-mixing property requires exponential moments of the number of points in bounded sets. For any dimension $d \geqslant 1$, examples of such point processes are DPP's (Example 1.10), Poisson cluster processes (Example 1.7) if the number of daughter points has an exponential moment and certain Cox processes (Example 1.6) as well as Gibbsian processes (Example 1.11) under suitable restrictions, see Ruelle [1969]. For $d = 1$, renewal processes with an exponentially decaying interrenewal density, see Heinrich and Schmidt [1985], among them the Erlang process and the Macchi process, see Daley and Vere-Jones [2003] (p. 144), are strongly Brillinger-mixing.

Now, one way to show (4.20) consists of expressing $\alpha^{[m]}(\cdot)$ by factorial cumulant measures $\gamma^{[k]}(\cdot), k = 1, \dots, m$ as in (1.5) and assuming that Ψ is strongly Brillinger-mixing (strongly L_q-Brillinger-mixing, resp. strongly L_q^*-Brillinger-mixing).

Lemma 4.2 (Lemma 3 in Flimmel and Heinrich [2021+]).
If the stationary point process $\Psi \sim P$ is strongly Brillinger-mixing with $b < \frac{1}{2}$ and $\mathbb{E}\, R_0 < \infty$, then

$$\sum_{m=1}^{\infty} \frac{1}{m!} \int_{\mathbb{R}^m} \prod_{j=1}^{m} w_{\rho x}(p_j) \alpha^{[m]}(\mathrm{d}(p_1, \dots, p_m)) \leqslant \frac{b}{1-2b}\big(\exp\{a\lambda \mathbb{E}\,|\Xi_0|_1\} - 1\big), \quad (4.21)$$

which immediately implies (4.20). If $\Psi \sim P$ is strongly L_q-Brillinger-mixing for some $q > 1$ such that $b_q < \frac{1}{2}(\mathbb{E}\,|\Xi_0|)^{\frac{1}{q}-1}$, then the estimate (4.21) remains valid with a and b replaced by $a_q(\mathbb{E}\,|\Xi_0|)^{\frac{1}{q}-1}$ and $b_q(\mathbb{E}\,|\Xi_0|)^{1-\frac{1}{q}}$, respectively.

Proof. Using the representation (1.5), we obtain

$$\begin{aligned}
&\frac{1}{m!} \int_{\mathbb{R}^m} \prod_{j=1}^{m} w_{\rho x}(p_j) \alpha^{[m]}(\mathrm{d}(p_1, \dots, p_m)) \\
&= \frac{1}{m!} \sum_{\ell=1}^{m} \sum_{\substack{K_1 \cup \dots \cup K_\ell \\ = \{1,\dots,m\}}} \prod_{j=1}^{\ell} \int_{\mathbb{R}^{\#K_j}} \prod_{i \in K_j} w_{\rho x}(p_i) \gamma^{[\#K_j]}(\mathrm{d}(p_i : i \in K_j)) \\
&= \frac{1}{m!} \sum_{\ell=1}^{m} \frac{1}{\ell!} \sum_{\substack{k_1+\dots+k_\ell=m \\ k_i \geqslant 1, i=1\dots,\ell}} \frac{m!}{k_1! \cdots k_\ell!} \prod_{j=1}^{\ell} f(k_j) = \sum_{\ell=1}^{m} \frac{1}{\ell!} \sum_{\substack{k_1+\dots+k_\ell=m \\ k_i \geqslant 1, i=1\dots,\ell}} \prod_{i=1}^{\ell} \frac{f(k_i)}{k_i!}, \quad (4.22)
\end{aligned}$$

where

$$\begin{aligned}
f(k) &:= \int_{\mathbb{R}^k} \prod_{i=1}^{k} w_{\rho x}(p_i) \gamma^{[k]}(\mathrm{d}(p_1, \dots, p_k)) \\
&= \lambda \int_{\mathbb{R}^1} w_{\rho x}(p_1) \int_{\mathbb{R}^{k-1}} \prod_{i=2}^{k} w_{\rho x}(p_i + p_1) \gamma_{red}^{[k]}(\mathrm{d}(p_2, \dots, p_k)) \mathrm{d}p_1
\end{aligned}$$

for $k = 1, \ldots, m$. The equality (4.22) is justified by the invariance of $\gamma^{[k]}(\times_{i=1}^k B_i)$ against permutations of the bounded sets $B_1, \ldots, B_k \in \mathcal{B}(\mathbb{R}^1)$ for each $k \in \mathbb{N}$. We proceed with

$$\begin{aligned}|f(k)| &\leqslant \lambda \int\limits_{\mathbb{R}^1} w_{\rho x}(p_1) \int\limits_{\mathbb{R}^{k-1}} |\gamma_{red}^{[k]}|(\mathrm{d}(p_2, \ldots, p_k))\mathrm{d}p_1 \\ &= \lambda \mathbb{E}\,|\Xi_0|_1 \|\gamma_{red}^{[k]}\|_{TV} \\ &\leqslant a\lambda \mathbb{E}\,|\Xi_0|_1 b^k k!.\end{aligned}$$

Here, we have used Fubini's theorem combined with $w_{\rho x}(p) \leqslant 1$ for $p \in \mathbb{R}^1$ and $x \in \mathbb{R}^2$ so that

$$\int\limits_{\mathbb{R}^1} w_{\rho x}(p)\mathrm{d}p = \int\limits_{\mathbb{R}^1} \mathbb{P}\big(p \in \Xi_0 + \rho\langle v(\Phi_0), x\rangle\big)\mathrm{d}p = \int\limits_{\mathbb{R}^1} \mathbb{P}\big(p \in \Xi_0\big)\mathrm{d}p = \mathbb{E}\,|\Xi_0|_1.$$

Hence, together with the combinatorial relations

$$\sum_{\substack{k_1+\cdots+k_\ell=m \\ k_i \geqslant 1, i=1\ldots,\ell}} 1 = \binom{m-1}{\ell-1} \quad \text{and} \quad \sum_{\ell=1}^{m} \binom{m-1}{\ell-1} = 2^{m-1}$$

we arrive at

$$\begin{aligned}\sum_{\ell=1}^{m} \frac{1}{\ell!} \sum_{\substack{k_1+\cdots+k_\ell=m \\ k_i \geqslant 1, i=1\ldots,\ell}} \prod_{i=1}^{\ell} \frac{|f(k_i)|}{k_i!} &\leqslant b^m \sum_{\ell=1}^{m} \frac{(a\lambda \mathbb{E}\,|\Xi_0|_1)^\ell}{\ell!} \binom{m-1}{\ell-1} \\ &\leqslant b^m 2^{m-1} \max_{1 \leqslant \ell \leqslant m} \frac{(a\lambda \mathbb{E}\,|\Xi_0|_1)^\ell}{\ell!} \\ &\leqslant \frac{1}{2}\big(\exp\{a\lambda \mathbb{E}\,|\Xi_0|_1\} - 1\big)(2b)^m. \qquad (4.23)\end{aligned}$$

By combining (4.22) and (4.23) with $b < 1/2$ the relation (4.21) follows immediately. Under the strong L_q-Brillinger-mixing condition we may rewrite $f(k)$ for $k \geqslant 2$ as follows:

$$f(k) = \lambda \int\limits_{\mathbb{R}^1} w_{\rho x}(p_1) \mathbb{E} \int\limits_{\mathbb{R}^{k-1}} \prod_{i=2}^{k} \mathbf{1}_{\Xi_i + \rho\langle v(\Phi_i), x\rangle}(p_i + p_1) c_{red}^{(k)}(p_2, \ldots, p_k)\mathrm{d}(p_2, \ldots, p_k)\mathrm{d}p_1,$$

where $\Xi_i = [-R_i, R_i]$ and $(R_2, \Phi_2), \ldots, (R_k, \Phi_k)$ are i.i.d. random vectors with same distribution as (R_0, Φ_0). Applying Hölder's inequality for $q > 1$ and $p = q/(q-1)$, Lyapunov's inequality $\mathbb{E}\,|\Xi_0|^{\frac{1}{p}} \leqslant (\mathbb{E}\,|\Xi_0|)^{\frac{1}{p}} = (\mathbb{E}\,|\Xi_0|)^{1-\frac{1}{q}}$ and the condition $\|c_{red}^{(k)}\|_q \leqslant a_q b_q^k k!$, we obtain that

$$\begin{aligned}|f(k)| &\leqslant \lambda \|c_{red}^{(k)}\|_q \mathbb{E}\,|\Xi_1|_1 \prod_{i=2}^{k} \mathbb{E}\,|\Xi_i|_1^{\frac{1}{p}} \\ &\leqslant \lambda \|c_{red}^{(k)}\|_q (\mathbb{E}\,|\Xi_0|_1)^{1+\frac{k-1}{p}} \\ &\leqslant \lambda a_q (\mathbb{E}\,|\Xi_0|_1)^{\frac{1}{q}} (b_p (\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}})^k k!.\end{aligned}$$

By repeating the foregoing steps with the latter bound the proof of Lemma 3 is finished.

□

Further, we can use the representation of $\mathbf{Cum}_k(|\Xi \cap \rho K|_2)$ (4.16) together with the expansion of the probability generating functional (4.19). Under the assumption of the strong type of Brillinger-mixing property of Ψ, we are ready to determine the asymptotic variance of $|\Xi \cap \rho K|_2$ and central limit theorem based on Lemma 4.2.

4.3 Asymptotic properties of a planar cylinder process

Asymptotic properties of cylinder processes in expanding domains using the cumulant method were studied under the Poisson setting in Heinrich and Spiess [2009], Heinrich and Spiess [2013]. Some of the results were then generalized for cylinder processes constructed from a strong Brillinger-mixing point process in Flimmel and Heinrich [2021+].

In order to prove the asymptotic normality of $\rho^{-3/2}\big(|\Xi \cap \rho K|_2 - \mathbb{E}\,|\Xi \cap \rho K|_2\big)$ as $\rho \to \infty$ one has to show that

$$\rho^{-3k/2}\mathbf{Cum}_k(|\Xi \cap \rho K|_2) \xrightarrow[\rho\to\infty]{} 0, \quad \text{for any } k \geqslant 3. \tag{4.24}$$

Indeed, (4.24) is sufficient by the homogeneity (4.5) and shift invariance (4.4) of the cumulants of the second or higher orders. Then by the method of cumulants (see the text after Example 4.1), we have that

$$\rho^{-3/2}\big(|\Xi \cap K_\rho|_2 - \mathbb{E}\,|\Xi \cap K_\rho|_2\big) \xrightarrow{\mathcal{D}} N(0, \sigma_P^2(K, F, G)),$$

where

$$\sigma_P^2(K, F, G) := \lim_{\rho\to\infty} \frac{\operatorname{Var}|\Xi \cap \rho K|_2}{\rho^3}. \tag{4.25}$$

denotes the asymptotic variance, if the limit exists. Note that the order ρ^3 of the growth of $\operatorname{Var}(|\Xi \cap \rho K|_2)$ is much faster than the growth of the area $|\rho K|_2 = \rho^2 |K|_2$ which reveals a typical feature of long-range dependencies within the random set Ξ.

Poisson setting

First, we shall briefly mention the results under the Poisson setting as the starting point for a more detailed analysis of cylinder processes driven by strong Brillinger-mixing point processes.

The results are formulated in Heinrich and Spiess [2013] for a general dimension. The cylinder process in $\mathbb{R}^d$ is constructed as a union of k-flats $(0 \leqslant k \leqslant d-1)$ which are dilated by an independent and identically distributed random compact cylinder base taken from the corresponding $(d-k)$-dimensional orthogonal complement. More precisely, it can be introduced as follows. By $\mathbb{G}(d, k)$ we denote the Grassmannian of k-dimensional subspaces of $\mathbb{R}^d$, where $k = 0, \ldots, d-1$. For

$L \in \mathbb{G}(d,k)$ (*direction space*) and a set B in the orthogonal complement $L^\perp$ (*base*), we define the corresponding cylinder as $L \oplus B$.

Let $\{e_1,\ldots,e_d\}$ be the usual orthonormal basis of $\mathbb{R}^d$ and E_k:=span$\{e_1,\ldots,e_k\}$, $E_k^\perp :=$ span$\{e_{d-k+1},\ldots,e_d\}$ orthogonal subspaces, $k \in \{0,\ldots,d-1\}$. Then for a given $L \in \mathbb{G}(d,k)$, there exists an equivalence class of orthogonal matrices $\mathcal{O} \in \mathbb{R}^{d\times d}$ with det $\mathcal{O} = 1$ and $\mathcal{O}E_k = L$. Further, each equivalence class can be identified with a single representative $\mathcal{O}_L$. It follows from a fact from differential geometry that the dimension $\dim(\mathbb{G}(d,k)) = (d-k)k$ implies that there exists a parametric representation of the matrices $\mathcal{O}_L$ over some subset of $\mathbb{R}^{(d-k)k}$. In particular, if $d = 2$ and $k = 1$, a suitable representation is the one in (4.9). At last, we denote by $\mathbb{SO}_k^d$ the family of all such $\mathcal{O}_L$. In this way, each random subspace $L \in \mathbb{G}(d,k)$ corresponds with a unique random matrix $\Theta(L) \in \mathbb{SO}_k^d$.

Assume that $\eta_m = \sum_{i\geqslant 1} \delta_{[P_i,(\Theta_i,\Xi_i)]}$ is a stationary, independently marked Poisson point process on $\mathbb{R}^d$ with intensity $\lambda > 0$ and marks with values in $\mathbb{SO}_k^d \times \mathcal{C}^{d-k}$, where $\mathcal{C}^j$ denoted the space of all compact subsets of $\mathbb{R}^j$, $j \in \mathbb{N}$. Similarly as in the Example 1.15, the cylinder process is defined as a stationary random set

$$\Xi_{\lambda,k} = \bigcup_{i\geqslant 1} \Theta_i((\Xi_i + P_i) \times \mathbb{R}^k). \tag{4.26}$$

Remark. When $k = 0$, the union set (4.26) coincides with the Boolean model (see Example 1.14).

The following results give the asymptotic normality of the random variable $|\Xi_{\lambda,k} \cap \rho K|_d$ with increasing ρ, where K is assumed to be a compact star-shaped (w.r.t the origin $\mathbf{o} \in \mathbb{R}^d$) subset of $\mathbb{R}^d$ such that $B(\mathbf{o},\varepsilon_K) \subset K \subset B(\mathbf{o},1)$ for some $\varepsilon_K \in (0,1]$.

Theorem 4.1 (Theorem 1 in Heinrich and Spiess [2013]).
Assume that the typical cylinder base Ξ_0 is a.s. compact and $0 < \mathbb{E}\,|\Xi_0|_{d-k}^2 < \infty$. Then

$$\frac{|\Xi_{\lambda,k} \cap \rho K|_d - \rho^d |K|_d (1 - e^{-\lambda \mathbb{E}\,|\Xi_0|_{d-k}})}{\sqrt{\mathrm{Var}\,(|\Xi_{\lambda,k} \cap \rho K|_d)}} \xrightarrow[\rho\to\infty]{\mathcal{D}} N(0,1).$$

Moreover, by Lemma 1 of Heinrich and Spiess [2009], the variance of $|\Xi_{\lambda,k} \cap \rho K|_d$ increases proportionally to the $(d+k)$-th power of ρ, i.e. there are constants c_1, c_2 not depending on ρ such that

$$c_1 \rho^{d+k} \leqslant \mathrm{Var}\,(|\Xi_{\lambda,k} \cap \rho K|_d) \leqslant c_2 \rho^{d+k}, \qquad \text{for all } \rho \geqslant 1.$$

The exact form of the asymptotic variance

$$\sigma_{\lambda,k}^2(K) := \lim_{\rho\to\infty} \frac{\mathrm{Var}\,(|\Xi_{\lambda,k} \cap \rho K|_d)}{\rho^{d+k}}$$

is given by the following theorem.

Theorem 4.2 (Theorem 2 in Heinrich and Spiess [2013]).
Let the assumptions of Theorem 4.1 hold. Then,

(a) *if* Θ_0 *is discretely distributed on some at most countable set* $\{\theta_i \in \mathbb{SO}_k^d, i \in I\}$, *then*

$$\sigma^2_{\lambda,k}(K) = e^{-2\lambda\mathbb{E}\,|\Xi_0|_{d-k}} \sum_{i\in I} I_k(\theta_i^\top K) \int\limits_{\mathbb{R}^{d-k}} \left(e^{\lambda f(y,\theta_i)} - 1\right) \mathrm{d}y,$$

where $f(y,\theta_i) := \mathbb{E}\left(|\Xi_0 \cap (\Xi_0 - y)|_{d-k} \mathbf{1}\{\Theta_0 = \theta_i\}\right)$ *for* $i \in I$ *and*

$$I_k(\theta^\top K) := \int\limits_{\mathbb{R}^k} |\theta^\top K \cap (\theta^\top K - (\mathbf{o}, x)^\top)|_d \mathrm{d}x, \qquad \theta \in \mathbb{SO}_k^d,$$

(b) *if* Θ_0 *is continuously distributed, then*

$$\sigma^2_{\lambda,k}(K) = \lambda e^{-2\lambda\mathbb{E}\,|\Xi|_{d-k}} \mathbb{E}\left(I_k(\Theta_0^\top K)|\Xi_0|^2_{d-k}\right).$$

Asymptotic variance in the strong Brillinger-mixing setting

This and the subsequent section is devoted to the application of the method of cumulant in order to generalize Theorem 4.1 and Theorem 4.2 to cylinder processes driven by strong Brillinger-mixing processes. In both sections we admit only the planar case as it was described in Example 1.15.

Under comparatively strong conditions on the higher-order cumulant measures of the unmarked (ground) point process we are able to prove first, a mean-square limit of the relative part of the area of an expanding star-shaped window covered by the union of cylinders, and second, we derive an explicit formula for the asymptotic variance (4.25) of this area using the representations (4.13) and (4.15). The latter is an important first step in proving the asymptotic normality of the covered area which shall follow in the subsequent section.

The limit (4.25) is positive and finite (if $\mathbb{E}\,|\Xi_0|^2 = 4\mathbb{E}\,R_0^2 < \infty$) and depends on the shape of K, the first and second moment of F and the distribution function G which is assumed to be continuous (not necessarily absolutely continuous).

Lemma 4.3 (Lemma 4 in Flimmel and Heinrich [2021+]).
Let $\Psi \sim P$ *be a stationary point process on* $\mathbb{R}^1$ *satisfying* $\max_{2\leqslant k\leqslant m} \|\gamma^{[k]}_{red}\|_{TV} < \infty$ *for some fixed* $m \geqslant 2$. *If* $\mathbb{E}\,R_0 < \infty$ *and* $\Phi_0 \sim G$ *has a continuous distribution function* G *then, for* $m \geqslant 2$ *not necessarily distinct points* $x_1, \ldots, x_m \in \mathbb{R}^2\backslash\{\mathbf{o}\}$,

$$\int\limits_{\mathbb{R}^m} \prod_{j=1}^m w_{\rho x_j}(p_j)\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m)) \xrightarrow[\rho\to\infty]{} \lambda^m \prod_{j=1}^m \int\limits_{\mathbb{R}^1} w_{\rho x_j}(p)\mathrm{d}p = (\lambda\mathbb{E}\,|\Xi_0|_1)^m. \tag{4.27}$$

Remark. A purely discrete distribution function G yields different expressions for the asymptotic variance $\sigma_P^2(K,F,G)$ even if $\Psi = \eta$ is a stationary Poisson point process with intensity $\lambda > 0$ (see Theorem 4.2). A distribution function G without jumps implies that $\mathbb{P}(\Phi_0 = \Phi_1) = 0$ if the angles $\Phi_0, \Phi_1 \sim G$ are independent.

Proof. We use the representation (1.5) for $k = m$ to rewrite the difference of left-hand and right-hand sides of (4.27) as follows:

$$\sum_{\ell=1}^{m-1} \sum_{\substack{K_1 \cup \cdots \cup K_\ell \\ =\{1,\ldots,m\}}} \prod_{j=1}^{\ell} \int_{\mathbb{R}^{\#K_j}} \prod_{i \in K_j} w_{\rho x_i}(p_i) \gamma^{[\#K_j]}(\mathrm{d}(p_i : i \in K_j)).$$

Hence, the limit (4.27) is shown if and only if the finite sum in the latter line disappears as $\rho \to \infty$ and this in turn follows by showing that, for $k = 2, \ldots, m$,

$$\int_{\mathbb{R}^k} \prod_{i=1}^{k} w_{\rho x_i}(p_i) \gamma^{[k]}(\mathrm{d}(p_1, \ldots, p_k))$$
$$= \lambda \int_{\mathbb{R}^k} w_{\rho x_1}(p_1) \prod_{i=2}^{k} w_{\rho x_i}(p_i + p_1) \gamma_{red}^{[k]}(\mathrm{d}(p_2, \ldots, p_k)) \mathrm{d}p_1 \xrightarrow[\rho \to \infty]{} 0.$$

In view of $0 \leqslant w_{\rho x_i}(p_i + p_1) \leqslant 1$ for $i = 3, \ldots, k$ it is sufficient to prove that

$$\int_{\mathbb{R}^{k-1}} \int_{\mathbb{R}^1} \mathbb{P}(p_1 \in \Xi_0 + \rho\langle v(\Phi_0), x_1\rangle) \mathbb{P}(p_1 \in \Xi_0 + \rho\langle v(\Phi_0), x_2\rangle - p_2) \mathrm{d}p_1 |\gamma_{red}^{[k]}|(\mathrm{d}(p_2, \ldots, p_k))$$

asymptotically disappears as $\rho \to \infty$. Since the total variation measure $|\gamma_{red}^{[k]}|(\cdot)$ is bounded on $\mathbb{R}^{k-1}$ and the inner integral over $\mathbb{R}^1$ is less than or equal to $\mathbb{E}\,|\Xi_0|_1$, we have only to verify that the inner integral disappears as $\rho \to \infty$. For this purpose, we rewrite its integrand as expectation $\mathbb{E}\mathbf{1}_{\{\Xi_1 + \rho\langle v(\Phi_1), x_1\rangle\}}(p_1)\mathbf{1}_{\{\Xi_2 + \rho\langle v(\Phi_2), x_2\rangle - p_2\}}(p_1)$, where $\Xi_i = [-R_i, R_i]$ and Φ_i for $i = 1, 2$ have the same distribution as $\Xi_0 = [-R_0, R_0]$ and Φ_0, respectively, and R_1, R_2, Φ_1, Φ_2 are independent of each other. By Fubini's theorem and the shift-invariance of the Lebesgue measure, we arrive at

$$\int_{\mathbb{R}^1} \mathbb{P}(p_1 \in \Xi_0 + \rho\langle v(\Phi_0), x_1\rangle) \mathbb{P}(p_1 \in \Xi_0 + \rho\langle v(\Phi_0), x_2\rangle - p_2) \mathrm{d}p_1$$
$$= \int_{\mathbb{R}^1} \mathbb{E}\,\mathbf{1}_{\{\Xi_1 + \rho\langle v(\Phi_1), x_1\rangle\}}(p_1) \mathbf{1}_{\{\Xi_2 + \rho\langle v(\Phi_2), x_2\rangle - p_2\}}(p_1) \mathrm{d}p_1$$
$$= \mathbb{E}\left|\Xi_1 \cap \left(\Xi_2 - p_2 + \rho(\langle v(\Phi_2), x_2\rangle - \langle v(\Phi_1), x_1\rangle)\right)\right|_1 \xrightarrow[\rho \to \infty]{} 0.$$

The limit in the last line can verified as follows: Let us take two fixed points $x_i = \|x_i\|(\cos(\alpha_i), \sin(\alpha_i)) \in \mathbb{R}^2, i = 1, 2$, and two points $v(\varphi_i) = (\cos(\varphi_i), \sin(\varphi_i))$, $i = 1, 2$, on the unit circle line. It can be easily seen that the equality $\langle v(\varphi_1), x_1\rangle = \langle v(\varphi_2), x_2\rangle$, i.e. $\|x_1\| \cos(\varphi_1 - \alpha_1) = \|x_2\| \cos(\varphi_2 - \alpha_2)$ holds for at most a finite number of pairs $\varphi_1, \varphi_2 \in [0, \pi]$. Hence, for two independent random angles Φ_1, Φ_2 with common atomless distribution function $G(\cdot)$ we have

$$\mathbb{P}(\langle v(\Phi_1), x_1\rangle \neq \langle v(\Phi_2), x_2\rangle) = 1$$

for any two points $x_1, x_2 \in \mathbb{R}^2$ with $\|x_1\| + \|x_2\| > 0$.

□

From Lemma 4.3 and (4.19) we obtain the behaviour of the expectation of $|\Xi \cap \rho K|_2$ as $\rho \to \infty$.

Corollary 4.2 (Corollary 2 in Flimmel and Heinrich [2021+]).
Let $\Psi \sim P$ be a Brillinger-mixing point process on $\mathbb{R}^1$. If $\mathbb{E}\, R_0 < \infty$ and $\Phi_0 \sim G$ has a continuous distribution function G, then

$$\frac{\mathbb{E}\,|\Xi \cap \rho K|_2}{|\rho K|_2} \xrightarrow[\rho\to\infty]{} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} (\lambda \mathbb{E}\,|\Xi_0|_1)^k = 1 - \exp\{-\lambda \mathbb{E}\,|\Xi_0|_1\}.$$

Proof. An application of (4.27) for $x_1 = \cdots = x_m = x \neq \mathbf{o}$ to the inequality (4.19) yields

$$\left| \lim_{\rho\to\infty} \big(1 - G_\Psi[1 - w_{\rho x}(\cdot)]\big) - \sum_{k=1}^{m-1} \frac{(-1)^{k-1}(\lambda \mathbb{E}\,|\Xi_0|_1)^k}{k!} \right| \leqslant \frac{(\lambda \mathbb{E}\,|\Xi_0|_1)^m}{m!} \tag{4.28}$$

for any $m \geqslant 1$. Combining this with (4.18) leads to

$$\frac{\mathbb{E}\,|\Xi \cap \rho K|_2}{|\rho K|_2} = \frac{1}{|K|_2} \int_K \big(1 - G_\Psi[1 - w_{\rho x}(\cdot)]\big) \mathrm{d}x \xrightarrow[\rho\to\infty]{} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}(\lambda \mathbb{E}\,|\Xi_0|_1)^k}{k!}$$

which immediately gives the assertion of Corollary 4.2.

□

The following result can be considered as a planar mean-square ergodic theorem which implies a weak law of large numbers for $|\Xi \cap \rho K|$ in the Euclidean plane $\mathbb{R}^2$.

Theorem 4.3 (Theorem 1 in Flimmel and Heinrich [2021+]).
Assume that the stationary point process Ψ on $\mathbb{R}^1$ is Brillinger-mixing. Further suppose that $\mathbb{E}\, R_0 < \infty$ and $\Phi_0 \sim G$ has a continuous distribution function G. Then

$$\mathbb{E}\left(\frac{|\Xi \cap \rho K|_2}{|\rho K|_2} - \big(1 - \exp\{-\lambda \mathbb{E}\,|\Xi_0|_1\}\big) \right)^2 \xrightarrow[\rho\to\infty]{} 0 \quad \textit{with} \quad \Xi_0 := [-R_0, R_0]. \tag{4.29}$$

Proof. The expectation on the left-hand side of (4.29) can be expressed as follows:

$$\frac{\mathrm{Var}\,(|\Xi \cap \rho K|_2)}{|\rho K|_2^2} + \left(\frac{\mathbb{E}\,|\Xi \cap \rho K|_2}{|\rho K|_2} - \Big(1 - \exp\{-\lambda \mathbb{E}\,|\Xi_0|_1\}\Big) \right)^2.$$

In view of Corollary 4.2 it remains to prove that $\rho^{-4}\mathrm{Var}\,(|\Xi \cap \rho K|_2) \xrightarrow[\rho\to\infty]{} 0$. Using the representation (4.15), we get

$$\begin{aligned}
\rho^{-4}\mathrm{Var}\,(|\Xi \cap \rho K|_2) &= \rho^{-4} \int_{\rho K}\int_{\rho K} \Big(G_\Psi\big[1 - w^{\cup}_{x_1,x_2}(\cdot)\big] - \prod_{i=1}^{2} G_\Psi\big[1 - w_{x_i}(\cdot)\big] \Big) \mathrm{d}x_1 \mathrm{d}x_2 \\
&= \int_K\int_K \Big(G_\Psi\big[1 - w^{\cup}_{\rho x_1,\rho x_2}(\cdot)\big] - G_\Psi\big[1 - w_{\rho x_1}(\cdot)\big] G_\Psi\big[1 - w_{\rho x_2}(\cdot)\big] \Big) \mathrm{d}x_1 \mathrm{d}x_2.
\end{aligned}$$

Thus, we just have to show that the integrand disappears as $\rho \to \infty$ for distinct points $x_1, x_2 \in K \backslash \{\mathbf{o}\}$, that is,

$$G_\Psi\big[1 - w^{\cup}_{\rho x_1, \rho x_2}(\cdot)\big] - G_\Psi\big[1 - w_{\rho x_1}(\cdot)\big] G_\Psi\big[1 - w_{\rho x_2}(\cdot)\big] \xrightarrow[\rho \to \infty]{} 0. \tag{4.30}$$

We make use of the finite expansion (4.19) of the pgf $G_\Psi\big[1 - w_{\rho x}(\cdot)\big]$ with remainder term, where $w_{\rho x}$ can be replaced by any Borel measurable function $w : \mathbb{R}^1 \mapsto [0,1]$. For brevity, we put

$$S_m(w) := \sum_{k=0}^{m-1} (-1)^k T_k(w)$$

with

$$T_0(w) := 1 \text{ and } T_k(w) := \frac{1}{k!} \int\limits_{\mathbb{R}^k} \prod_{j=1}^{k} w(p_j) \alpha^{[k]}(d(p_1, \ldots, p_k))$$

for $1 \leqslant k \leqslant m \in \mathbb{N}$. Hence, (4.19) reads as $\big|G_\Psi\big[1 - w(\cdot)\big] - S_m(w)\big| \leqslant T_m(w)$ which leads us to the following estimate for $m \geqslant 2$:

$$\begin{aligned}
&\Big|G_\Psi\big[1 - w^{\cup}_{\rho x_1, \rho x_2}(\cdot)\big] - G_\Psi\big[1 - w_{\rho x_1}(\cdot)\big] G_\Psi\big[1 - w_{\rho x_2}(\cdot)\big] \\
&\qquad\qquad - \Big(S_m(w^{\cup}_{\rho x_1, \rho x_2}) - S_m(w_{\rho x_1}) S_m(w_{\rho x_2})\Big)\Big| \\
&\leqslant \Big|G_\Psi\big[1 - w^{\cup}_{\rho x_1, \rho x_2}(\cdot)\big] - S_m(w^{\cup}_{\rho x_1, \rho x_2})\Big| \\
&\qquad\qquad + \Big|G_\Psi\big[1 - w_{\rho x_1}(\cdot)\big] - S_m(w_{\rho x_1})\Big| G_\Psi\big[1 - w_{\rho x_2}(\cdot)\big] \\
&\qquad\qquad + \Big|G_\Psi\big[1 - w_{\rho x_2}(\cdot)\big] - S_m(w_{\rho x_2})\Big| S_m(w_{\rho x_1}) \\
&\leqslant T_m(w^{\cup}_{\rho x_1, \rho x_2}) + T_m(w_{\rho x_1}) + T_m(w_{\rho x_2}) + T_m(w_{\rho x_1}) T_m(w_{\rho x_2}).
\end{aligned} \tag{4.31}$$

Here, we have additionally used that $G_\Psi\big[1 - w(\cdot)\big] \leqslant 1$ and $\big|S_m(w)\big| \leqslant G_\Psi\big[1 - w(\cdot)\big] + T_m(w)$.
We are now in a position to apply the limit (4.22) under the assumptions of Lemma 4.3. This yields for $i = 1, 2$ and $m \in \mathbb{N}$

$$T_m(w_{\rho x_i}) \xrightarrow[\rho \to \infty]{} \frac{(\lambda \mathbb{E}\, |\Xi_0|_1)^m}{m!}$$

and

$$S_m(w_{\rho x_i}) \xrightarrow[\rho \to \infty]{} \sum_{k=0}^{m-1} \frac{(-\lambda \mathbb{E}\, |\Xi_0|_1)^k}{k!} = e^{-\lambda \mathbb{E}\, |\Xi_0|_1} + \theta_1 \frac{(\lambda \mathbb{E}\, |\Xi_0|_1)^m}{m!}$$

for some $\theta_1 \in [-1, 1]$ in accordance with $\Big|e^{-x} - \sum_{k=0}^{m-1} \frac{(-x)^k}{k!}\Big| \leqslant \frac{x^m}{m!}$ for any $m \in \mathbb{N}$ and $x \geqslant 0$.
Next, we have to find the limit of $T_m(w^{\cup}_{\rho x_1, \rho x_2})$ as $\rho \to \infty$. Using the relation $w^{\cup}_{x_1, x_2}(p) = w_{x_1}(p) + w_{x_2}(p) - w^{\cap}_{x_1, x_2}(p)$ and taking into account that the factorial moment measure $\alpha^{[m]}$ is invariant under permutation of its m components, we

may write

$$T_m(w^{\cup}_{\rho x_1,\rho x_2}) = \frac{1}{m!}\int\limits_{\mathbb{R}^m}\prod_{j=1}^{m}\big(w_{\rho x_1}(p_j)+w_{\rho x_2}(p_j)-w^{\cap}_{\rho x_1,\rho x_2}(p_j)\big)\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m))$$

$$= \frac{1}{m!}\int\limits_{\mathbb{R}^m}\prod_{j=1}^{m}\big(w_{\rho x_1}(p_j)+w_{\rho x_2}(p_j)\big)\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m)) \tag{4.32}$$

$$+\frac{1}{m!}\sum_{\ell=1}^{m}\binom{m}{\ell}\int\limits_{\mathbb{R}^m}\prod_{i=1}^{\ell}w^{\cap}_{\rho x_1,\rho x_2}(p_i)\prod_{j=\ell+1}^{m}(w_{\rho x_1}(p_j)+w_{\rho x_2}(p_j))\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m)).$$

There is at least one term $w^{\cap}_{\rho x_1,\rho x_2}(p_i) = \mathbb{P}\big(p_i \in (\Xi_0+\rho\langle v(\Phi_0),x_1\rangle)\cap(\Xi_0+\rho\langle v(\Phi_0),x_2\rangle)\big)$ in each summand of the last line which will be integrated over $\mathbb{R}^1$ w.r.t. $\mathrm{d}p_i$ so that after expressing $\alpha^{[m]}$ by cumulant measures, see (1.5), the expectation $\mathbb{E}\,|\Xi_0\cap(\Xi_0+\rho\langle v(\Phi_0),x_2-x_1\rangle)|_1$ emerges and disappears as $\rho\to\infty$ if $x_1\neq x_2$. Thus, the last line disappears completely as $\rho\to\infty$, whereas the line (4.32) converges to the limit $(2\lambda\mathbb{E}\,|\Xi_0|_1)^m/m!$ as $\rho\to\infty$ by applying the limit (4.23) once more. Therefore, we obtain for any $m\in\mathbb{N}$ that $T_m(w^{\cup}_{\rho x_1,\rho x_2})\xrightarrow[\rho\to\infty]{}(2\lambda\mathbb{E}\,|\Xi_0|_1)^m/m!$ and

$$S_m(w^{\cup}_{\rho x_1,\rho x_2})\xrightarrow[\rho\to\infty]{}\sum_{k=0}^{m-1}\frac{(-2\lambda\mathbb{E}\,|\Xi_0|_1)^k}{k!}$$
$$= e^{-2\lambda\mathbb{E}\,|\Xi_0|_1}+\theta_2\frac{(2\lambda\mathbb{E}\,|\Xi_0|_1)^m}{m!}\quad\text{for some } \theta_2\in[-1,1].$$

The latter limit combined with above limits of $S_m(w_{\rho x_i})$ for $i=1,2$ leads to

$$\varlimsup_{\rho\to\infty}\Big|S_m(w^{\cup}_{\rho x_1,\rho x_2})-S_m(w_{\rho x_1})S_m(w_{\rho x_2})\Big|$$
$$\leqslant\frac{(2\lambda\mathbb{E}\,|\Xi_0|_1)^m}{m!}+2\frac{(\lambda\mathbb{E}\,|\Xi_0|_1)^m}{m!}+\frac{(\lambda\mathbb{E}\,|\Xi_0|_1)^{2m}}{(m!)^2}.$$

For any given $\varepsilon\in(0,1]$, we find some $m(\varepsilon)$ such that $\frac{(2\lambda\mathbb{E}\,|\Xi_0|_1)^m}{m!}\leqslant\varepsilon$ for all $m\geqslant m(\varepsilon)$.

Thus, the right-hand side of the last inequality does not exceed $2\varepsilon+\varepsilon^2$ for sufficiently large m. The same bound can be obtained for the limit (as $\rho\to\infty$) of the four summands in line (4.31).

Finally, after summarizing all ε-bounds of the above limiting terms we arrive at

$$\varlimsup_{\rho\to\infty}\Big|G_\Psi\big[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)\big]-G_\Psi\big[1-w_{\rho x_1}(\cdot)\big]G_\Psi\big[1-w_{\rho x_2}(\cdot)\big]\Big|\leqslant 2(2\varepsilon+\varepsilon^2)\leqslant 6\varepsilon.$$

This implies (4.30) completing the proof of Theorem 4.3.

□

Next, we provide the exact asymptotic behaviour of the variance of the area of the cylinder process (1.12) that is contained in a star-shaped set ρK which is growing unboundedly in all directions. For this purpose, in comparison with Theorem 4.3, we need a strengthening and quantification of the classical Brillinger-mixing condition.

Theorem 4.4 (Theorem 2 in Flimmel and Heinrich [2021+]).
Assume that the stationary point process Ψ on $\mathbb{R}^1$ is either strongly Brillinger-mixing with $b < 1/2$ or strongly L_q-Brillinger-mixing with $(\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}} b_q < 1/2$ and strongly L_q^*-Brillinger-mixing with $(\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}} b_q^* < 1/2$ for some $q > 1$, where $\Xi_0 := [-R_0, R_0]$. Further suppose that $\mathbb{E}\, R_0^2 < \infty$ and $\Phi_0 \sim G$ has a continuous distribution function G. Then

$$\sigma_P^2(K, F, G) = \lambda e^{-2\lambda \mathbb{E}\,|\Xi_0|_1} \Big((\mathbb{E}\,|\Xi_0|_1)^2 \gamma_{red}^{[2]}(\mathbb{R}^1) C_1^{G,K} + 2\mathbb{E}\,|\Xi_0|_1^2 C_2^{G,K} \Big), \tag{4.33}$$

where

$$C_1^{G,K} := \int_{\mathbb{R}^1} (\mathbb{E}\,|g(p, \Phi_0) \cap K|_1)^2 \mathrm{d}p$$

and

$$C_2^{G,K} := \int_0^{\pi} \int_0^{r_K(\varphi \pm \pi/2)} |K \cap (K + sv(\varphi \pm \frac{\pi}{2})|_2 \mathrm{d}s \mathrm{d}G(\varphi).$$

Remark. In the special case $K = b(\mathbf{o}, 1)$, one can show that $C_1^{G,K} = \frac{16}{3}$ and $C_2^{G,K} = \frac{8}{3}$ are independent of the distribution function G. If Φ_0 is uniformly distributed on $[0, \pi]$, then we get

$$C_1^{G,K} = \frac{1}{\pi^2} \int_{-\infty}^{\infty} \Big(\int_0^{\pi} |g(p, \varphi) \cap K|_1 \mathrm{d}\varphi \Big)^2 \mathrm{d}p,$$

$$\begin{aligned} C_2^{G,K} &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\infty} |K \cap (K + sv(\varphi))|_2 \mathrm{d}s \mathrm{d}\varphi = \frac{1}{2\pi} \int_{\mathbb{R}^2} |K \cap (K + x)|_2 \frac{\mathrm{d}x}{\|x\|} \\ &= \frac{1}{2\pi} \int_K \int_K \frac{\mathrm{d}x \mathrm{d}y}{\|x - y\|}. \end{aligned}$$

The latter double integral is known as *second-order chord power integral of* K, see e.g. Heinrich and Spiess [2013], p. 327, and Schneider and Weil [2008], Chapter 7, for integral geometric background.

In order to prove Theorem 4.4, the following two lemmas are essential. Interestingly, the assumptions to prove the following lemmas are rather mild in comparison with the Brillinger-mixing-type conditions in Theorems 4.3 and 4.4.

Lemma 4.4 (Lemma 5 in Flimmel and Heinrich [2021+]).
Let $\Psi \sim P$ be a second-order stationary point process on $\mathbb{R}^1$ with $\|\gamma_{red}^{[2]}\|_{TV} < \infty$. Further, suppose that $\mathbb{E}\, R_0 < \infty$ and $\Phi_0 \sim G$ with a not necessarily continuous distribution function G. Then

$$\rho \int_K \int_K \int_{\mathbb{R}^2} w_{\rho x_1}(p_1) w_{\rho x_2}(p_2) \gamma^{[2]}(\mathrm{d}(p_1, p_2)) \mathrm{d}x_1 \mathrm{d}x_2 \tag{4.34}$$

$$\xrightarrow[\rho \to \infty]{} \lambda (\mathbb{E}\,|\Xi_0|_1)^2 \gamma_{red}^{[2]}(\mathbb{R}^1) \int_{\mathbb{R}^1} \big(\mathbb{E}\,|g(p, \Phi_0) \cap K|_1 \big)^2 \mathrm{d}p.$$

Proof. By the stationarity of Ψ, we may write $\gamma^{[2]}(\mathrm{d}(p_1,p_2)) = \lambda\gamma^{[2]}_{red}(\mathrm{d}p_2 - p_1)\mathrm{d}p_1$ which gives

$$\rho \int\limits_{\mathbb{R}^2}\int\limits_K\int\limits_K w_{\rho x}(p_1)w_{\rho y}(p_2)\mathrm{d}x\mathrm{d}y\gamma^{[2]}(\mathrm{d}(p_1,p_2))$$
$$= \rho\lambda \int\limits_{\mathbb{R}^2}\int\limits_K\int\limits_K w_{\rho x}(p_1)w_{\rho y}(p_2+p_1)\mathrm{d}x\mathrm{d}y\gamma^{[2]}_{red}(\mathrm{d}p_2)\mathrm{d}p_1.$$

To determine the limit of the right-hand side as $\rho \to \infty$, we rewrite the probabilities $w_{\rho x}(p_1) = \mathbb{P}(p_1 \in \{\cdots\})$ and $w_{\rho y}(p_2+p_1) = \mathbb{P}(p_2+p_1 \in \{\cdots\})$ by means of the expectation (as integral over the product of probability measures) over the corresponding indicator function $\mathbf{1}_{\{\dots\}}$. We fix $\Xi_i = \xi_i$ (compact sets in $\mathbb{R}^1$) and $\Phi_i = \varphi_i$ (angles in $[0,\pi]$) for $i = 1, 2$ and omit the expectation which stands in front of all other integrals due to Fubini's theorem. The intensity λ will be suppressed. Further, we write $x = (x_1, x_2)$ and $y = (y_1, y_2)$. Thus, we only treat the integral

$$\begin{aligned} J_\rho(K,\xi_1,\varphi_1,\xi_2,\varphi_2) &:= \rho \int\limits_{\mathbb{R}^2}\int\limits_K\int\limits_K \mathbf{1}_{\xi_1+\rho(x_1\cos\varphi_1+x_2\sin\varphi_1)}(p_1) \\ &\quad\times \mathbf{1}_{\xi_2+\rho(y_1\cos\varphi_2+y_2\sin\varphi_2)}(p_2+p_1)\mathrm{d}(x_1,x_2)\mathrm{d}(y_1,y_2)\gamma^{[2]}_{red}(\mathrm{d}p_2)\mathrm{d}p_1 \\ &= \rho \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} \mathbf{1}_K(x_1,x_2)\mathbf{1}_K(y_1,y_2)\mathbf{1}_{\xi_1+\rho(x_1\cos\varphi_1+x_2\sin\varphi_1)}(p_1) \\ &\quad\times \mathbf{1}_{\xi_2+\rho(y_1\cos\varphi_2+y_2\sin\varphi_2)}(p_2+p_1)\mathrm{d}(x_1,x_2)\mathrm{d}(y_1,y_2)\gamma^{[2]}_{red}(\mathrm{d}p_2)\mathrm{d}p_1. \end{aligned} \tag{4.35}$$

Now, we substitute $(x_1,x_2)^T = \mathcal{O}(\varphi_1)(u_1,u_2)^T, (y_1,y_2)^T = \mathcal{O}(\varphi_2)(v_1,v_2)^T$, where $\mathcal{O}(\varphi_1)$ and $\mathcal{O}(\varphi_2)$ are defined by (4.9). Then $x_1 = u_1\cos\varphi_1 - u_2\sin\varphi_1$, $x_2 = u_1\sin\varphi_1 + u_2\cos\varphi_1$ and $y_1 = v_1\cos\varphi_2 - v_2\sin\varphi_2, y_2 = v_1\sin\varphi_2 + v_2\cos\varphi_2$. Hence, since $\mathcal{O}(\varphi_i)^{-1} = \mathcal{O}(-\varphi_i)$ for $i = 1, 2$, the integral $J_\rho(K,\xi_1,\varphi_1,\xi_2,\varphi_2)$ in (4.35) takes on the form

$$\begin{aligned} &\rho \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} \mathbf{1}_{\mathcal{O}(-\varphi_1)K}(u_1,u_2)\mathbf{1}_{\mathcal{O}(-\varphi_2)K}(v_1,v_2)\mathbf{1}_{\xi_1+\rho u_1}(p_1)\mathbf{1}_{\xi_2+\rho v_1}(p_2+p_1) \\ &\qquad\qquad\times \mathrm{d}(u_1,u_2)\mathrm{d}(v_1,v_2)\gamma^{[2]}_{red}(\mathrm{d}p_2)\mathrm{d}p_1 \\ &= \rho \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2} \mathbf{1}_{\mathcal{O}(-\varphi_1)K}(u_1,u_2)\mathbf{1}_{\mathcal{O}(-\varphi_2)K}(v_1,v_2)\mathbf{1}_{\xi_1+\rho(u_1-v_1)}(p_1)\mathbf{1}_{\xi_2}(p_2+p_1) \\ &\qquad\qquad\times \mathrm{d}(u_1,u_2)\mathrm{d}(v_1,v_2)\gamma^{[2]}_{red}(\mathrm{d}p_2)\mathrm{d}p_1. \end{aligned}$$

It is easy to see that the invariance properties of the one-dimensional Hausdorff measure on $\mathbb{R}^2$ (also denoted by $|\cdot|_1$) yield

$$\begin{aligned} \int\limits_{\mathbb{R}^1} \mathbf{1}_{\mathcal{O}(-\varphi_1)K}(u_1,u_2)\mathrm{d}u_2 &= |g(u_1,0) \cap \mathcal{O}(-\varphi_1)K|_1 \\ &= |\mathcal{O}(\varphi_1)g(u_1,0) \cap K|_1 \\ &= |g(u_1,\varphi_1) \cap K|_1. \end{aligned}$$

and likewise $\int\limits_{\mathbb{R}^1} \mathbf{1}_{\mathcal{O}(-\varphi_2)K}(v_1, v_2)\mathrm{d}v_2 = |g(v_1, \varphi_2) \cap K|_1$.

Therefore, the integral $J_\rho(K, \xi_1, \varphi_1, \xi_2, \varphi_2)$ is equal to

$$
\begin{aligned}
&\rho \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1} |g(u,\varphi_1)\cap K|_1 |g(v,\varphi_2)\cap K|_1 \mathbf{1}_{\xi_1+\rho(u-v)}(p_1)\mathbf{1}_{\xi_2}(p_2+p_1)\mathrm{d}u\mathrm{d}v\gamma_{red}^{[2]}(\mathrm{d}p_2)\mathrm{d}p_1 \\
&= \rho \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1} |g(w+v,\varphi_1)\cap K|_1 |g(v,\varphi_2)\cap K|_1 \mathbf{1}_{\xi_1+\rho w}(p_1)\mathbf{1}_{\xi_2}(p_2+p_1)\mathrm{d}w\mathrm{d}v \\
&\qquad\qquad \times \gamma_{red}^{[2]}(\mathrm{d}p_2)\mathrm{d}p_1 \\
&= \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1} |g(w/\rho+v,\varphi_1)\cap K|_1 |g(v,\varphi_2)\cap K|_1 \mathbf{1}_{\xi_1+w}(p_1)\mathbf{1}_{\xi_2}(p_2+p_1)\mathrm{d}w\mathrm{d}v \\
&\qquad\qquad \times \gamma_{red}^{[2]}(\mathrm{d}p_2)\mathrm{d}p_1 \\
&\xrightarrow[\rho\to\infty]{} \int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1} |g(v,\varphi_1)\cap K|_1 |g(v,\varphi_2)\cap K|_1 \mathbf{1}_{-\xi_1+p_1}(w)\mathbf{1}_{\xi_2-p_2}(p_1)\mathrm{d}w\mathrm{d}v \\
&\qquad\qquad \times \gamma_{red}^{[2]}(\mathrm{d}p_2)\mathrm{d}p_1 \\
&= |\xi_1|_1|\xi_2|_1\gamma_{red}^{[2]}(\mathbb{R}^1) \int\limits_{\mathbb{R}^1} \mathbf{1}_{[\ell(\varphi_1,K),r(\varphi_1,K)]}(v)|g(v,\varphi_1)\cap K|_1 \mathbf{1}_{[\ell(\varphi_2,K),r(\varphi_2,K)]}(v) \\
&\qquad\qquad \times |g(v,\varphi_2)\cap K|_1\mathrm{d}v,
\end{aligned}
$$

where the interval $[\ell(\varphi_i, K), r(\varphi_i, K)] = \{v \in \mathbb{R}^1 : g(v, \varphi_i) \cap K \neq \emptyset\}$ coincides with the orthogonal projection of $\mathcal{O}(-\varphi_i)K$ on the v-axis for $i = 1, 2$. To justify the above limit we have used that $|g(w/\rho + v, \varphi_1) \cap K|_1 \leqslant \mathrm{diam}(K)$ so that Lebesgue's dominated convergence theorem can be applied. Furthermore, it is easily seen that

$$
|J_\rho(K, \xi_1, \varphi_1, \xi_2, \varphi_2)| \leqslant \mathrm{diam}(K)|K|_2|\xi_1|_1|\xi_2|_1\|\gamma_{red}^{[2]}\|_{TV}. \tag{4.36}
$$

Hence, the limit of (4.34), i.e. limit of $\lambda\mathbb{E}\, J_\rho(K, \Xi_1, \Phi_1, \Xi_2, \Phi_2)$ as $\rho \to \infty$, exists and can be expressed by using the independence assumptions as follows:

$$
\lambda(\mathbb{E}\,|\Xi_0|_1)^2\gamma_{red}^{[2]}(\mathbb{R}^1) \int\limits_{\mathbb{R}^1} \big(\mathbb{E}\,\mathbf{1}_{[\ell(\Phi_0,K),r(\Phi_0,K)]}(v)|g(v,\Phi_0)\cap K|_1\big)^2\mathrm{d}v.
$$

Note that the indicator function $\mathbf{1}_{[\ell(\Phi_0,K),r(\Phi_0,K)]}(\cdot)$ can be omitted since the range of integration w.r.t. v is well-defined.

□

Lemma 4.5 (Lemma 6 in Flimmel and Heinrich [2021+]).
Assume that $\mathbb{E}\,R_0^2 < \infty$ and $\Phi_0 \sim G$ with a not necessarily continuous distribution function G. Then

$$
\rho \int\limits_{K}\int\limits_{K}\int\limits_{\mathbb{R}^1} w^{\cap}_{\rho x_1,\rho x_2}(p)\mathrm{d}p\mathrm{d}x_1\mathrm{d}x_2
\xrightarrow[\rho\to\infty]{} 2\mathbb{E}\,|\Xi_0|_1^2 \int\limits_0^{\pi} \int\limits_0^{r_K(\varphi\pm\frac{\pi}{2})} |K \cap (K + sv(\varphi \pm \frac{\pi}{2}))|_2\mathrm{d}s\mathrm{d}G(\varphi)
$$

with $r_K(\psi) := \max\{r \geqslant 0 : rv(\psi) \in K \oplus (-K)\}$. *Obviously, it is true that* $r_K(\psi) = r_K(\psi \pm \pi)$.

Proof. With the abbreviation $\Xi_0 = [-R_0, R_0]$ we obtain that

$$
\begin{aligned}
J_\rho(K) :=& \rho \int_K \int_K \int_{\mathbb{R}^1} w^{\cap}_{\rho x, \rho y}(p) \mathrm{d}p \mathrm{d}x \mathrm{d}x \\
=& \rho \int_K \int_K \int_{\mathbb{R}^1} \mathbb{P}\big(p \in \Xi_0 \cap (\Xi_0 + \rho\langle v(\Phi_0), y - x\rangle)\big) \mathrm{d}p \mathrm{d}x \mathrm{d}y \\
=& \rho \int_{\mathbb{R}^2} \mathbf{1}_{K \oplus (-K)}(y) |K \cap (K - y)|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho\langle v(\Phi_0), y\rangle)|_1 \mathrm{d}y \\
=& \rho \int_0^{2\pi} \int_0^{\infty} \mathbf{1}_{K \oplus (-K)}(sv(\psi)) |K \cap (K - sv(\psi))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s \cos(\Phi_0 - \psi))|_1 s \mathrm{d}s \mathrm{d}\psi,
\end{aligned}
$$

where we have substituted $y = sv(\psi)$ with $v(\psi) = (\cos\psi, \sin\psi)^T$ and with $r_K(\psi) = \max\{s \geqslant 0 : sv(\psi) \in K \oplus (-K)\} (= r_K(\psi \pm \pi)$ due to symmetry reasons). Moreover, using the independence of Φ_0 and R_0, the latter expression equals

$$
\begin{aligned}
&\rho \int_0^{2\pi} \int_0^{r_K(\psi)} |K \cap (K - sv(\psi))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s \cos(\Phi_0 - \psi))|_1 s \mathrm{d}s \mathrm{d}\psi \\
=& \rho \mathbb{E} \int_{-\Phi_0}^{2\pi - \Phi_0} \int_0^{r_K(\psi + \Phi_0)} |K \cap (K - sv(\psi + \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s \cos(\psi))|_1 s \mathrm{d}s \mathrm{d}\psi \\
=& \rho \mathbb{E} \int_0^{2\pi} \int_0^{r_K(\psi + \Phi_0)} |K \cap (K - sv(\psi + \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s \cos(\psi))|_1 s \mathrm{d}s \mathrm{d}\psi \\
=& 2\rho \mathbb{E} \int_0^{\pi} \int_0^{r_K(\psi + \Phi_0)} |K \cap (K + sv(\psi + \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s \cos(\psi))|_1 s \mathrm{d}s \mathrm{d}\psi,
\end{aligned}
$$

where we additionally used first, $\int_{-\Phi_0}^{0} (\cdots) \mathrm{d}\psi = \int_{2\pi - \Phi_0}^{2\pi} (\cdots) \mathrm{d}\psi$ due to $v(\psi) = v(\psi + 2\pi)$ and second, the shift-invariance of $|\cdot|_1$, the motion-invariance of $|\cdot|_2$ and $v(\psi + \pi) = -v(\psi)$.

By definition of $r_K(\psi)$, we have that $s > r_K(\psi)$ if and only if $sv(\psi) \notin K \oplus (-K)$ which happens if and only if $K \cap (K + sv(\psi)) = \varnothing$. Thus, the inner integral $\int_0^{r_K(\psi + \Phi_0)}$ in the above double integral can be replaced by $\int_0^{\infty}$ showing that

$$
\begin{aligned}
J_\rho(K) &= 2\rho \mathbb{E} \int_0^{\pi} \int_0^{\infty} |K \cap (K + sv(\psi + \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s \cos(\psi))|_1 s \mathrm{d}s \mathrm{d}\psi \\
&= 2\rho \int_{-1}^{1} \mathbb{E} \int_0^{\infty} |K \cap (K + sv(\arccos(y) + \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho s y)|_1 s \mathrm{d}s \frac{\mathrm{d}y}{\sqrt{1 - y^2}}
\end{aligned}
$$

by substituting $y = \cos(\psi) \in [-1, 1]$ so that we obtain $\psi = \arccos(y)$ and $(\arccos(y))' = -\frac{1}{\sqrt{1-y^2}}$. So, the latter equals

$$2\mathbb{E} \int_0^\infty \int_{-s}^{s} |K \cap (K + sv(\arccos(\frac{z}{s}) + \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho z)|_1 \frac{\rho \mathrm{d}z \mathrm{d}s}{\sqrt{s^2 - z^2}}$$

by substituting $z = sy \in [-s, s]$ so that $y = z/s$ and changing the order of integration. Interchanging again the integration over z and s, we can proceed with the abbreviation

$$\begin{aligned} h(s, z, \Phi_0) &:= sv(\arccos\big(\frac{z}{s}\big) + \Phi_0) \\ &= \big(z \cos\Phi_0 - \sqrt{s^2 - z^2} \sin\Phi_0, z \sin\Phi_0 + \sqrt{s^2 - z^2} \cos\Phi_0\big), \end{aligned}$$

where $0 \leqslant \|h(s, z, \Phi_0)\| = s \leqslant r_K := \max\{r_K(\varphi) : 0 \leqslant \varphi \leqslant \pi\} \leqslant \mathrm{diam}(K)$, leading to

$$\begin{aligned} J_\rho(K) &= 2\mathbb{E} \int_{\mathbb{R}^1} \int_{|z|}^{r_K} |K \cap (K + h(s, z, \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + \rho z)|_1 \frac{s \mathrm{d}s \rho \mathrm{d}z}{\sqrt{s^2 - z^2}} \\ &= 2\mathbb{E} \int_{\mathbb{R}^1} \int_{|u|/\rho}^{r_K} |K \cap (K + h(s, u/\rho, \Phi_0))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + u)|_1 \frac{s \mathrm{d}s \mathrm{d}u}{\sqrt{s^2 - (u/\rho)^2}} \end{aligned}$$

by substituting $u = \rho z$ so that $z = u/\rho$. Thus,

$$J_\rho(K) \xrightarrow[\rho\to\infty]{} 2\mathbb{E} \int_{\mathbb{R}^1} \int_0^{r_K} |K \cap (K + sv(\Phi_0 + \pi/2))|_2 \mathbb{E}\, |\Xi_0 \cap (\Xi_0 + u)|_1 \mathrm{d}s \mathrm{d}u.$$

We could apply Lebesgue's dominated convergence theorem since

$$\int_{|u|/\rho}^{r_K} |K \cap (K + h(s, u/\rho, \Phi_0))|_2 \frac{s \mathrm{d}s}{\sqrt{s^2 - (u/\rho)^2}} \leqslant |K|_2 \frac{1}{2} \int_0^{r_K^2 - \frac{u^2}{\rho^2}} \frac{\mathrm{d}t}{\sqrt{t}} \leqslant |K|_2 \mathrm{diam}(K).$$

Further, we use the continuity of the function $z \mapsto h(s, z, \varphi)$, $\arccos(0) = \pi/2$ and $h(s, 0, \varphi) = sv(\varphi + \pi/2) = s(-\sin\varphi, \cos\varphi)^T (= -sv(\varphi - \pi/2))$ and the relation $\int_{\mathbb{R}^1} |\Xi_0 \cap (\Xi_0 + u)|_1 \mathrm{d}u = |\Xi_0|_1^2 = 4R_0^2$ combined with a multiple application of Fubini's theorem. Finally, we arrive at

$$\begin{aligned} J_\rho(K) &= \rho \int_K \int_K \int_{\mathbb{R}^1} w^{\cap}_{\rho x_1, \rho x_2}(p) \mathrm{d}p \mathrm{d}x_1 \mathrm{d}x_2 \\ &\xrightarrow[\rho\to\infty]{} 2\mathbb{E}\, |\Xi_0|_1^2 \int_0^\pi \int_0^{r_K(\varphi \pm \frac{\pi}{2})} \big|K \cap \big(K + sv\big(\varphi \pm \frac{\pi}{2}\big)\big)\big|_2 \mathrm{d}s \mathrm{d}G(\varphi). \end{aligned}$$

□

Proof of Theorem 4.4. Recall that

$$\rho^{-3}\mathrm{Var}\left(|\Xi\cap\rho K|_2\right) = \int_K\int_K \rho\Big(G_\Psi\big[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)\big]-G_\Psi\big[1-w_{\rho x_1}(\cdot)\big]G_\Psi\big[1-w_{\rho x_2}(\cdot)\big]\Big)\mathrm{d}x_1\mathrm{d}x_2. \tag{4.37}$$

Instead of using the factorial moment expansion of the pgf's $G_\Psi[1-w^{\cup}_{\rho x_1,\rho x_2}]$, $G_P[1-w_{\rho x_1}]$ and $G_P[1-w_{\rho x_2}]$ as in (4.18) and (4.19), we first rewrite the integrand of the right-hand side of the foregoing equality as follows:

$$\begin{aligned}&\rho\Big(G_\Psi\big[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)\big]-G_\Psi\big[1-w_{\rho x_1}(\cdot)\big]G_\Psi\big[1-w_{\rho x_2}(\cdot)\big]\Big)\\ &=\rho G_\Psi\big[1-w_{\rho x_1}(\cdot)\big]G_\Psi\big[1-w_{\rho x_2}(\cdot)\big] \qquad (4.38)\\ &\times\Big(\exp\Big\{\log G_\Psi[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)]-\log G_\Psi[1-w_{\rho x_1}(\cdot)]-\log G_\Psi[1-w_{\rho x_2}(\cdot)]\Big\}-1\Big).\end{aligned}$$

In order to evaluate the exponent in (4.38), we use an expansion of $\log G_\Psi\big[1-w(\cdot)\big]$ in terms of the factorial cumulant measures $\gamma^{[k]}$ of $\Psi\sim P$, see Theorem 1.6, which is as follows:

$$\log G_\Psi\big[1-w(\cdot)\big]=\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_{\mathbb{R}^k}\prod_{j=1}^{k}w(p_j)\gamma^{[k]}(\mathrm{d}(p_1,\dots,p_k)), \tag{4.39}$$

provided the sum in (4.39) is convergent. In what follows we will show that

$$\overline{\lim_{\rho\to\infty}}\rho\int_K\int_K\Big|\log G_\Psi[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)]-\log G_\Psi[1-w_{\rho x_1}(\cdot)]-\log G_\Psi[1-w_{\rho x_2}(\cdot)]\Big|\mathrm{d}x_1\mathrm{d}x_2 \tag{4.40}$$

is finite. Before proving this, we note that the relation (4.28) implies that

$$\lim_{\rho\to\infty}G_\Psi[1-w_{\rho x}(\cdot)]=\sum_{k=0}^{m-1}\frac{(-1)^k(\lambda\mathbb{E}\,|\Xi_0|_1)^k}{k!}+\theta\frac{(\lambda\mathbb{E}\,|\Xi_0|_1)^m}{m!}\xrightarrow[m\to\infty]{}\exp\{-\lambda\mathbb{E}\,|\Xi_0|_1\} \tag{4.41}$$

for some $\theta\in[-1,1]$ uniformly for all $x\neq\mathbf{o}$. Furthermore, it is rapidly seen that the limit (4.30) (which has been proved under the assumptions of Theorem 4.3) holds if and only if

$$\lim_{\rho\to\infty}\Big(\log G_\Psi[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)]-\log G_\Psi[1-w_{\rho x_1}(\cdot)]-\log G_\Psi[1-w_{\rho x_2}(\cdot)]\Big)=0$$

for distinct points $x_1,x_2\in K\backslash\{\mathbf{o}\}$. Finally, the latter limit combined with (4.40) proves the equality

$$\begin{aligned}&\lim_{\rho\to\infty}\rho\int_K\int_K\Big(\exp\Big\{\log G_\Psi[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)]-\log G_\Psi[1-w_{\rho x_1}(\cdot)]\\ &\qquad\qquad -\log G_\Psi[1-w_{\rho x_2}(\cdot)]\Big\}-1\Big)\mathrm{d}x_1\mathrm{d}x_2\\ &=\lim_{\rho\to\infty}\rho\int_K\int_K\Big(\log G_\Psi[1-w^{\cup}_{\rho x_1,\rho x_2}(\cdot)]-\log G_\Psi[1-w_{\rho x_1}(\cdot)]\\ &\qquad\qquad -\log G_\Psi[1-w_{\rho x_2}(\cdot)]\Big)\mathrm{d}x_1\mathrm{d}x_2.\end{aligned}$$

The equality of both limits results from the inequality $|e^x - 1 - x| \leqslant \frac{x^2}{2} e^{\max(x,0)}$ and Lebesgue's dominated convergence theorem.

Combining the latter equality with (4.38), (4.40), (4.41) and the integral representation (4.37), we can state the relation

$$\lim_{\rho\to\infty} \rho^{-3} \mathrm{Var}\left(|\Xi \cap \rho K|_2\right) = e^{-2\lambda \mathbb{E}\,|\Xi_0|_1} \lim_{\rho\to\infty} \int\limits_K \int\limits_K \rho\Big(\log G_\Psi[1 - w^{\cup}_{\rho x_1, \rho x_2}(\cdot)] - \log G_\Psi[1 - w_{\rho x_1}(\cdot)] - \log G_\Psi[1 - w_{\rho x_2}(\cdot)]\Big) \mathrm{d}x_1 \mathrm{d}x_2. \tag{4.42}$$

By using the expansion (4.39), the double integral on the right-hand side of (4.42) takes the form

$$\int\limits_K \int\limits_K \rho\Big(\log G_\Psi[1 - w^{\cup}_{\rho x, \rho y}(\cdot)] - \log G_\Psi[1 - w_{\rho x}(\cdot)] - \log G_\Psi[1 - w_{\rho y}(\cdot)]\Big) \mathrm{d}x\mathrm{d}y = \sum_{n=1}^{\infty} \frac{(-1)^n T_n^{(\rho)}(K)}{n!},$$

where $T_n^{(\rho)}(K)$ for $n \in \mathbb{N}$ is defined by

$$T_n^{(\rho)}(K) := \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^n} \rho\Big(\prod_{j=1}^n w^{\cup}_{\rho x, \rho y}(p_j) - \prod_{j=1}^n w_{\rho x}(p_j) - \prod_{j=1}^n w_{\rho y}(p_j)\Big) \gamma^{[n]}(\mathrm{d}(p_1, \dots, p_n)) \mathrm{d}x\mathrm{d}y. \tag{4.43}$$

Since $\gamma^{[1]}(\mathrm{d}p) = \lambda \mathrm{d}p$ and $w^{\cup}_{\rho x, \rho y}(p) - w_{\rho x}(p) - w_{\rho y}(p) = -w^{\cap}_{\rho x, \rho y}(p)$, we get

$$-T_1^{(\rho)}(K) = \lambda \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^1} \rho w^{\cap}_{\rho x, \rho y}(p) \mathrm{d}p \mathrm{d}x \mathrm{d}y = \lambda J_\rho(K) \xrightarrow[m\to\infty]{} 2\lambda \mathbb{E}\,|\Xi_0|_1^2 C_2^{G,K},$$

where the limit is just the assertion of Lemma 4.5. The above proof of Lemma 4.5 reveals that $|T_1^{(\rho)}(K)| \leqslant \lambda J_\rho(K) \leqslant 2\lambda \mathbb{E}\,|\Xi_0|_1^2 |K|_2 \mathrm{diam}(K)$. In the next step, we derive a uniform bound of $T_2^{(\rho)}(K)$ as well as its limit as $\rho \to \infty$. For doing this, we rewrite

$$\prod_{j=1}^2 w^{\cup}_{\rho x, \rho y}(p_j) - \prod_{j=1}^2 w_{\rho x}(p_j) - \prod_{j=1}^2 w_{\rho y}(p_j) = w_{\rho x}(p_1) w_{\rho y}(p_2) + w_{\rho y}(p_1) w_{\rho x}(p_2) - w^{\cup}_{\rho x, \rho y}(p_1) w^{\cap}_{\rho x, \rho y}(p_2) - w^{\cap}_{\rho x, \rho y}(p_1)\big(w_{\rho x}(p_2) + w_{\rho y}(p_2)\big)$$

and by regarding the symmetry in x, y and p_1, p_2 we get

$$\begin{aligned} T_2^{(\rho)}(K) &= \rho \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} \Big(\prod_{j=1}^2 w^{\cup}_{\rho x, \rho y}(p_j) - \prod_{j=1}^2 w_{\rho x}(p_j) - \prod_{j=1}^2 w_{\rho y}(p_j)\Big) \gamma^{[2]}(\mathrm{d}(p_1, p_2)) \mathrm{d}x\mathrm{d}y \\ &= \rho \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} \Big(2w_{\rho x}(p_1) w_{\rho y}(p_2) - \big(w^{\cup}_{\rho x, \rho y}(p_2) + 2w_{\rho x}(p_2)\big) w^{\cap}_{\rho x, \rho y}(p_1)\Big) \gamma^{[2]}(\mathrm{d}(p_1, p_2)) \mathrm{d}x\mathrm{d}y \\ &= 2\rho \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w_{\rho x}(p_1) w_{\rho y}(p_2) \gamma^{[2]}(\mathrm{d}(p_1, p_2)) \mathrm{d}x\mathrm{d}y + \widetilde{T}_2^{(\rho)}(K), \end{aligned} \tag{4.44}$$

where

$$|\widetilde{T}_2^{(\rho)}(K)| \leqslant 3\lambda\rho \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^1} \int\limits_{\mathbb{R}^1} w^{\cap}_{\rho x,\rho y}(p_1) w_{\rho x}(p_2+p_1) |\gamma^{[2]}_{red}|(\mathrm{d}p_2)\mathrm{d}p_1 \mathrm{d}x\mathrm{d}y$$
$$= 3\lambda\rho \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^1} w^{\cap}_{\rho x,\rho y}(p_1) \mathbb{E}\, |\gamma^{[2]}_{red}| \big(\Xi_0 + \rho\langle v(\Phi_0), x\rangle - p_1\big) \mathrm{d}p_1 \mathrm{d}x\mathrm{d}y. \tag{4.45}$$

Clearly, we have $\infty > |\gamma^{[2]}_{red}|(\mathbb{R}^1) \geqslant \mathbb{E}\, |\gamma^{[2]}_{red}| \big(\Xi_0 + \rho\langle v(\Phi_0), x\rangle - p_1\big) \xrightarrow[\rho\to\infty]{} 0$ for $x \neq \mathbf{o}$. Together with the arguments used in the proof of Lemma 4.5, among them the uniform estimate $J_\rho(K) \leqslant 2\mathbb{E}\,|\Xi_0|^2 |K|_2 \mathrm{diam}(K)$, it follows that $\widetilde{T}_2^{(\rho)}(K) \xrightarrow[\rho\to\infty]{} 0$. Finally, Lemma 4.4 and (4.44) show that

$$\frac{T_2^{(\rho)}(K)}{2} \xrightarrow[\rho\to\infty]{} \lambda(\mathbb{E}\,|\Xi_0|_1)^2 \gamma^{[2]}_{red}(\mathbb{R}^1) \int\limits_{\mathbb{R}^1} \big(\mathbb{E}\,|g(p,\Phi_0) \cap K|_1\big)^2 \mathrm{d}p$$
$$= \lambda(\mathbb{E}\,|\Xi_0|_1)^2 \gamma^{[2]}_{red}(\mathbb{R}^1) C_1^{G,K}.$$

In addition, we can derive a uniform bound of $T_2^{(\rho)}(K)$. From (4.45) and the above bound of $T_1^{(\rho)}(K)$ we get

$$|\widetilde{T}_2^{(\rho)}(K)| \leqslant 3\|\gamma^{[2]}_{red}\|_{TV} |T_1^{(\rho)}(K)| \leqslant 6\lambda |K|_2 \mathrm{diam}(K) \|\gamma^{[2]}_{red}\|_{TV} \mathbb{E}\,|\Xi_0|_1^2.$$

Hence, we see from (4.36) and (4.43) that, for two independent pairs $(\Xi_i, \Phi_i), i = 1, 2$, with the same distribution as (Ξ_0, Φ_0), the following estimate holds:

$$|T_2^{(\rho)}(K)| \leqslant 2\lambda |J_\rho(K, \Xi_1, \Phi_1, \Xi_2, \Phi_2)| + \widetilde{T}_2^{(\rho)}(K)$$
$$\leqslant 8\lambda |K|_2 \mathrm{diam}(K) \mathbb{E}\,|\Xi_0|_1^2 \|\gamma^{[2]}_{red}\|_{TV}.$$

Obviously, the limit (4.33) coincides with $\lim_{\rho\to\infty}(-T_1^{(\rho)}(K) + \frac{1}{2}T_2^{(\rho)}(K))$. Thus, the proof of Theorem 4.4 is accomplished if we show that

$$\lim_{\rho\to\infty} T_n^{(\rho)}(K) = 0 \text{ and } \sup_{\rho\geqslant 1} \frac{|T_n^{(\rho)}(K)|}{n!} \leqslant C_n^K \text{ for } n \geqslant 3 \text{ such that } \sum_{n\geqslant 3} C_n^K < \infty. \tag{4.46}$$

This means that we have to find suitable upper bounds of the integrals (4.43) for each $n \geqslant 3$ which are uniform w.r.t. ρ and disappear as $\rho \to \infty$. Using the reduced factorial cumulant measures $\gamma^{[n]}_{red}$ defined (in differential notation) by $\gamma^{[n]}(\mathrm{d}(p_1,\dots,p_n)) = \lambda\gamma^{[n]}_{red}((\mathrm{d}p_i - p_j : i \neq j))\mathrm{d}p_j$ for any $j = 1,\dots,n$, the boundedness of the total variation measure $|\gamma^{[n]}_{red}|(\cdot)$ on $\mathbb{R}^{n-1}$ and obvious relations

$$\prod_{i=1}^n \big(w_{\rho x}(p_i) + w_{\rho y}(p_i)\big) - \prod_{i=1}^n w_{\rho x}(p_i) - \prod_{i=1}^n w_{\rho y}(p_i)$$
$$= \sum_{k=1}^{n-1} \sum_{1\leqslant i_1<\cdots<i_k\leqslant n} \prod_{\ell=1}^k w_{\rho x}(p_{i_\ell}) \prod_{\substack{j=1 \\ j\neq i_1,\dots,i_k}}^n w_{\rho y}(p_j)$$

and

$$\prod_{i=1}^{n}\big(w_{\rho x}(p_i)+w_{\rho y}(p_i)\big)-\prod_{i=1}^{n}w^{\cup}_{\rho x,\rho y}(p_i)$$
$$=\sum_{k=1}^{n}w^{\cap}_{\rho x,\rho y}(p_k)\prod_{i=1}^{k-1}w^{\cup}_{\rho x,\rho y}(p_i)\prod_{j=k+1}^{n}\big(w_{\rho x}(p_j)+w_{\rho y}(p_j)\big)$$
$$\leqslant\sum_{k=1}^{n}w^{\cap}_{\rho x,\rho y}(p_k)\prod_{\substack{j=1\\ j\neq k}}^{n}\big(w_{\rho x}(p_j)+w_{\rho y}(p_j)\big),$$

we obtain the following estimates

$$\Big|\int_K\int_K\int_{\mathbb{R}^n}\rho\Big(\prod_{i=1}^{n}\big(w_{\rho x}(p_i)+w_{\rho y}(p_i)\big)-\prod_{i=1}^{n}w_{\rho x}(p_i)-\prod_{i=1}^{n}w_{\rho y}(p_i)\Big)\gamma^{[n]}(\mathrm{d}(p_1,\dots,p_n))\mathrm{d}x\mathrm{d}y\Big|$$
$$=\Big|\sum_{k=1}^{n-1}\binom{n}{k}\int_K\int_K\int_{\mathbb{R}^n}\rho\prod_{i=1}^{k}w_{\rho x}(p_i)\prod_{j=k+1}^{n}w_{\rho y}(p_j)\gamma^{[n]}(\mathrm{d}(p_1,\dots,p_n))\mathrm{d}x\mathrm{d}y\Big|$$
$$\leqslant T^{(\rho)}_{n,1}(K):=\lambda\sum_{k=1}^{n-1}\binom{n}{k}\int_K\int_K\int_{\mathbb{R}^1}\rho w_{\rho x}(p_1)\int_{\mathbb{R}^{n-1}}\prod_{i=2}^{k}w_{\rho x}(p_i+p_1)\prod_{j=k+1}^{n}w_{\rho y}(p_j+p_1)$$
$$\times\big|\gamma^{[n]}_{red}\big|(\mathrm{d}(p_2,\dots,p_n))\mathrm{d}p_1\mathrm{d}x\mathrm{d}y \tag{4.47}$$

and

$$\Big|\int_K\int_K\int_{\mathbb{R}^n}\rho\Big(\prod_{i=1}^{n}\big(w_{\rho x}(p_i)+w_{\rho y}(p_i)\big)-\prod_{i=1}^{n}w^{\cup}_{\rho x,\rho y}(p_i)\Big)\gamma^{[n]}(\mathrm{d}(p_1,\dots,p_n))\mathrm{d}x\mathrm{d}y\Big|$$
$$\leqslant\lambda n\int_K\int_K\int_{\mathbb{R}^1}\rho w^{\cap}_{\rho x,\rho y}(p_1)\int_{\mathbb{R}^{n-1}}\prod_{j=2}^{n}\big(w_{\rho x}(p_j+p_1)+w_{\rho y}(p_j+p_1)\big)\big|\gamma^{[n]}_{red}\big|(\mathrm{d}(p_2,\dots,p_n))\mathrm{d}p_1\mathrm{d}x\mathrm{d}y$$
$$\leqslant T^{(\rho)}_{n,2}(K):=\lambda n\sum_{k=1}^{n}\binom{n-1}{k-1}\int_K\int_K\int_{\mathbb{R}^1}\rho w^{\cap}_{\rho x,\rho y}(p_1)\int_{\mathbb{R}^{n-1}}\prod_{i=2}^{k}w_{\rho x}(p_i+p_1)$$
$$\times\prod_{j=k+1}^{n}w_{\rho y}(p_j+p_1)\big|\gamma^{[n]}_{red}\big|(\mathrm{d}(p_2,\dots,p_n))\mathrm{d}p_1\mathrm{d}x\mathrm{d}y. \tag{4.48}$$

Obviously, we have $|T^{(\rho)}_{n}(K)|\leqslant T^{(\rho)}_{n,1}(K)+T^{(\rho)}_{n,2}(K)$ for $n\geqslant 3$. Let us first, rewrite the integral terms in (4.47). For this purpose we introduce the abbreviation

$$I^{(\rho)}_{n,k}(K):=\int_K\int_K\int_{\mathbb{R}^1}\rho w_{\rho x}(p_1)\int_{\mathbb{R}^{n-1}}\prod_{i=2}^{k}w_{\rho x}(p_i+p_1)\prod_{j=k+1}^{n}w_{\rho y}(p_j+p_1)$$
$$\times\big|\gamma^{[n]}_{red}\big|(\mathrm{d}(p_2,\dots,p_n))\mathrm{d}p_1\mathrm{d}x\mathrm{d}y$$

for $k=2,\dots,n-1$. As in (4.35) we substitute $x=\mathcal{O}(\Phi_1)u$ and $y=\mathcal{O}(\Phi_n)w$ with $\mathcal{O}(\cdot)$ as defined in (4.9). Since $\mathcal{O}^{-1}(\varphi)=\mathcal{O}(-\varphi)$ and $\det(\mathcal{O}(\varphi))=1$ it follows that $u=\mathcal{O}(-\Phi_1)x$, $w=\mathcal{O}(-\Phi_n)y$ and $\langle v(\Phi_i),x\rangle=\langle v(\Phi_i),\mathcal{O}(\Phi_1)u\rangle=\langle v(\Phi_i-\Phi_1),u\rangle$ for $i=1,\dots,k$ and $\langle v(\Phi_j),y\rangle=\langle v(\Phi_j-\Phi_n),w\rangle$ for $j=k+1,\dots,n$. Note that $\langle v(\Phi_1), x\rangle = u_1$ and $\langle v(\Phi_n), y\rangle = w_1$ for $u = (u_1, u_2)^T$ and $w = (w_1, w_2)^T$, respectively.

Similarly as in the proof of Lemma 4.2, let $(R_1, \Phi_1), \ldots, (R_n, \Phi_n)$ be independent copies of the random vector (R_0, Φ_0) and $\Xi_1, \ldots, \Xi_n$ independent copies of the random interval $\Xi_0 = [-R_0, R_0]$. Then the product $w_{\rho x}(p) \prod_{i=2}^k w_{\rho x}(p_i + p_1) \prod_{j=k+1}^n w_{\rho y}(p_j + p_1)$ can be expressed as the expectation

$$\mathbb{E}\Big(\mathbf{1}_{\Xi_1+\rho\langle\Phi_1,x\rangle}(p_1)\prod_{i=2}^k \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i),x\rangle}(p_i+p_1)\prod_{j=k+1}^n \mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j),y\rangle}(p_j+p_1)\Big),$$

which, together with the above transformations of $x, y \in \mathbb{R}^2$ and Fubini's theorem, allows us to write $I_{n,k}^{(\rho)}(K)$ in the form

$$\begin{aligned}
&\mathbb{E}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^1}\int_{\mathbb{R}^{n-1}}\rho\Big(\mathbf{1}_{\Xi_1+\rho u_1}(p_1)\prod_{i=2}^k \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i-\Phi_1),u\rangle}(p_i+p_1)\\
&\qquad\times\prod_{j=k+1}^{n-1}\mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle}(p_j+p_1)\mathbf{1}_{\Xi_n}(p_n+p_1-\rho w_1)\Big)\\
&\qquad\times\left|\gamma_{red}^{[n]}\right|(\mathrm{d}(p_2,\ldots,p_n))\mathrm{d}p_1\mathbf{1}_{\mathcal{O}(-\Phi_1)K}(u)\mathbf{1}_{\mathcal{O}(-\Phi_n)K}(w)\mathrm{d}(u_1,u_2)\mathrm{d}(w_1,w_2)\\
=&\mathbb{E}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^1}\int_{\mathbb{R}^{n-1}}\rho\Big(\mathbf{1}_{\Xi_1+\rho(u_1-w_1)}(p_1-\rho w_1)\prod_{i=2}^k \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i-\Phi_1),u\rangle-\rho w_1}(p_i+p_1-\rho w_1)\\
&\qquad\times\prod_{j=k+1}^{n-1}\mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle-\rho w_1}(p_j+p_1-\rho w_1)\mathbf{1}_{\Xi_n}(p_n+p_1-\rho w_1)\Big)\\
&\qquad\times\left|\gamma_{red}^{[n]}\right|(\mathrm{d}(p_2,\ldots,p_n))\mathrm{d}p_1\mathbf{1}_{\mathcal{O}(-\Phi_1)K}(u)\mathbf{1}_{\mathcal{O}(-\Phi_n)K}(w)\mathrm{d}(u_1,u_2)\mathrm{d}(w_1,w_2)\\
=&\mathbb{E}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^1}\int_{\mathbb{R}^{n-1}}\rho\Big(\prod_{i=2}^k \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i-\Phi_1),u\rangle-\rho w_1}(p_i+p_1)\\
&\qquad\times\prod_{j=k+1}^{n-1}\mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle-\rho w_1}(p_j+p_1)\mathbf{1}_{\Xi_1+\rho(u_1-w_1)}(p_1)\mathbf{1}_{\Xi_n}(p_n+p_1)\Big)\\
&\qquad\times\left|\gamma_{red}^{[n]}\right|(\mathrm{d}(p_2,\ldots,p_n))\mathrm{d}p_1\mathbf{1}_{\mathcal{O}(-\Phi_1)K}(u)\mathbf{1}_{\mathcal{O}(-\Phi_n)K}(w)\mathrm{d}(u_1,u_2)\mathrm{d}(w_1,w_2)\\
=&\mathbb{E}\int_{\mathbb{R}^2}\int_{\mathbb{R}^2}\int_{\mathbb{R}^1}\int_{\mathbb{R}^{n-1}}\rho\Big(\prod_{i=2}^k \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i-\Phi_1),(z_1+w_1,z_2)\rangle-\rho w_1}(p_i+p_1)\\
&\qquad\times\prod_{j=k+1}^{n-1}\mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle-\rho w_1}(p_j+p_1)\mathbf{1}_{\Xi_1+\rho_1}(p)\mathbf{1}_{\Xi_n}(p_n+p)\Big)\\
&\qquad\times\left|\gamma_{red}^{[n]}\right|(\mathrm{d}(p_2,\ldots,p_n))\mathrm{d}p\mathbf{1}_{\mathcal{O}(-\Phi_1)K}((z_1+w_1,z_2))\mathbf{1}_{\mathcal{O}(-\Phi_n)K}(w)\\
&\qquad\times\mathrm{d}(z_1,z_2)\mathrm{d}(w_1,w_2)
\end{aligned}$$

$$
\begin{aligned}
=&\mathbb{E}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^{n-1}}\Big(\prod_{i=2}^{k}\mathbf{1}_{\Xi_i+\langle v(\Phi_i-\Phi_1),(z_1+\rho w_1,\rho z_2)\rangle-\rho w_1}(p_i+p_1)\\
&\qquad\times\prod_{j=k+1}^{n-1}\mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle-\rho w_1}(p_j+p_1)\mathbf{1}_{\Xi_n}(p_n+p_1)\Big)\\
&\qquad\times|\gamma_{red}^{[n]}|(\mathrm{d}(p_2,\dots,p_n))\mathbf{1}_{\Xi_1+z_1}(p_1)\mathrm{d}p_1\mathbf{1}_{\mathcal{O}(-\Phi_1)K}((\frac{z_1}{\rho}+w_1,z_2))\\
&\qquad\times\mathbf{1}_{\mathcal{O}(-\Phi_n)K}((w_1,w_2))\mathrm{d}(z_1,z_2)\mathrm{d}(w_1,w_2).
\end{aligned}\tag{4.49}
$$

Replacing the two products of indicator functions in (4.49) by 1 leads to the following bound of $I_{n,k}^{(\rho)}(K)$ provided that $|\gamma_{red}^{[n]}|(\mathbb{R}^{n-1})<\infty$:

$$
\begin{aligned}
I_{n,k}^{(\rho)}(K)\leqslant&\ \mathbb{E}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^{n-1}}\Big(\mathbf{1}_{-\Xi_1+p_1}(z_1)\mathbf{1}_{\Xi_n-p_n}(p_1)\Big)|\gamma_{red}^{[n]}|(\mathrm{d}(p_2,\dots,p_n))\mathrm{d}p_1\\
&\qquad\times\mathbf{1}_{\mathcal{O}(-\Phi_1)K}((\frac{z_1}{\rho}+w_1,z_2))\mathrm{d}(z_1,z_2)\mathbf{1}_{\mathcal{O}(-\Phi_n)K}((w_1,w_2))\mathrm{d}(w_1,w_2)\\
=&\ \mathbb{E}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^{n-1}}\Big(\mathbf{1}_{-\Xi_1+p_1}(z_1)\mathbf{1}_{\Xi_n-p_n}(p_1)\Big)|\gamma_{red}^{[n]}|(\mathrm{d}(p_2,\dots,p_n))\mathrm{d}p_1\\
&\qquad\times|g(\frac{z_1}{\rho}+w_1,\Phi_1)\cap K|_1|g(w_1,\Phi_n)\cap K|_1\mathrm{d}z_1\mathrm{d}w_1\\
\leqslant&\ \mathrm{diam}(K)\mathbb{E}\int\limits_{\mathbb{R}^1}|g(w_1,\Phi_n)\cap K|_1\mathrm{d}w_1\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1}\mathbf{1}_{-\Xi_1+p_1}(z_1)\mathbf{1}_{\Xi_n-p_n}(p_1)\mathrm{d}z_1\mathrm{d}p_1\\
&\qquad\times|\gamma_{red}^{[n]}|(\mathbb{R}^{n-1})\\
=&\ \mathrm{diam}(K)|K|_2\mathbb{E}\,|\Xi_1|_1\mathbb{E}\,|\Xi_n|_1|\gamma_{red}^{[n]}|(\mathbb{R}^{n-1})\\
=&\ \mathrm{diam}(K)|K|_2(\mathbb{E}\,|\Xi_0|_1)^2\|\gamma_{red}^{[n]}\|_{TV}.
\end{aligned}\tag{4.50}
$$

Here, we have used arguments that have already been applied to prove (4.36). Also, the product of the indicator functions in the first line of (4.49) disappears as $\rho\to\infty$ $\mathbb{P}$-a.s. and for a.a. $(w_1,w_2),(z_1,z_2),p_1,(p_2,\dots,p_n)\in\mathbb{R}^{n+4}$ w.r.t. the corresponding product measure. Therefore, again by Lebesgue's dominated convergence theorem,

$$
\lim_{\rho\to\infty}I_{n,k}^{(\rho)}(K)=0\quad\text{for}\quad k=2,\dots,n,n\geqslant 3.\tag{4.51}
$$

Next, we derive a further bound of $I_{n,k}^{(\rho)}(K)$ that depends more on the mean thickness $\mathbb{E}\,|\Xi_0|_1$ of the typical cylinder. For this, we need the Radon–Nikodym density $|c_{red}^{(n)}(p_2,\dots,p_n)|$ of $|\gamma_{red}^{[n]}(\cdot)|$ w.r.t. to Lebesgue measure on $\mathbb{R}^{n-1}$. Hence, by using Fubini's theorem, we replace the integral (4.50) over $\mathbb{R}^{n-1}$ by two iterated integrals. The first integral over $(p_2,\dots,p_{n-1})\in\mathbb{R}^{n-2}$ can be estimated by Hölder's inequality as follows:

$$\int\limits_{\mathbb{R}^{n-2}} \prod_{i=2}^{k} \mathbf{1}_{\Xi_i+\langle v(\Phi_i-\Phi_1),(z_1+\rho w_1,\rho z_2)\rangle-\rho w_1-p_1}(p_i) \prod_{j=k+1}^{n-1} \mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle-\rho w_1-p_1}(p_j)$$
$$\times \left|c_{red}^{(n)}(p_2,\ldots,p_{n-1},p_n)\right| \mathrm{d}(p_2,\ldots,p_{n-1})$$
$$\leqslant \Big(\int\limits_{\mathbb{R}^{n-2}} \prod_{i=2}^{k} \mathbf{1}_{\Xi_i+\langle v(\Phi_i-\Phi_1),(z_1+\rho w_1,\rho z_2)\rangle-\rho w_1-p_1}(p_i) \prod_{j=k+1}^{n-1} \mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j-\Phi_n),w\rangle-\rho w_1-p_1}(p_j)$$
$$\times \mathrm{d}(p_2,\ldots,p_{n-1})\Big)^{\frac{q-1}{q}} \Big(\int\limits_{\mathbb{R}^{n-2}} \left|c_{red}^{(n)}(p_2,\ldots,p_{n-1},p_n)\right|^q \mathrm{d}(p_2,\ldots,p_{n-1})\Big)^{\frac{1}{q}}$$
$$= \Big(\prod_{i=2}^{n-1} |\Xi_i|_1\Big)^{\frac{q-1}{q}} \|c_{red}^{(n)}(\cdot,p_n)\|_q \tag{4.52}$$

for any $q > 1$, where $\|c_{red}^{(n)}(\cdot,p_n)\|_q$ coincides with the term before the equal sign in (4.52). Combining the estimates (4.47) and (4.52) with $|g(p,\varphi)\cap K|_1 \leqslant \mathrm{diam}(K)$ for $(p,\varphi) \in \mathbb{R}^1 \times [0,\pi]$, $\int_{\mathbb{R}^1} |g(p,\varphi)\cap K|_1 \mathrm{d}p = |K|_2$, switching the order of integration and finally applying Lyapunov's inequality we arrive at

$$I_{n,k}^{(\rho)}(K) \leqslant \mathbb{E} \int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1} \Big(\prod_{i=2}^{n-1} |\Xi_i|_1\Big)^{\frac{q-1}{q}} \|c_{red}^{(n)}(\cdot,p_n)\|_q \mathbf{1}_{-\Xi_1+p_1}(z_1)\mathbf{1}_{\Xi_n-p_n}(p_1)\mathrm{d}p_1\mathrm{d}p_n$$
$$\times |g(\frac{z_1}{\rho}+w_1,\Phi_1)\cap K|_1 |g(w_1,\Phi_n)\cap K|_1 \mathrm{d}z_1\mathrm{d}w_1$$
$$\leqslant \mathrm{diam}(K)|K|_2 \mathbb{E}\Big(\prod_{i=2}^{n-1} |\Xi_i|_1\Big)^{\frac{q-1}{q}} \int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1}\int\limits_{\mathbb{R}^1} \|c_{red}^{(n)}(\cdot,p_n)\|_q \mathbf{1}_{-\Xi_1+p_1}(z_1)\mathbf{1}_{\Xi_n-p_n}(p_1)\mathrm{d}z_1\mathrm{d}p_1\mathrm{d}p_n$$
$$= \mathrm{diam}(K)|K|_2 \int\limits_{\mathbb{R}^1} \|c_{red}^{(n)}(\cdot,p)\|_q \mathrm{d}p \big(\mathbb{E}\,|\Xi_0|_1\big)^{\frac{n(q-1)}{q}+\frac{2}{q}}.$$

Applying the same arguments as above, the estimate (4.51) reveals that (4.51) remains true if, instead of $\|\gamma_{red}^{[n]}\|_{TV} < \infty$, we assume that the L_q^*-norm defined by $\|c_{red}^{(n)}\|_q^* := \int_{\mathbb{R}^1} \|c_{red}^{(n)}(\cdot,p)\|_q \mathrm{d}p$ is finite for some $q > 1$ and $n \geqslant 3$. Hence, we have

$$T_{n,1}^{(\rho)}(K) = \lambda \sum_{k=1}^{n-1} \binom{n}{k} I_{n,k}^{(\rho)}(K)$$
$$\leqslant \lambda \mathrm{diam}(K)|K|_2(2^n-2)\big(\mathbb{E}\,|\Xi_0|_1\big)^{\frac{n(q-1)}{q}+\frac{2}{q}} \|c_{red}^{(n)}\|_q^*.$$

Together with the strong L_q^*-Brillinger mixing condition with $b_q^*(\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}} < 1/2$, we get

$$\sum_{n\geqslant 3} \frac{T_{n,1}^{(\rho)}(K)}{n!} \leqslant \lambda a_q^*(\mathbb{E}\,|\Xi_0|_1)^{\frac{2}{q}} \mathrm{diam}(K)|K|_2 \sum_{n\geqslant 3} \big(2b_q^*(\mathbb{E}\,|\Xi_0|_1)^{\frac{(q-1)}{q}}\big)^n$$
$$\leqslant \frac{\lambda a_q^*(\mathbb{E}\,|\Xi_0|_1)^{\frac{2}{q}} \mathrm{diam}(K)|K|_2}{1-2b_q^*(\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}}}.$$

Next, we derive two different bounds for the sum $T^{(\rho)}_{n,2}(K)$ defined in (4.48). For this purpose, in analogy to $I^{(\rho)}_{n,k}(K)$, we need uniform bounds of

$$J^{(\rho)}_{n,k}(p) := \int\limits_{\mathbb{R}^{n-1}} \prod_{i=2}^{k} w_{\rho x}(p_i + p_1) \prod_{j=k+1}^{n} w_{\rho y}(p_j + p)\big|\gamma^{[n]}_{red}\big|(\mathrm{d}(p_2, \dots, p_n)).$$

It is easily seen that

$$\begin{aligned} J^{(\rho)}_{n,k}(p) &= \mathbb{E} \int\limits_{\mathbb{R}^{n-1}} \prod_{i=2}^{k} \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i),x\rangle - p_1}(p_i) \prod_{j=k+1}^{n} \mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j),y\rangle-p_1}(p_j)\big|\gamma^{[n]}_{red}\big|\mathrm{d}(p_2, \dots, p_n) \\ &\leqslant \big|\gamma^{[n]}_{red}\big|(\mathbb{R}^{n-1}) \end{aligned}$$

and, for any $q > 1$ such that $\|c^{(n)}_{red}\|_q < \infty$

$$\begin{aligned} J^{(\rho)}_{n,k}(p) &= \mathbb{E} \int\limits_{\mathbb{R}^{n-1}} \prod_{i=2}^{k} \mathbf{1}_{\Xi_i+\rho\langle v(\Phi_i),x\rangle-p}(p_i) \prod_{j=k+1}^{n} \mathbf{1}_{\Xi_j+\rho\langle v(\Phi_j),y\rangle-p}(p_j) c^{(n)}_{red}(p_2, \dots, p_n) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \mathrm{d}(p_2, \dots, p_n) \\ &\leqslant \mathbb{E} \prod_{i=2}^{n} \big|\Xi_i\big|_1^{\frac{q-1}{q}} \Big(\int\limits_{\mathbb{R}^{n-1}} \big|c^{(n)}_{red}(p_2, \dots, p_n)\big|^q \mathrm{d}(p_2, \dots, p_n)\Big)^{\frac{1}{q}} \\ &\leqslant (\mathbb{E}\,|\Xi_0|_1)^{(n-1)\frac{q-1}{q}} \|c^{(n)}_{red}\|_q. \end{aligned}$$

The foregoing estimates show that

$$\lim_{\rho\to\infty} J^{(\rho)}_{n,k}(p) = 0 \quad \text{for} \quad k = 2, \dots, n, n \geqslant 3. \tag{4.53}$$

Further, from the definition of $T^{(\rho)}_{n,2}(K)$, see (4.48), and the integral $J_\rho(K)$ introduced and estimated in the proof of Lemma 4.5 with the uniform upper bound $2\,\mathrm{diam}(K)|K|_2\mathbb{E}\,|\Xi_0|_1^2$, we see that

$$\begin{aligned} T^{(\rho)}_{n,2}(K) &\leqslant \lambda n \sum_{k=1}^{n} \binom{n-1}{k-1} \int\limits_{K}\int\limits_{K}\int\limits_{\mathbb{R}^1} \rho w^{\cap}_{\rho x,\rho y}(p_1)\mathrm{d}p_1\mathrm{d}x\mathrm{d}y \max_{2\leqslant k\leqslant n} \sup_{p\in\mathbb{R}^1} J^{(\rho)}_{n,k}(p) \\ &= \lambda n 2^{n-1} J_\rho(K) \max_{2\leqslant k\leqslant n} \sup_{p\in\mathbb{R}^1} J^{(\rho)}_{n,k}(p) \\ &\leqslant \lambda n 2^{n} \mathrm{diam}(K)|K|_2\mathbb{E}\,|\Xi_0|_1^2 \max_{2\leqslant k\leqslant n} \sup_{p\in\mathbb{R}^1} J^{(\rho)}_{n,k}(p) \end{aligned}$$

Under the assumption that $\Psi \sim P$ is either strongly Brillinger-mixing with $b < 1/2$ or strongly L_q-Brillinger-mixing with $b_q(\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}} < 1/2$ we obtain the inequalities

$$\begin{aligned} \sum_{n\geqslant 3} \frac{T^{(\rho)}_{n,2}(K)}{n!} &\leqslant 2\lambda ab\mathbb{E}\,|\Xi_0|_1^2 \mathrm{diam}(K)|K|_2 \sum_{n\geqslant 3} n\big(2b\big)^{n-1} \\ &\leqslant \frac{2\lambda ab\mathbb{E}\,|\Xi_0|_1^2 \mathrm{diam}(K)|K|_2}{(1-2b)^2} \end{aligned}$$

and

$$\sum_{n\geqslant 3}\frac{T^{(\rho)}_{n,2}(K)}{n!}\leqslant 2\lambda a_q b_q \mathbb{E}\,|\Xi_0|_1^2 \mathrm{diam}(K)|K|_2\sum_{n\geqslant 3} n\big(2b_q(\mathbb{E}\,|\Xi_0|_1)^{\frac{q-1}{q}}\big)^{n-1}$$
$$\leqslant \frac{2\lambda a_q b_q \mathbb{E}\,|\Xi_0|_1^2 \mathrm{diam}(K)|K|_2}{\big(1-2b_q(\mathbb{E}\,|\Xi_0|_1)^{1-\frac{1}{q}}\big)^2}.$$

Finally, summarizing the above-proved relations (4.51), (4.53) and the convergence of the series $\sum_{n\geqslant 3} T^{(\rho)}_{n,i}(K)/n!$ for $i = 1, 2$ shows the validity of (4.46) which in turn implies (4.40). Thus, the proof of Theorem 4.4 is complete.

□

Remark. Note that in Theorems 4.3 and 4.4, the interval $\Xi_0 := [-R_0, R_0]$ with $\mathbb{E}\, R_0^k < \infty$ can be replaced by a finite union of random closed intervals $\Xi_0 \subset \mathbb{R}^1$ satisfying $\inf \Xi_0 \leqslant 0 \leqslant \sup \Xi_0$ and $\mathbb{E}\,|\Xi_0|_1^k < \infty$ for $k = 1$ or $k = 2$, respectively. This restriction is based on the definition of a process of cylinders with non-convex bases, see e.g. Spiess and Spodarev [2011]. In Lemma 4.4 and Lemma 4.5, the cross-section (or base) Ξ_0 of the typical cylinder can be chosen as random compact set satisfying $0 < \mathbb{E}\,|\Xi_0|_1 < \infty$ or $\mathbb{E}\,|\Xi_0|_1^2 < \infty$, respectively.

Central limit theorem in the strong Brillinger-mixing setting

To obtain the asymptotic normality, we need much more strict assumptions, in particular imposed on the reduced factorial cumulant measures of Ψ.

Theorem 4.5.
Assume that there are constants $b > 0, a \geqslant b^{-1}$ such that the reduced factorial cumulant measures of Ψ satisfy $\|\gamma^{[k]}_{red}\|_{TV} \leqslant ab^k$ for all $k \in \mathbb{N}$. Moreover, suppose that $\mathbb{E}\, R_0^k < \infty$ for all $k \in \mathbb{N}$ and that Φ_0 has a continuous distribution function G. Then

$$\rho^{-\frac{3}{2}}(|\Xi \cap K_\rho|_2 - \mathbb{E}\,|\Xi \cap K_\rho|_2) \xrightarrow[\rho\to\infty]{\mathcal{D}} N(0, \sigma_P^2(K, F, G)),$$

where $\sigma_P^2(K, F, G) > 0$ is the asymptotic variance from Theorem 4.4.

The assumption $a \geqslant b^{-1}$ is required so that $\|\gamma^{[1]}_{red}\|_{TV} := 1 \leqslant ab$. Theorem 4.5 generalizes some of the results obtained in Heinrich and Spiess [2013], in particular Theorem 4.1, for stationary Poisson cylinder processes, yet under significant expenses of the generality of the dimension. We believe that the result can be transferred to higher dimensions, but probably not by the approach used here since it is very computationally demanding even for the planar case.

Our main tool in proving Theorem 4.5 is the application of the factorial moment measure expansion, i.e. Theorem 1.4 on the pfg's $G_\Psi[1 - w^{\cup}_{x_1,\dots,x_k}]$ for $k \in \mathbb{N}$ and $x_1, \dots, x_k \in \mathbb{R}^2$. Recall that

$$G_\Psi[1-w] = 1 + \sum_{k=1}^{\infty}\frac{(-1)^k}{k!}\int_{\mathbb{R}^k} w(p_1)\cdots w(p_k)\alpha^{[k]}(\mathrm{d}(p_1,\dots,p_k)) \qquad (4.54)$$

for some function $w : \mathbb{R} \to [0, 1]$. The condition ensuring the convergence of (4.54) for $w = w^{\cup}_{x_1,\ldots,x_k}$ for some $x_1, \ldots, x_k \in \mathbb{R}^2$ is verified by Lemma 4.2.

1. Step in proving Theorem 4.5: Expansion of $\rho^{-9/2}\mathbf{Cum}_3|\Xi \cap \rho K|_2$ into an infinite sum of asymptotically vanishing terms:

Using the factorial moment measure expansion (4.54) of pgf's $G[1-w^{\cup}_{x_1,x_2,x_3}]$, $G[1-w^{\cup}_{x_1,x_2}], G[1-w^{\cup}_{x_1,x_3}], G[1-w^{\cup}_{x_2,x_3}], G[1-w_{x_1}], G[1-w_{x_2}], G[1-w_{x_3}]$ and applying the Cauchy product formula, we obtain

$$\rho^{-9/2}\mathbf{Cum}_3|\Xi \cap \rho K|_2 = \int\limits_K\int\limits_K\int\limits_K \sum_{k=1}^{\infty} \frac{(-1)^k}{k!} T_k^{(\rho)}(x_1,x_2,x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3,$$

where

$$\begin{aligned}
T_k^{(\rho)}(x_1,x_2,x_3) := \rho^{3/2}\Bigg(&\int\limits_{\mathbb{R}^k}\prod_{i=1}^{k} w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_i)\alpha^{[k]}(\mathrm{d}(p_1,\ldots,p_k)) \qquad (4.55)\\
&-\sum_{l=0}^{k}\binom{k}{l}\int\limits_{\mathbb{R}^l}\prod_{i=1}^{l} w_{\rho x_1}(p_i)\alpha^{[l]}(\mathrm{d}(p_1,\ldots,p_l))\int\limits_{\mathbb{R}^{k-l}}\prod_{i=1}^{k-l} w^{\cup}_{\rho x_2,\rho x_3}(p_i)\alpha^{[k-l]}(\mathrm{d}(p_1,\ldots,p_{k-l}))\\
&-\sum_{l=0}^{k}\binom{k}{l}\int\limits_{\mathbb{R}^l}\prod_{i=1}^{l} w_{\rho x_2}(p_i)\alpha^{[l]}(\mathrm{d}(p_1,\ldots,p_l))\int\limits_{\mathbb{R}^{k-l}}\prod_{i=1}^{k-l} w^{\cup}_{\rho x_1,\rho x_3}(p_i)\alpha^{[k-l]}(\mathrm{d}(p_1,\ldots,p_{k-l}))\\
&-\sum_{l=0}^{k}\binom{k}{l}\int\limits_{\mathbb{R}^l}\prod_{i=1}^{l} w_{\rho x_3}(p_i)\alpha^{[l]}(\mathrm{d}(p_1,\ldots,p_l))\int\limits_{\mathbb{R}^{k-l}}\prod_{i=1}^{k-l} w^{\cup}_{\rho x_1,\rho x_2}(p_i)\alpha^{[k-l]}(\mathrm{d}(p_1,\ldots,p_{k-l}))\\
&+2\sum_{l=0}^{k}\sum_{m=0}^{l}\binom{k}{l}\binom{l}{m}\int\limits_{\mathbb{R}^m}\prod_{i=1}^{m} w_{\rho x_1}(p_i)\alpha^{[m]}(\mathrm{d}(p_1,\ldots,p_m))\\
&\times\int\limits_{\mathbb{R}^{l-m}}\prod_{i=1}^{l-m} w_{\rho x_2}(p_i)\alpha^{[l-m]}(\mathrm{d}(p_1,\ldots,p_{l-m}))\int\limits_{\mathbb{R}^{k-l}}\prod_{i=1}^{k-l} w_{\rho x_3}(p_i)\alpha^{[k-l]}(\mathrm{d}(p_1,\ldots,p_{k-l}))\Bigg).
\end{aligned}$$

Lemma 4.6.
Assume that $\mathbb{E}\,R_0^3 < \infty$ and Φ_0 has a continuous distribution function G. Then

$$\lim_{\rho\to\infty}\int\limits_K\int\limits_K\int\limits_K T_1^{(\rho)}(x_1,x_2,x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = 0.$$

Proof. We have that

$$\begin{aligned}
&\frac{1}{\lambda}\int\limits_K\int\limits_K\int\limits_K T_1^{(\rho)}(x_1,x_2,x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\\
&\qquad = \rho^{3/2}\int\limits_K\int\limits_K\int\limits_K\int\limits_{\mathbb{R}}\mathbb{P}\Big(p \in \bigcap_{i=1}^{3}(\Xi_0 + \rho\langle v(\Phi_0), x_i\rangle)\Big)\mathrm{d}p\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3.
\end{aligned}$$

To determine the limit of the right-hand side as $\rho \to \infty$, we rewrite the probability within the integral by means of the expectation over the corresponding indicator function leading to

$$
\begin{aligned}
&\rho^{3/2}\mathbb{E}\int\limits_K\int\limits_K\int\limits_K \mathbf{1}_{\Xi_0+\rho\langle v(\Phi_0),x_1\rangle\cap\Xi_0+\rho\langle v(\Phi_0),x_2\rangle\cap\Xi_0+\rho\langle v(\Phi_0),x_3\rangle}(p)\mathrm{d}p\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\\
=&\rho^{3/2}\mathbb{E}\int\limits_K\int\limits_K\int\limits_K |\Xi_0\cap(\Xi_0+\rho\langle v(\Phi_0),x_2-x_1\rangle)\cap(\Xi_0+\rho\langle v(\Phi_0),x_3-x_1\rangle)|_1\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\\
=&\rho^{3/2}\mathbb{E}\int\limits_{K\oplus(-K)}\int\limits_{K\oplus(-K)} |K\cap(K-y_1)\cap(K-y_2)|_2\\
&\qquad\qquad\times|\Xi_0\cap(\Xi_0+\rho\langle v(\Phi_0),y_1\rangle)\cap(\Xi_0+\rho\langle v(\Phi_0),y_2\rangle)|_1\mathrm{d}y_1\mathrm{d}y_2\\
\leqslant&\rho^{3/2}\mathbb{E}\int\limits_{K\oplus(-K)}\int\limits_{K\oplus(-K)} |K|_2|\Xi_0\cap(\Xi_0+\rho\langle v(\Phi_0),y_1\rangle)\cap(\Xi_0+\rho\langle v(\Phi_0),y_2\rangle)|_1\mathrm{d}y_1\mathrm{d}y_2
\end{aligned}
$$

after the substitution $y_1 = x_2 - x_1, y_2 = x_3 - x_1, y_3 = x_1$ and estimate $|K\cap(K-y_1)\cap(K-y_2)|_2 \leqslant |K|_2$. Further,we put $y_1 = \mathcal{O}(\Phi_0)(u_1,u_2)^T, y_2 = \mathcal{O}(\Phi_0)(v_1,v_2)^T$ (recall (4.9)), so that the last line equals

$$
\begin{aligned}
&\rho^{3/2}|K|_2\mathbb{E}\int\limits_{[\mathcal{O}(-\Phi_0)K\oplus(-K)]^2}|\Xi_0\cap(\Xi_0+\rho u_1)\cap(\Xi_0+\rho v_1)|_1\mathrm{d}(u_1,u_2)\mathrm{d}(v_1,v_2)\\
&\leqslant \rho^{3/2}|K|_2\mathbb{E}\int\limits_{[\mathcal{O}(-\Phi_0)K\oplus(-K)]^2}|\Xi_0|_1\mathbf{1}_{[0,2R_0]}(|\rho u_1|)\mathbf{1}_{[0,2R_0]}(|\rho v_1|)\mathrm{d}(u_1,u_2)\mathrm{d}(v_1,v_2)\\
&= \rho^{-1/2}|K|_2\mathbb{E}\,|\Xi_0|_1\int\limits_{\mathbb{R}^2}\int\limits_{\mathbb{R}^2}\mathbf{1}_{\mathcal{O}(-\Phi_0)K\oplus(-K)}(\frac{z_1}{\rho},u_2)\mathbf{1}_{\mathcal{O}(-\Phi_0)K\oplus(-K)}(\frac{z_2}{\rho},v_2)\\
&\qquad\qquad\times\mathbf{1}_{[0,2R_0]}(|z_1|)\mathbf{1}_{[0,2R_0]}(|z_2|)\mathrm{d}(z_1,u_2)\mathrm{d}(z_2,v_2),
\end{aligned}
$$

where we put $z_1 = \rho u_1, z_2 = \rho v_1$. By integrating w.r.t u_2 and v_2, the latter expression equals

$$
\begin{aligned}
&\rho^{-1/2}|K|_2\mathbb{E}\,|\Xi_0|_1\int\limits_{\mathbb{R}}\int\limits_{\mathbb{R}}|g(\frac{z_1}{\rho},\Phi_0)\cap K\oplus(-K)|_1|g(\frac{z_2}{\rho},\Phi_0)\cap K\oplus(-K)|_1\\
&\qquad\qquad\times\mathbf{1}_{[0,2R_0]}(|z_1|)\cdot\mathbf{1}_{[0,2R_0]}(|z_2|)\mathrm{d}z_1\mathrm{d}z_2\\
&\leqslant |K|_2\rho^{-1/2}(\mathrm{diam}(K\oplus(-K)))^2 32\mathbb{E}\,R_0^3.
\end{aligned}
$$

Since $\mathbb{E}\,R_0^3$ was assumed to be finite, the later term converges to 0 as $\rho\to\infty$.

□

The integration steps used to prove Lemma 4.6 will be repeated several times in the rest of the paper, so some details will be omitted. In particular, it can be shown that, assuming $\mathbb{E}\,|\Xi_0|_1^k < \infty$, we would have

$$
\begin{aligned}
&\lim_{\rho\to\infty}\rho^{k/2}\int\limits_{K^k}\int\limits_{\mathbb{R}} w^{\cap}_{\rho x_1,\ldots,\rho x_k}(p)\mathrm{d}p\mathrm{d}(x_1,\ldots,x_k)\\
&\quad=\lim_{\rho\to\infty}\rho^{k/2}\int\limits_{K^k}\int\limits_{\mathbb{R}}\mathbb{P}(p\in\bigcap_{i=1}^{k}(\Xi_0+\rho\langle v(\Phi_0),x_i\rangle))\mathrm{d}p\mathrm{d}(x_1,\ldots,x_k)\\
&\quad=0.
\end{aligned}
$$

Lemma 4.7.
If $\Psi \sim P$ is Brillinger-mixing, $\mathbb{E}\, R_0^3 < \infty$ and Φ_0 has a continuous distribution function, then
$$\lim_{\rho\to\infty} \int\limits_K \int\limits_K \int\limits_K T_2^{(\rho)}(x_1, x_2, x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3 = 0.$$

Proof. From (4.55), we have for $k = 2$ that

$$\begin{aligned}
&T_2^{(\rho)}(x_1, x_2, x_3)\\
&= \rho^{3/2} \int\limits_{\mathbb{R}^2} \Bigg(\prod_{i=1}^{2} w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_i) - \sum_{\substack{s,r\in\{1,2,3\}\\ s\neq r}} \prod_{i=1}^{2} w^{\cup}_{\rho x_s,\rho x_r}(p_i) + \sum_{s=1}^{3}\prod_{i=1}^{2} w_{\rho x_s}(p_i) \Bigg)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \alpha^{[2]}(\mathrm{d}(p_1,p_2))\\
&+ \rho^{3/2} \int\limits_{\mathbb{R}^2} \Bigg(4 \sum_{\substack{s,r\in\{1,2,3\}\\ s\neq r}} w_s(p_1)w_r(p_2) - 2\sum_{s=1}^{3} w_{\rho x_s}(p_1) w^{\cup}_{\{\rho x_r, r\in\{1,2,3\}\setminus s\}}(p_2) \Bigg)\\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times \alpha^{[1]}(\mathrm{d}p_1)\alpha^{[1]}(\mathrm{d}p_2).
\end{aligned}$$

Using the inclusion-exclusion principle, the fact that $\alpha^{[2]}$ is invariant under permutation of its components and the representations $\alpha^{[1]}(B) = \lambda|B|_1$ and $\alpha^{[2]}(B_1\times B_2) = \gamma^{[2]}(B_1 \times B_2) + \lambda^2|B_1|_1|B_2|_1$, we arrive at

$$\begin{aligned}
&T_2^{(\rho)}(x_1, x_2, x_3)\\
&= \rho^{3/2} \int\limits_{\mathbb{R}^2} \big(w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p_1) w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_2) - 2w_{\rho x_1}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\\
&\qquad - 2w_{\rho x_2}(p_1)w^{\cap}_{\rho x_1,\rho x_3}(p_2) - 2w_{\rho x_3}(p_1)w^{\cap}_{\rho x_1,\rho x_2}(p_2) + 2w^{\cap}_{\rho x_1,\rho x_2}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\\
&\qquad +2w^{\cap}_{\rho x_1,\rho x_2}(p_1)w^{\cap}_{\rho x_1,\rho x_3}(p_2) + 2w^{\cap}_{\rho x_1,\rho x_3}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\big)\, \gamma^{[2]}(\mathrm{d}(p_1,p_2))\\
&+ \lambda^2\rho^{3/2} \int\limits_{\mathbb{R}^2} \big(w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p_1) w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_2) + 2w^{\cap}_{\rho x_1,\rho x_2}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\\
&\qquad +2w^{\cap}_{\rho x_1,\rho x_2}(p_1)w^{\cap}_{\rho x_1,\rho x_3}(p_2) + 2w^{\cap}_{\rho x_1,\rho x_3}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\big)\, \mathrm{d}p_1\mathrm{d}p_2.
\end{aligned}$$

Thus,

$$\begin{aligned}
&\int\limits_K\int\limits_K\int\limits_K T_2^{(\rho)}(x_1, x_2, x_3)\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\\
&= \rho^{3/2} \int\limits_K\int\limits_K\int\limits_K\int\limits_{\mathbb{R}^2} \big(w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p_1) w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_2) - 6w_{\rho x_1}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\\
&\qquad\qquad\qquad\qquad +6w^{\cap}_{\rho x_1,\rho x_2}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\big)\, \gamma^{[2]}(\mathrm{d}(p_1,p_2))\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\\
&+ \lambda^2\rho^{3/2} \int\limits_K\int\limits_K\int\limits_K\int\limits_{\mathbb{R}^2} \big(w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p_1) w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_2) + 6w^{\cap}_{\rho x_1,\rho x_2}(p_1)w^{\cap}_{\rho x_2,\rho x_3}(p_2)\big)\\
&\qquad\qquad\qquad\qquad \times \mathrm{d}p_1\mathrm{d}p_2\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3
\end{aligned}$$

The limit of $\int_K \int_K \int_K T_2^{(\rho)} \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3$ will be evaluated term by term. First, it is a direct consequence of Lemma 4.6 and the Brillinger-mixing property that

$$\lim_{\rho\to\infty} \rho^{3/2} \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p_1) w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_2) \gamma^{[2]}(\mathrm{d}(p_1,p_2)) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = 0$$

and

$$\lim_{\rho\to\infty} \rho^{3/2} \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p_1) w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_2) \mathrm{d}p_1 \mathrm{d}p_2 \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = 0.$$

For the rest of integrals, we use the same integration procedure as in the proof of Lemma 4.6, especially the substitutions. Hence, the details are omitted. First, we assume that $w^{\cap}_{\rho x_1,\rho x_2}(p) = \mathbb{P}(p \in \cap_{i=1,2}(\Xi_1 + \rho\langle v(\Phi_1), x_i\rangle))$ and that $w^{\cap}_{\rho x_2,\rho x_3}(p) = \mathbb{P}(p \in \cap_{i=1,3}(\Xi_2 + \rho\langle v(\Phi_2), x_i\rangle))$, where $\Xi_1 := [-R_1, R_1]$, $\Xi_2 := [-R_2, R_2]$, Φ_1, Φ_2 are mutually independent. Then we arrive at the estimate

$$\begin{aligned}
&\rho^{3/2} \left| \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w^{\cap}_{\rho x_1,\rho x_2}(p_1) w^{\cap}_{\rho x_2,\rho x_3}(p_2) \gamma^{(2)}(\mathrm{d}(p_1,p_2)) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \right| \\
&= \rho^{3/2}\lambda \left| \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w^{\cap}_{\rho x_1,\rho x_2}(p_1) w^{\cap}_{\rho x_2,\rho x_3}(p_1+p_2) \mathrm{d}p_1 \gamma^{[2]}_{red}(\mathrm{d}p_2) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \right| \\
&\leqslant \lambda \rho^{3/2} \mathbb{E} \int\limits_{\mathcal{O}(-\Phi_1)K\oplus(-K)} \int\limits_{\mathcal{O}(-\Phi_2)K\oplus(-K)} |K|_2 \mathbf{1}_{\Xi_1\cap(\Xi_1+\rho u_1)}(p_1) \\
&\qquad\qquad \times \mathbf{1}_{\Xi_2\cap(\Xi_2+\rho v_1)}(p_1+p_2) \mathrm{d}p_1 |\gamma^{[2]}_{red}|(\mathrm{d}(u_1,u_2)) \mathrm{d}(v_1,v_2) \\
&\leqslant \lambda \rho^{-1/2} |K|_2 \mathrm{diam}(K\oplus(-K))^2 \mathbb{E}|\Xi_1|_1 \|\gamma^{[2]}_{red}\|_{TV} \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} \mathbf{1}_{[0,2R_1]}(|z_1|) \mathbf{1}_{[0,2R_2]}(|z_2|) \mathrm{d}z_1 \mathrm{d}z_2
\end{aligned}$$

which tends to 0 as $\rho \to \infty$ having $\mathbb{E}\, R_0^2 < \infty$. Similarly,

$$\begin{aligned}
&\rho^{3/2} \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w^{\cap}_{\rho x_1,\rho x_2}(p_1) w^{\cap}_{\rho x_1,\rho x_3}(p_2) \mathrm{d}p_1 \mathrm{d}p_2 \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \\
&\leqslant \rho^{-1/2} |K|_2 \mathrm{diam}(K\oplus(-K))^2 \mathbb{E}\, |\Xi_1|_1 |\Xi_2|_1 \int\limits_{\mathbb{R}} \int\limits_{\mathbb{R}} \mathbf{1}_{[0,2R_1]}(|z_1|) \mathbf{1}_{[0,2R_2]}(|z_2|) \mathrm{d}z_1 \mathrm{d}z_2
\end{aligned} \tag{4.56}$$

and

$$\begin{aligned}
&\rho^{3/2} \left| \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w_{\rho x_1}(p_1) w^{\cap}_{\rho x_2,\rho x_3}(p_2) \gamma^{[2]}(\mathrm{d}(p_1,p_2)) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \right| \\
&= \lambda \rho^{3/2} \left| \int\limits_K \int\limits_K \int\limits_K \int\limits_{\mathbb{R}^2} w_{\rho x_1}(p_1) w^{\cap}_{\rho x_2,\rho x_3}(p_1+p_2) \mathrm{d}p_1 \gamma^{[2]}_{red}(\mathrm{d}p_2) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \right| \\
&\leqslant \lambda \rho^{-1/2} |K|_2 \mathrm{diam}(K) \mathrm{diam}(K\oplus(-K)) 2\mathbb{E}\, |\Xi_1|_1 |\Xi_2|_1^2 \gamma^{[2]}_{red}(\mathbb{R}).
\end{aligned} \tag{4.57}$$

Hence, expressions (4.56) and (4.57) converge to 0 with $\rho \to \infty$ and consequently,

$$\int_K \int_K \int_K T_2^{(\rho)}(x_1, x_2, x_3) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \xrightarrow[\rho\to\infty]{} 0.$$

□

Before stating a general result for arbitrary $k \geqslant 3$, we need to introduce some notation and other supplementary results. Denote by $\mathbf{S}$ the space of all measurable functions $e : \mathbb{R} \to \mathbb{R}$ and define operators $f, f_{S'}, f^{(k)}, k \in \mathbb{N}, S' \subset \mathbf{S}$ on $\mathbb{N} \times \mathbf{S}$, such that for $n \in \mathbb{N}$ and $S \subset \mathbf{S}$

$$f(n, S) \mapsto f(n, S)(p_1, \dots, p_n) := \sum_{(e_1,\dots,e_n)\in S} \prod_{i=1}^{n} e_i(p_i), \quad (p_1, \dots, p_n) \in \mathbb{R}^n,$$

$$f^{(k)}(n, S) \mapsto f^{(k)}(n, S)(p_1, \dots, p_n) := \sum_{\substack{(e_1,\dots,e_n)\in S \\ \exists e_{i_1} \neq \cdots \neq e_{i_k}}} \prod_{i=1}^{n} e_i(p_i), \quad (p_1, \dots, p_n) \in \mathbb{R}^n,$$

$$f_{S'}(n, S) \mapsto f_{S'}(n, S)(p_1, \dots, p_n) := \sum_{\substack{(e_1,\dots,e_n)\in S\cup S' \\ S' \subset \cup e_i}} \prod_{i=1}^{n} e_i(p_i), \quad (p_1, \dots, p_n) \in \mathbb{R}^n.$$

Note that $f^{(k)}(n, S) \equiv 0$ whenever $k > n$ and $f_{S'}(n, S) \equiv 0$ whenever $\#S' > n$.

Moreover, we define functions $g, g_{S'}, g^{(k)} : \mathbb{N} \times \mathbf{S} \to \mathbb{R}, k \in \mathbb{N}, S' \subset \mathbf{S}$ such that for $n \in \mathbb{N}$ and $S \subset \mathbf{S}$

$$g(n, S) = \int_{\mathbb{R}^n} f(n, S)(p_1, \dots, p_n) \gamma^{[n]} \mathrm{d}(p_1, \dots, p_n),$$

$$g^{(k)}(n, S) = \int_{\mathbb{R}^n} f^{(k)}(n, S)(p_1, \dots, p_n) \gamma^{[n]} \mathrm{d}(p_1, \dots, p_n),$$

$$g_{S'}(n, S) = \int_{\mathbb{R}^n} f_{S'}(n, S)(p_1, \dots, p_n) \gamma^{[n]} \mathrm{d}(p_1, \dots, p_n).$$

It is easy to see that for $k, n \in \mathbb{N}$ and $e_1, \dots, e_k \in \mathbf{S}$,

$$f(n, \{e_1, \dots, e_k\}) = \sum_{i=1}^{k} f(n, e_i) + \sum_{i\neq j} f^{(2)}(n, \{e_i, e_j\}) + \cdots + f^{(k)}(n, \{e_1, \dots, e_k\}) \tag{4.58}$$

and

$$f(n, e_1 + \cdots + e_k) = f(n, \{e_1, \dots, e_k\}). \tag{4.59}$$

Similar relations hold for function g.

From now on, we will refer to a universal constant denoted by $C^{(\rho)}(x_1, x_2, x_3)$ such that

$$\rho^{3/2} \int_K \int_K \int_K C^{(\rho)}(x_1, x_2, x_3) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \to 0 \tag{4.60}$$

as $\rho \to \infty$.

Lemma 4.8.
Let $n, k \in \mathbb{N}$ and E, E' be any subsets of $\{w_{\rho x_r}, w^{\cap}_{\rho x_r,\rho x_s}, w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}\}_{r,s=1,2,3}$. Then, under the assumptions of Lemma 4.7, we have that

1. $g_{w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}}(n,E) = C^{(\rho)}(x_1,x_2,x_3)$,
2. $g_{w^{\cap}_{\rho x_r,\rho x_s}}(n,E) g_{w^{\cap}_{\rho x_s,\rho x_t}}(k,E') = C^{(\rho)}(x_1,x_2,x_3), \quad r \neq s, s \neq t$,
3. $g_{w^{\cap}_{\rho x_r,\rho x_s}, w^{\cap}_{\rho x_s,\rho x_t}}(n,E) = C^{(\rho)}(x_1,x_2,x_3), \quad r \neq s, s \neq t, r \neq t$,
4. $g_{w_{\rho x_r}, w^{\cap}_{\rho x_s,\rho x_t}}(n,E) = C^{(\rho)}(x_1,x_2,x_3), \quad r \neq s, s \neq t, r \neq t$,
5. $g_{w_{\rho x_1}, w_{\rho x_2}, w_{\rho x_3}}(n,E) = C^{(\rho)}(x_1,x_2,x_3)$,
6. $g_{w_{\rho x_r}, w_{\rho x_s}}(n,E) g_{w^{\cap}_{\rho x_s,\rho x_t}}(k,E') = C^{(\rho)}(x_1,x_2,x_3), \quad r \neq s, s \neq t$.

Proof. 1. It is enough to show for $n = 1$. This case, however, was shown in the proof of Lemma 4.6.

2.-3. For $n = 1$, it was shown in the proof of Lemma 4.7 (see (4.56)).

5. For $n = 1, 2$ the term $g_{w_{\rho x_1}, w_{\rho x_2}, w_{\rho x_3}}(n,E)$ is equal to 0 by definition. We will show the result for $n = 3$:

$$\rho^{3/2}\frac{1}{\lambda}\left|\int\limits_{K^3}\int\limits_{\mathbb{R}^3} w_{\rho x_1}(p_1) w_{\rho x_2}(p_2) w_{\rho x_3}(p_3)\gamma^{[3]}(\mathrm{d}(p_1,p_2,p_3))\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3\right|$$
$$\leqslant \mathbb{E}\,\rho^{3/2}\int\limits_{K^3}\int\limits_{\mathbb{R}^3} \mathbf{1}_{\Xi_1+\rho\langle v(\Phi_1),x_1\rangle}(p_1)\mathbf{1}_{\Xi_2+\rho\langle v(\Phi_2),x_1\rangle}(p_1+p_2)\mathbf{1}_{\Xi_3+\rho\langle v(\Phi_3),x_1\rangle}(p_1+p_3)$$
$$\times\, \mathrm{d}p_1|\gamma^{[3]}_{red}|(\mathrm{d}(p_2,p_3))\mathrm{d}x_1\mathrm{d}x_2\mathrm{d}x_3$$
$$= \mathbb{E}\,\rho^{3/2}\int\limits_{\times_{i=1}^3\mathcal{O}(-\Phi_i)K}\int\limits_{\mathbb{R}^3} \mathbf{1}_{\Xi_1+\rho u_1}(p_1)\mathbf{1}_{\Xi_2+\rho v_1}(p_1+p_2)\mathbf{1}_{\Xi_3+\rho w_1}(p_1+p_3)$$
$$\times\, \mathrm{d}p_1|\gamma^{[3]}_{red}|(\mathrm{d}(p_2,p_3))\mathrm{d}(u_1,u_2)\mathrm{d}(v_1,v_2)\mathrm{d}(w_1,w_2)$$
$$= \mathbb{E}\,\rho^{3/2}\int\limits_{\times_{i=1}^3\mathcal{O}(-\Phi_i)K}\int\limits_{\mathbb{R}^3} \mathbf{1}_{\Xi_1+\rho(u_1-w_1)}(p_1)\mathbf{1}_{\Xi_2+\rho(v_1-w_1)}(p_1+p_2)\mathbf{1}_{\Xi_3}(p_1+p_3)$$
$$\times\, \mathrm{d}p_1|\gamma^{[3]}_{red}|(\mathrm{d}(p_2,p_3))\mathrm{d}(u_1,u_2)\mathrm{d}(v_1,v_2)\mathrm{d}(w_1,w_2)$$
$$\leqslant \mathrm{diam}(K)^2\mathbb{E}\,\rho^{3/2}\int\limits_{\mathbb{R}^3}\int\limits_{\mathbb{R}^3} |g(w_1,\Phi_3)\cap K|_1\mathbf{1}_{\Xi_1+\rho(u_1-w_1)}(p_1)\mathbf{1}_{\Xi_2+\rho(v_1-w_1)}(p_1+p_2)$$
$$\times\, \mathbf{1}_{\Xi_3}(p_1+p_3)\mathrm{d}p_1|\gamma^{[3]}_{red}|(\mathrm{d}(p_2,p_3))\mathrm{d}u_1\mathrm{d}v_1\mathrm{d}w_1$$
$$= \mathrm{diam}(K)^2\mathbb{E}\,\rho^{-1/2}\int\limits_{\mathbb{R}^3}\int\limits_{\mathbb{R}^3} |g(z_1/\rho+w_1,\Phi_3)\cap K|_1\mathbf{1}_{\Xi_1+z_1}(p_1)\mathbf{1}_{\Xi_2+z_2}(p_1+p_2)$$
$$\times\, \mathbf{1}_{\Xi_3}(p_1+p_3)\mathrm{d}p_1|\gamma^{[3]}_{red}|(\mathrm{d}(p_2,p_3))\mathrm{d}z_1\mathrm{d}z_2\mathrm{d}w_1$$

$$\leqslant \operatorname{diam}(K)^3 \mathbb{E}\, \rho^{-1/2} \int_{\mathbb{R}^2} \int_{\mathbb{R}^3} \mathbf{1}_{\Xi_1+z_1}(p_1) \mathbf{1}_{\Xi_2+z_2}(p_1+p_2) \times \mathbf{1}_{\Xi_3}(p_1+p_3) \mathrm{d}p_1 |\gamma_{red}^{[3]}|(\mathrm{d}(p_2,p_3)) \mathrm{d}z_1 \mathrm{d}z_2$$

$$= \rho^{-1/2} \operatorname{diam}(K)^3 (\mathbb{E}\,|\Xi_0|_1)^3 \|\gamma_{red}^{[3])}\|_{TV} \xrightarrow{\rho\to\infty} 0.$$

6. We will show it for $n = 2, k = 1$. By following the same steps as in the proof of the previous point combined with the steps in the proof of Lemma 4.6, we arrive at

$$\int_K \int_K \int_K g_{w_{\rho x_r}, w_{\rho x_s}}(2, E) g_{w^{\cap}_{\rho x_s, \rho x_t}}(1, E') \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3$$
$$\leqslant \rho^{-1/2} 2\lambda (\operatorname{diam}(K))^2 \operatorname{diam}(K \oplus (-K)) \|\gamma_{red}^{[2]}\|_{TV} \mathbb{E}\,|\Xi_1|_1 |\Xi_2|_1 |\Xi_3|_1 \xrightarrow{\rho\to\infty} 0.$$

The case of general $n, k \in \mathbb{N}$ can be treated similarly for 1. – 6. Let us demonstrate on the case 5. for $n \geqslant 3$ and $e_1, \dots, e_n \subset \{w_{\rho x_r}, w^{\cap}_{\rho x_r, \rho x_s}, w^{\cap}_{\rho x_1, \rho x_2, \rho x_3}\}_{r,s=1,2,3}$ such that $\{w_{\rho x_1}, w_{\rho x_2}, w_{\rho x_3}\} \subset \{e_1, \dots, e_n\}$. Since $e_i, i = 1, \dots, n$ are probabilities, we come to the following estimate by following the exact steps as for $n = 3$:

$$\rho^{3/2} \left| \int_{K^3} \int_{\mathbb{R}^n} e_1(p_1) \cdots e_n(p_n) \gamma^{[n]}(\mathrm{d}(p_1, \dots, p_n)) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \right|$$
$$\leqslant \rho^{3/2} \left| \int_{K^3} \int_{\mathbb{R}^n} w_{\rho x_1}(p_1) w_{\rho x_2}(p_2) w_{\rho x_3}(p_3) \gamma^{[n]}(\mathrm{d}(p_1, \dots, p_n)) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 \right|$$
$$\leqslant \rho^{-1/2} \lambda \operatorname{diam}(K)^4 (\mathbb{E}\,|\Xi_0|_1)^3 \|\gamma_{red}^{[n]}\|_{TV}.$$

The latter expression goes to 0 with $\rho \to \infty$ since we assumed $\|\gamma_{red}^{[n]}\|_{TV} < \infty$. □

Lemma 4.9.
Under the assumptions of Lemma 4.7,

$$\lim_{\rho\to\infty} \int_K \int_K \int_K T_k^{(\rho)}(x_1, x_2, x_3) \mathrm{d}x_1 \mathrm{d}x_2 \mathrm{d}x_3 = 0, \qquad k \geqslant 3.$$

Proof. By expressing the factorial moment measures $\alpha^{[k]}, k \in \mathbb{N}$ in terms of factorial cumulant measures $\gamma^{[k]}$ as in (1.5), we have

$$\rho^{-3/2} T_k^{(\rho)}(x_1, x_2, x_3)$$
$$= \sum_{n=1}^{k} \sum_{\substack{K_1 \cup \cdots \cup K_n \\ = \{1,\dots,k\}}} \prod_{j=1}^{n} \int_{\mathbb{R}^{\# K_j}} \prod_{i \in K_j} w^{\cup}_{\rho x_1, \rho x_2, \rho x_3}(p_i) \gamma^{[\# K_j]}(\mathrm{d}(p_i, i \in K_j))$$

$$-\sum_{r=1}^{3}\sum_{l=0}^{k}\binom{k}{l}\sum_{n_1=1}^{l}\sum_{\substack{K_1\cup\cdots\cup K_{n_1}\\=\{1,\dots,l\}}}\prod_{j=1}^{n_1}\int_{\mathbb{R}^{\#K_j}}\prod_{i\in K_j}w_{\rho x_r}(p_i)\gamma^{[\#K_j]}(\mathrm{d}(p_i,i\in K_j))$$

$$\times\sum_{n_2=1}^{k-l}\sum_{\substack{K_1\cup\cdots\cup K_{n_2}\\=\{1,\dots,k-l\}}}\prod_{j=1}^{n_2}\int_{\mathbb{R}^{\#K_j}}\prod_{i\in K_j}w^{\cup}_{\{\rho x_s,s\neq r\}}(p_i)\gamma^{[\#K_j]}(\mathrm{d}(p_i,i\in K_j))$$

$$+2\sum_{l=0}^{k}\sum_{m=0}^{l}\binom{k}{l}\binom{l}{m}\sum_{n_1=1}^{m}\sum_{\substack{K_1\cup\cdots\cup K_{n_1}\\=\{1,\dots,m\}}}\prod_{j=1}^{n_1}\int_{\mathbb{R}^{\#K_j}}\prod_{i\in K_j}w_{\rho x_1}(p_i)\gamma^{[\#K_j]}(\mathrm{d}(p_i,i\in K_j))$$

$$\times\sum_{n_2=1}^{l-m}\sum_{\substack{K_1\cup\cdots\cup K_{n_2}\\=\{1,\dots,l-m\}}}\prod_{j=1}^{n_2}\int_{\mathbb{R}^{\#K_j}}\prod_{i\in K_j}w_{\rho x_2}(p_i)\gamma^{[\#K_j]}(\mathrm{d}(p_i,i\in K_j))$$

$$\times\sum_{n_3=1}^{k-l}\sum_{\substack{K_1\cup\cdots\cup K_{n_3}\\=\{1,\dots,k-l\}}}\prod_{j=1}^{n_3}\int_{\mathbb{R}^{\#K_j}}\prod_{i\in K_j}w_{\rho x_3}(p_i)\gamma^{[\#K_j]}(\mathrm{d}(p_i,i\in K_j)).$$

Next, we use the inclusion-exclusion principle to rewrite $w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p)=w_{\rho x_1}(p)+w_{\rho x_2}(p)+w_{\rho x_3}(p)-w^{\cap}_{\rho x_1,\rho x_2}(p)-w^{\cap}_{\rho x_1,\rho x_3}(p)-w^{\cap}_{\rho x_2,\rho x_3}(p)+w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}(p)$ and $w^{\cup}_{\rho x_r,\rho x_s}=w_{\rho x_r}+w_{\rho x_s}-w^{\cap}_{\rho x_r,\rho x_s}$, $r,s\in\{1,2,3\}$. Further, we denote

$$S:=\{w_{\rho x_1},w_{\rho x_2},w_{\rho x_3},-w^{\cap}_{\rho x_1,\rho x_2},-w^{\cap}_{\rho x_1,\rho x_3},-w^{\cap}_{\rho x_2,\rho x_3},w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}\},$$
$$S':=\{w_{\rho x_1},w_{\rho x_2},w_{\rho x_3},-w^{\cap}_{\rho x_1,\rho x_2},-w^{\cap}_{\rho x_1,\rho x_3},-w^{\cap}_{\rho x_2,\rho x_3}\}.$$

Then using the relations (4.58) and (4.59),

$$T_k^{(\rho)}(x_1,x_2,x_3)=\rho^{3/2}\sum_{\substack{\cup K_i=\{1,\dots,k\}\\L=\{K_1,\dots\}}}(I_1^L-\sum_{r=1}^{3}I_{2,r}^L+2I_3^L)(x_1,x_2,x_3),$$

where

$$I_1^L(x_1,x_2,x_3):=\prod_{K\in L}\int_{\mathbb{R}^{\#K}}\prod_{i\in K}w^{\cup}_{\rho x_1,\rho x_2,\rho x_3}(p_i)\gamma^{[\#K]}(\mathrm{d}(p_i,i\in K))=\prod_{K\in L}g(\#K,S)$$

$$=\prod_{K\in L}\Bigg\{\sum_{r=1}^{3}g(\#K,w_{\rho x_r})+\sum_{r\neq s}g^{(2)}(\#K,\{w_{\rho x_r},w_{\rho x_s}\})+g^{(3)}(\#K,\{w_{\rho x_1},w_{\rho x_2},w_{\rho x_3}\})$$
$$+\sum_{r\neq s}g(\#K,-w^{\cap}_{\rho x_r,\rho x_s})+\sum_{r\neq s}g^{(2)}(\#K,\{w_{\rho x_r},-w^{\cap}_{\rho x_r,\rho x_s}\})$$
$$+\sum_{r\neq s,s\neq t,r\neq t}g^{(2)}(\#K,\{w_{\rho x_r},-w^{\cap}_{\rho x_s,\rho x_t}\})+\sum_{r\neq s,s\neq t,r\neq t}g^{(2)}(\#K,\{-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})$$
$$+\sum_{r\neq s}g^{(3)}(\#K,\{w_{\rho x_r},w_{\rho x_s},-w^{\cap}_{\rho x_r,\rho x_s}\})+\sum_{r\neq s,s\neq t,r\neq t}g^{(3)}(\#K,\{w_{\rho x_r},w_{\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})$$
$$+\sum_{r\neq s,s\neq t,r\neq t}g^{(3)}(\#K,\{w_{\rho x_r},-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})$$
$$+\sum_{r\neq s,s\neq t,r\neq t}g^{(3)}(\#K,\{w_{\rho x_s},-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})+g^{(4)}(\#K,S')$$
$$+g_{w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}}(\#K,S)\Bigg\}$$

$$
\begin{aligned}
=&f\,\big(\#L,\{g(\cdot,w_{\rho x_r})_{r=1,2,3},g^{(2)}(\cdot,\{w_{\rho x_r},w_{\rho x_s}\})_{r\neq s},g(\cdot,-w^{\cap}_{\rho x_r,\rho x_s})_{r\neq s},\\
&g^{(2)}(\cdot,\{w_{\rho x_r},-w^{\cap}_{\rho x_r,\rho x_s}\})_{r\neq s}\}),g^{(3)}(\cdot,\{w_{\rho x_r},w_{\rho x_s},-w^{\cap}_{\rho x_r,\rho x_s}\})_{r\neq s},\\
&g^{(3)}(\cdot,\{w_{\rho x_1},w_{\rho x_2},w_{\rho x_3}\}),g^{(2)}(\cdot,\{w_{\rho x_r},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s\neq t}\}),\\
&g^{(3)}(\cdot,\{w_{\rho x_r},w_{\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq r,r\neq t},g^{(2)}(\cdot,\{w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq t}),\\
&g^{(3)}(\cdot,\{w_{\rho x_r},-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq t,r\neq t},\\
&g^{(3)}(\cdot,\{w_{\rho x_s},-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq t,r\neq t},g^{(4)}(\cdot,S'),\\
&g_{w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}}(\cdot,S)\}\big)\,(\#K,K\in L).
\end{aligned}
$$

Similarly,

$$
\begin{aligned}
I^L_{2,r}(x_1,x_2,x_3):=&\sum_{L_1\cup L_2=L}\prod_{K\in L_1}\int_{\mathbb{R}^{\#K}}\prod_{i\in K}w_{\rho x_r}(p_i)\gamma^{[\#K]}(\mathrm{d}(p_i,i\in K))\\
&\times\prod_{K'\in L_2}\int_{\mathbb{R}^{\#K'}}\prod_{i\in K'}w^{\cup}_{\{\rho x_s,s\neq r\}}(p_i)\gamma^{[\#K']}(\mathrm{d}(p_i,i\in K'))\\
=f\Big(&\#L,\{g(\cdot,w_{\rho x_s})_{s=1,2,3},g(\cdot,-w^{\cap}_{\rho x_s,\rho x_t})_{s\neq t,s\neq r,t\neq r},\\
&g^{(2)}(\cdot,\{w_{\rho x_s},w_{\rho x_t}\})_{s\neq t,s\neq r,t\neq r},g^{(2)}(\cdot,\{w_{\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{s\neq t\neq r}\},\\
&g^{(3)}(\cdot,\{w_{\rho x_s},w_{\rho x_t},-w^{\cap}_{\rho x_s,\rho x_t}\})_{s\neq t,s\neq r,t\neq r}\Big)(\#K,K\in L)
\end{aligned}
$$

and

$$
\begin{aligned}
I^L_3(x_1,x_2,x_3):=&\sum_{L_1\cup L_2\cup L_3=L}\prod_{K\in L_1}\int_{\mathbb{R}^{\#K}}\prod_{i\in K}w_{\rho x_1}(p_i)\gamma^{[\#K]}(\mathrm{d}(p_i,i\in K))\\
&\cdot\prod_{K'\in L_2}\int_{\mathbb{R}^{\#K'}}\prod_{i\in K'}w_{\rho x_2}(p_i)\gamma^{[\#K']}(\mathrm{d}(p_i,i\in K'))\\
&\cdot\prod_{K''\in L_3}\int_{\mathbb{R}^{\#K''}}\prod_{i\in K''}w_{\rho x_3}(p_i)\gamma^{[\#K'']}(\mathrm{d}(p_i,i\in K''))\\
=&\ f(\#L,\{g(\cdot,w_{\rho x_r})_{r=1,2,3}\})(\#K,K\in L).
\end{aligned}
$$

We shall fix a set L and study the expression $I^L_1-\sum_{r=1}^3 I^L_{2,r}+2I^L_3$. Denote by $\mathcal{S}$ the set of functions in the argument of the function f in the latter expression for I^L_1, i.e

$$
\begin{aligned}
\mathcal{S}:=&\{g(\cdot,w_{\rho x_r})_{r=1,2,3},g^{(2)}(\cdot,\{w_{\rho x_r},w_{\rho x_s}\})_{r\neq s},g(\cdot,-w^{\cap}_{\rho x_r,\rho x_s})_{r\neq s},\\
&g^{(2)}(\cdot,\{w_{\rho x_r},-w^{\cap}_{\rho x_r,\rho x_s}\})_{r\neq s}\}),g^{(3)}(\cdot,\{w_{\rho x_r},w_{\rho x_s},-w^{\cap}_{\rho x_r,\rho x_s}\})_{r\neq s},\\
&g^{(3)}(\cdot,\{w_{\rho x_1},w_{\rho x_2},w_{\rho x_3}\}),g^{(2)}(\cdot,\{w_{\rho x_r},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s\neq t}\}),\\
&g^{(3)}(\cdot,\{w_{\rho x_r},w_{\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq r,r\neq t},g^{(2)}(\cdot,\{w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq t}),\\
&g^{(3)}(\cdot,\{w_{\rho x_r},-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq t,r\neq t},\\
&g^{(3)}(\cdot,\{w_{\rho x_s},-w^{\cap}_{\rho x_r,\rho x_s},-w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s,s\neq t,r\neq t},g^{(4)}(\cdot,S'),g_{w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}}(\cdot,S)\}
\end{aligned}
$$

and, moreover,

$$\begin{aligned}\mathcal{S}' := \{&g^{(3)}(\cdot,\{w_{\rho x_1}, w_{\rho x_2}, w_{\rho x_3}\}), g^{(2)}(\cdot,\{w_{\rho x_r}, -w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s\neq t}\}),\\ &g^{(3)}(\cdot,\{w_{\rho x_r}, w_{\rho x_s}, -w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s, s\neq r, r\neq t}, g^{(2)}(\cdot,\{w^{\cap}_{\rho x_r,\rho x_s}, -w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s, s\neq t}),\\ &g^{(3)}(\cdot,\{w_{\rho x_r}, -w^{\cap}_{\rho x_r,\rho x_s}, -w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s, s\neq t, r\neq t},\\ &g^{(3)}(\cdot,\{w_{\rho x_s}, -w^{\cap}_{\rho x_r,\rho x_s}, -w^{\cap}_{\rho x_s,\rho x_t}\})_{r\neq s, s\neq t, r\neq t}, g^{(4)}(\cdot, S'), g_{w^{\cap}_{\rho x_1,\rho x_2,\rho x_3}}(\cdot, S)\}.\end{aligned}$$

Then, the remaining terms of $I_1^L - \sum_{r=1}^3 I_{2,r}^L + 2I_3^L$ can be estimated by

$$\begin{aligned}\left|I_1^L - \sum_{r=1}^3 I_{2,r}^L + 2I_3^L\right| \leqslant &\sum_{r\neq s, s\neq t} f_{g_{w^{\cap}_{\rho x_r,\rho x_s}}(\cdot,S), g_{w^{\cap}_{\rho x_s,\rho x_t}}(\cdot,S)}(\#L,\mathcal{S})(\#K, K\in L)\\ &+\sum_{r\neq s, s\neq t} f_{g_{w_{\rho x_r}, w_{\rho x_s}}(\cdot,S), g_{w^{\cap}_{\rho x_s,\rho x_t}}(\cdot,S)}(\#L,\mathcal{S})(\#K, K\in L)\\ &+\sum_{g\in\mathcal{S}'} f_g(\#L,\mathcal{S})(\#K, K\in L).\end{aligned}$$

According to Lemma 4.8, the right-hand side consists of finitely many terms being equal to $C^{(\rho)}(x_1,x_2,x_3)$, where this universal constant is defined by (4.60).

□

2. Step in proving Theorem 4.5: Expansion of $\rho^{-3k/2}\mathbf{Cum}_k|\Xi\cap\rho K|_2$ into an infinite sum of asymptotically vanishing terms:

We will follow the exact strategy as for proving the convergence of the terms of the third order cumulant. Henceforth, we will skip some of the details. First, using the factorial moment measure expansion of pgf's of the type $G_\Psi[1-w^{\cup}_Q]$, $Q\subset\{\rho x_1,\dots,\rho x_k\}$ and Cauchy product of i infinite series, $i=2,\dots,k$, we arrive at the expression

$$\begin{aligned}&\rho^{-\frac{3k}{2}}\mathbf{Cum}_k(|\Xi\cap\rho K|_2)\\ =&\rho^{\frac{k}{2}}\int_{K^k}\sum_{l=1}^k(-1)^{l-1}(l-1)!\sum_{\substack{Q_1\cup\dots\cup Q_l\\=\{\rho x_1,\dots,\rho x_k\}}}\prod_{j=1}^l G_P[1-w^{\cup}_{Q_j}]\mathrm{d}x_1\dots\mathrm{d}x_k\\ =&\rho^{\frac{k}{2}}\int_{K^k}\sum_{m=1}^\infty\frac{(-1)^m}{m!}\sum_{l=1}^k(-1)^{l-1}(l-1)!\sum_{\substack{Q_1\cup\dots\cup Q_l\\=\{\rho x_1,\dots,\rho x_k\}}}\sum_{n_1=0}^m\sum_{n_2=0}^{n_1}\dots\sum_{n_{l-1}=0}^{n_{l-2}}\binom{m}{n_1}\binom{n_1}{n_2}\dots\binom{n_{l-2}}{n_{l-1}}\\ &\qquad\times\int_{\mathbb{R}^{m-n_1}}\prod_{i=1}^{m-n_1}w^{\cup}_{Q_1}(p_i)\alpha^{[m-n_1]}(\mathrm{d}(p_1,\dots,p_{m-n_1}))\\ &\qquad\times\int_{\mathbb{R}^{n_1-n_2}}\prod_{i=1}^{n_1-n_2}w^{\cup}_{Q_2}(p_i)\alpha^{[n_1-n_2]}(\mathrm{d}(p_1,\dots,p_{n_1-n_2}))\\ &\qquad\times\dots\times\int_{\mathbb{R}^{n_{l-1}}}\prod_{i=1}^{n_{l-1}}w^{\cup}_{Q_l}(p_i)\alpha^{[n_{l-1}]}(\mathrm{d}(p_1,\dots,p_{n_{l-1}}))\mathrm{d}x_1\dots\mathrm{d}x_k\end{aligned}$$

$$=\rho^{\frac{k}{2}}\int\limits_{K^k}\sum_{m=1}^{\infty}\frac{(-1)^m}{m!}\sum_{l=1}^{k}(-1)^{l-1}(l-1)!\sum_{\substack{Q_1\cup\dots\cup Q_l\\=\{\rho x_1,\dots,\rho x_k\}}}\sum_{\substack{\cup K_i=\{1,\dots,m\}\\ L=\{K_1,\dots\}}}\sum_{L_1\cup\dots\cup L_l=L}$$

$$\times\prod_{j=1}^{l}\left(\prod_{K\in L_j}\int\limits_{\mathbb{R}^{\#K}}\prod_{i\in K}w^{\cup}_{Q_l}(p_i)\gamma^{[\#K]}(\mathrm{d}(p_i,i\in K))\right)\mathrm{d}x_1\dots\mathrm{d}x_k$$

$$=\int\limits_{K^k}\sum_{m=1}^{\infty}\frac{(-1)^m}{m!}T^{(\rho)}_{m,k}(x_1,\dots,x_k)\mathrm{d}x_1\dots\mathrm{d}x_k,$$

where

$$T^{(\rho)}_{m,k}(x_1,\dots,x_k):=\rho^{\frac{k}{2}}\sum_{\substack{\cup K_i=\{1,\dots,m\}\\ L=\{K_1,\dots\}}}\sum_{l=1}^{k}(-1)^{l-1}(l-1)!$$
$$\times\sum_{\substack{Q_1\cup\dots\cup Q_l\\=\{\rho x_1,\dots,\rho x_k\}}}f(\#L,\{g(\cdot,\{(-1)^{\#q+1}w^{\cap}_q\}_{q\subset Q_i})\}_{i=1,\dots,l})(\#K\in L).\tag{4.61}$$

Lemma 4.10.
Let $k\geqslant 2$. Assume $\Psi\sim P$ is Brillinger-mixing, $\mathbb{E}\,R_0^k<\infty$ and Φ_0 has a continuous distribution function G. Then

$$\lim_{\rho\to\infty}\int\limits_{K^k}T^{(\rho)}_{m,k}(x_1,\dots,x_k)\mathrm{d}x_1\dots\mathrm{d}x_k=0,\qquad\forall m\geqslant 1.$$

Proof. The strategy is to fix a given partition L in (4.61) and study which of the summands cancel out. Denote by

$$S:=\{g^{(n)}(\cdot,\{(-1)^{\#q_1+1}w^{\cap}_{q_1},\dots,(-1)^{\#q_n+1}w^{\cap}_{q_n}\}),n\in\mathbb{N},q_1,\dots,q_n\subset\{\rho x_1,\dots,\rho x_k\}\}$$
$$=\left\{\int\limits_{\mathbb{R}^n}(-1)^{\#q_1+1}w^{\cap}_{q_1}(p_1)\cdots(-1)^{\#q_n+1}w^{\cap}_{q_n}(p_n)\gamma^{[n]}_{red}(\mathrm{d}(p_1,\dots,p_n)),\right.$$
$$\left.n\in\mathbb{N},q_1,\dots,q_n\subset\{\rho x_1,\dots,\rho x_k\}\right\}$$

the set of all possible terms that appear in the operators f in $T^{(\rho)}_{m,k}$. Let us take arbitrary functions $e_1,\dots,e_{\#L}\in S$. We will study how many times the term $f^{(\#L)}(\#L,\{e_1,\dots,e_{\#L}\})$ $(\#K,K\in L)$ appears in (4.61) and whether, under certain conditions, cancels out. Before that, let us show the simplest example of the choice $e_1,\dots,e_{\#L}$ to give the reader better understanding about the ideas.

Example. Let $e_j\in\{g(\cdot,w_{\rho x_i}),i=1,\dots,k\},j=1,\dots,\#L$. The term $f(\#L,\{e_1,\dots,e_{\#L}\})(\#K\in L)$ is present in all the summands (i.e. exist for all partitions $Q_1\cup\dots\cup Q_l=\{\rho x_1,\dots,\rho x_k\},l=1,\dots,k$). Hence, for fixed L, $T^{(\rho)}_{m,k}(x_1,\dots,x_k)$ contains

$$\sum_{l=1}^{k}(-1)^{l-1}(l-1)!\sum_{\substack{Q_1\cup\dots\cup Q_l\\=\{\rho x_1,\dots,\rho x_k\}}}f(\#L,\{e_1,\dots,e_{\#L}\})(\#K\in L).$$

This is equal to

$$f(\#L,\{e_1,\ldots,e_{\#L}\})(\#K\in L)\sum_{l=1}^{k}(-1)^{l-1}(l-1)!\frac{1}{l!}\sum_{j=0}^{l}(-1)^j\binom{l}{j}(l-j)^k.$$

Here, the part $\frac{1}{l!}\sum_{j=0}^{l}(-1)^j\binom{l}{j}(l-j)^k$ is the Stirling number of the second kind (see Definition 1.10). Using the relation $\left\{{k+1 \atop l}\right\}=l\left\{{k \atop l}\right\}+\left\{{k \atop l-1}\right\}$, we can see that $f(\#L,\{e_1,\ldots,e_{\#L}\})(\#K\in L)$ cancels out, because

$$\begin{aligned}
&\sum_{l=1}^{k+1}(-1)^{l-1}(l-1)!\left\{{k+1 \atop l}\right\}=\sum_{l=1}^{k+1}(-1)^{l-1}(l-1)!l\left\{{k \atop l}\right\}+\sum_{l=1}^{k+1}(-1)^{l-1}(l-1)!\left\{{k \atop l-1}\right\}\\
&\quad=\sum_{l=1}^{k+1}(-1)^{l-1}l!\left\{{k \atop l}\right\}+\sum_{l=0}^{k}(-1)^{l}l!\left\{{k \atop l}\right\}\\
&\quad=\sum_{l=1}^{k}(-1)^{l-1}l!\left\{{k \atop l}\right\}+(-1)^k(k+1)!\left\{{k \atop k+1}\right\}+\sum_{l=1}^{k}(-1)^{l}l!\left\{{k \atop l}\right\}+\left\{{k \atop 0}\right\}\\
&\quad=0.
\end{aligned}$$

Note that this phenomenon corresponds to what we have seen while studying the third cumulant.

Going back to the general choice of $e_1,\ldots,e_{\#L}$, it is clear that we need to first, find all partitions $Q_1\cup\cdots\cup Q_l$ that generate $f^{(\#L)}(\#L,\{e_1,\ldots,e_{\#L}\})(\#K,K\in L)$ and second, see under which conditions on $e_1,\ldots,e_{\#L}$ it eventually cancels out. For each $i=1,\ldots,\#L$ there exist $n_i\in\mathbb{N}$ and $q_1^i,\ldots,q_{\#L}^i\subset\{\rho x_1,\ldots,\rho x_k\}$ such that

$$e_i=g^{(n_i)}(\cdot,\{(-1)^{\#q_1^i+1}w^{\cap}_{q_1^i},\ldots,(-1)^{\#q_{n_i}^i+1}w^{\cap}_{q_{n_i}^i}\}).$$

For e_i we denote $R_i:=q_1^i\cup\cdots\cup q_{n_i}^i$. Moreover, denote $X^1:=R_1\cup\cdots\cup R_{\#L}$ and $X^2:=\{\rho x_1,\ldots,\rho x_k\}\backslash X^1$.

Step 1: Constructing the maximal partition of X^1. Next, we present an algorithm to find the largest $m\in\mathbb{N}$ such that there is a partition $Q_1\cup\cdots\cup Q_m=X^1$ allowing to generate $f^{(\#L)}(\#L,\{e_1,\ldots,e_{\#L}\})(\#K,K\in L)$. Then all its *subpartitions* (i.e. partitions created by making unions among $Q_1,\ldots,Q_m$) allow to generate $f^{(\#L)}(\#L,\{e_1,\ldots,e_{\#L}\})(\#K,K\in L)$.

1. Set $m=1, Q_1=R_1$ and $\mathbf{R}=\{R_2,\ldots,R_{\#L}\}$.

2. For each $x\in Q_1$ find all $R\in\mathbf{R}$ such that $x\in R$ and make the following updates: $Q_1=Q_1\cup R$ and $\mathbf{R}=\mathbf{R}\backslash\{R\}$.

3. Repeat step (2) until either (i) $Q_1\cap\cup_{R\in\mathbf{R}}R=\varnothing$ or (ii) $\mathbf{R}=\varnothing$. If (i) but not (ii), then set $m=2$, if (ii), then end the algorithm here.

4. Take any $R\in\mathbf{R}$ and put $Q_m=R$. For each $x\in Q_m$ find all $R'\in\mathbf{R}$ such that $x\in R'$ and make the following updates: $Q_m=Q_m\cup R'$ and $\mathbf{R}=\mathbf{R}\backslash\{R'\}$.

5. Repeat step (4) until either (i) $Q_m\cap\cup_{R\in\mathbf{R}}R=\varnothing$ or (ii) $\mathbf{R}=\varnothing$. If (i) but not (ii), go to step (4) with $m=m+1$, if (ii), then end the algorithm here.

Since $\#L$ is finite, the algorithm always stops resulting in a disjoint partition $Q_1 \cup \cdots \cup Q_m = X^1$ and m being the maximal integer such that there is a partition generating $f^{(\#L)}(\#L, \{e_1, \ldots, e_{\#L}\})(\#K, K \in L)$. Moreover, this partition is unique up to permutation and for $n = 1, \ldots, m$ there exist $\left\{{m \atop n}\right\}$ subpartitions $Q_1^{(n)} \cup \cdots \cup Q_n^{(n)} = X^1$.

Step 2: Adding the elements of X^2 into partition of X^1. To track all the partitions of $\{\rho x_1, \ldots, \rho x_k\}$ that generate $f^{(\#L)}(\#L, \{e_1, \ldots, e_{\#L}\})(\#K, K \in L)$, we need to add elements of X^2 to each partition of X^1 from Step 1. To do so, denote $p = \#X^2$. The maximal partition of $\{\rho x_1, \ldots, \rho x_k\}$ is of the size $m + p$ where m is the maximum size of partition of X^1 and p corresponds to a situation where each $x \in X^2$ belongs to an individual set. The minimal size of the partition is obviously equal to 1.

Fix $l \in \{1, \ldots, m + p\}$ and partition of X^1 from Step 1 of the size n such that $n \leqslant \min(m, l)$. The number of ways how to add elements of X^2 to create a partition that generates $f^{(\#L)}(\#L, \{e_1, \ldots, e_{\#L}\})(\#K, K \in L)$ is

$$\begin{bmatrix} p \\ l-n \end{bmatrix} := \frac{1}{(l-n)!} \sum_{i=0}^{l-n} (-1)^i \binom{l-n}{i} (l-i)^p.$$

This can be shown by the inclusion-exclusion principle.

Step 3: The role of $f^{(\#L)}(\#L, \{e_1, \ldots, e_{\#L}\})(\#K, K \in L)$ in $T^{(\rho)}_{m,k}$. If we put together Step 1, Step 2 and the expression (4.61), then for fixed L the term $f^{(\#L)}$ $(\#L, \{e_1, \ldots, e_{\#L}\})(\#K, K \in L)$ appears in $T^{(\rho)}_{m,k}$ exactly $\nu(m,p)$ times, where

$$\nu(m,p) := \sum_{l=1}^{m+p} (-1)^{l-1}(l-1)! \sum_{n=\max(1,l-p)}^{\min(m,l)} \left\{ {m \atop n} \right\} \begin{bmatrix} p \\ l-n \end{bmatrix}.$$

Our goal now is to study, for which parameters m, p is $\nu(m,p)$ equal to 0 and how the parameters m, p correspond to the choice of $e_1, \ldots, e_{\#L}$. It is easy to see that

$$\nu(1,0) = 1.$$

Take $m > 1$ and $p = 0$, then

$$\nu(m,0) = \sum_{l=1}^{m} (-1)^{l-1}(l-1)! \left\{ {m \atop l} \right\} = 0$$

as was shown in the example within this proof. To see similar result also for $m > 1, p > 0$ we first show a recurrence relation for $\left[{p \atop l-n}\right]$:

If $p > 0$ and $n \in \{\max(1, l-p-1), \ldots, \min(m,l)\}$, then

$$\begin{bmatrix} p+1 \\ l-n \end{bmatrix} = \begin{bmatrix} p \\ l-n-1 \end{bmatrix} + l \begin{bmatrix} p \\ l-n \end{bmatrix}. \tag{4.62}$$

This relation can be seen similarly as the recurrence relation for the Stirling number of the second kind. We extract a $(p+1)$-th element and divide it into two situations.

(a) $(p+1)$-th element from X^2 creates a singleton in the partition of the set $\{\rho x_1, \dots, \rho x_k\}$, i.e. a set in the partition where only this element belongs. Hence, the set containing this element does not contain any element from X^2, but neither from X^1. There are $\left[{p \atop l-n-1}\right]$ of such partitions, because we need to place the remaining p elements into $l-1$ sets where $l-n-1$ are non-empty.

(b) $(p+1)$-th element from X^2 does not create a singleton in the partition of $\{\rho x_1, \dots, \rho x_k\}$. Then we need to place the remaining p elements into l sets such that $l-n$ are non-empty and after that add the $(p+1)$-th element in any of the l sets. There are $l\left[{p \atop l-n}\right]$ of such partitions.

Eventually, for $m > 1, p > 0$ using (4.62), we have

$$\nu(m,p+1) = \sum_{l=1}^{m+p+1} (-1)^{l-1}(l-1)! \sum_{n=\max(1,l-p-1)}^{\min(m,l)} \left\{{m \atop n}\right\} \left(\left[{p \atop l-n-1}\right] + l\left[{p \atop l-n}\right] \right)$$
$$= \sum_{l=0}^{m+p} (-1)^l l! \sum_{n=\max(1,l-p)}^{\min(m,l+1)} \left\{{m \atop n}\right\} \left[{p \atop l-n}\right] - \sum_{l=1}^{m+p+1} (-1)^l l! \sum_{n=\max(1,l-p-1)}^{\min(m,l)} \left\{{m \atop n}\right\} \left[{p \atop l-n}\right]$$
$$= \sum_{l=1}^{m-1} (-1)^l l! \left\{{m \atop l+1}\right\} \left[{p \atop -1}\right] - \sum_{l=p+2}^{m+p} (-1)^l l! \left\{{m \atop l-p-1}\right\} \left[{p \atop p+1}\right]$$
$$- (-1)^{m+p+1}(m+p+1)! \left[{p \atop p+1}\right] = 0,$$

because $\left[{p \atop p+1}\right] = 0$ and $\left[{p \atop -1}\right] = 0$.

We conclude that $\nu(m,p) = 1$ if and only if $m = 1$ and $p = 0$. In other cases $\nu(m,p) = 0$. In the language of our algorithm above, the parameters $m = 1, p = 0$ correspond to the situation when $Q_1 = \{\rho x_1, \dots, \rho x_k\}$. This happens when $R_1 \cup \dots \cup R_{\#L} = \{\rho x_1, \dots, \rho x_k\}$ and $R_1 \cap \dots \cap R_{\#L} \neq \varnothing$. In fact, here the corresponding term $f^{(\#L)}(\#L, \{e_1, \dots, e_{\#L}\})(\#K, K \in L)$ can be generated only by the first term in $T^{(\rho)}_{m,k}$ (i.e. when $l = 1$). This class of terms includes e.g.

- $f_{g^{(n)}(\cdot,\{(-1)^{\#q_1+1}w^{\cap}_{q_1},\dots,(-1)^{\#q_n+1}w^{\cap}_{q_n}\})}(\#L, S)(\#K, K \in L), n \in \mathbb{N}, q_1, \dots, q_n \subset \{\rho x_1, \dots, \rho x_k\}$ such that $\cup_{i=1}^n q_i = \{\rho x_1, \dots, \rho x_k\}$,
- $f_{\{g(\cdot,w^{\cap}_{q_1}),\dots,g(\cdot,w^{\cap}_{q_n})\}}(\#L, S)(\#K, K \in L), n \in \mathbb{N}, q_1, \dots, q_n \subset \{\rho x_1, \dots, \rho x_k\}$ such that $\cup_{i=1}^n q_i = \{\rho x_1, \dots, \rho x_k\}$ and $\cap_{i=1}^n q_i \neq \varnothing$.

Lemma 4.8 can be extended to see that for $e_1, \dots, e_{\#L}$ for which $m = 1$ and $p = 0$

$$\rho^{k/2} \int_{K^k} f^{(\#L)}(\#L, \{e_1, \dots, e_{\#L}\})(\#K, K \in L)\mathrm{d}x_1 \dots \mathrm{d}x_k \xrightarrow[\rho\to\infty]{} 0$$

if $\mathbb{E}\,|\Xi_0|_1^k < \infty$. The convergence does not depend on the choice of L. Hence,

$$\int_{K^k} T^{(\rho)}_{m,k}(x_1, \dots, x_k)\mathrm{d}x_1 \dots \mathrm{d}x_k \xrightarrow[\rho\to\infty]{} 0.$$

□

Denote by $\mathbb{F}_k$ the set of all functions $e_1 \cdot \ldots \cdot e_l, l \in \mathbb{N}$, such that e_i is of the form

$$e_i = g^{(n_i)}(n_i, \{(-1)^{\#q_1^i+1} w^{\cap}_{q_1^i}, \ldots, (-1)^{\#q_{n_i}^i+1} w^{\cap}_{q_{n_i}^i}\}), \tag{4.63}$$

where $n_i \in \mathbb{N}$, $q_1^i, \ldots, q_{n_i}^i \subset \{\rho x_1, \ldots, \rho x_k\}$ and $\sum_i n_i = l$ such that the algorithm in the proof of Lemma 4.10 for $e_1, \ldots, e_l$ returns $m = 1$ and $p = 0$. Then

$$\rho^{k/2} \int\limits_{K^k} h(x_1, \ldots, x_k) \mathrm{d}x_1 \ldots \mathrm{d}x_k \xrightarrow{\rho \to \infty} 0, \quad \text{for all } h \in \mathbb{F}_k,$$

assuming $\mathbb{E}\,|\Xi|_1^k < \infty$.

We say that the function $h = e_1 \cdot \ldots \cdot e_l \in \mathbb{F}_k$ is of the *basic form*, if $e_1, \ldots, e_l$ are pairwise different functions of the form (4.63) and for each $i = 1, \ldots, l$, $q_1^i, \ldots, q_{n_i}^i$ are pairwise different subsets of $\{\rho x_1, \ldots, \rho x_k\}$. Let the set of all functions of the basic form be denoted by $\mathbb{F}_{k,basic}$. Note that $\#(\mathbb{F}_{k,basic}) < \infty$ and the number of elements depends only on k.

Lemma 4.11.
Let $h = e_1 \cdot \ldots \cdot e_l \in \mathbb{F}_k$ be such that each e_i is of the form of (4.63). Then

$$\left| \rho^{k/2} \int\limits_{K^k} h(x_1, \ldots, x_k) \mathrm{d}x_1 \ldots \mathrm{d}x_k \right| \leqslant C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \ldots, \mathbb{E}\,|\Xi_0|_1^k)(\max\{1, \lambda \mathbb{E}\,|\Xi_0|_1\})^l \prod_{i=1}^{l} \|\gamma_{red}^{[n_i]}\|_{TV},$$

where $C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \ldots, \mathbb{E}\,|\Xi_0|_1^k) < \infty$ is a constant that does not depend on h, hence l, and $C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \ldots, \mathbb{E}\,|\Xi_0|_1^k) \to 0$ as $\rho \to \infty$ if $\mathbb{E}\,|\Xi|_1^k < \infty$.

Proof. The proof consists of a generalization of a step we made in the proof of Lemma 4.8, point 5, when transitioning from $n = 3$ to general $n \in \mathbb{N}$.

For each $i = 1, \ldots, l$, we find the greatest $m_i \in \mathbb{N}$ such that $\{q'^i_1, \ldots, q'^i_{m_i}\} \subset \{q_1^i, \ldots, q_{n_i}^i\}$ is a set of pairwise different subsets of $\{\rho x_1, \ldots, \rho x_k\}$. Denote by e_i' the function

$$e_i' = g^{(m_i)}(n_i, \{w^{\cap}_{q'^i_1}, \ldots, w^{\cap}_{q'^i_{m_i}}\}).$$

Then

$$\rho^{k/2} \left| \int\limits_{K^k} e_1 \cdot \ldots \cdot e_l \mathrm{d}x_1 \ldots \mathrm{d}x_k \right| \leqslant \rho^{k/2} \left| \int\limits_{K^k} e_1' \cdot \ldots \cdot e_l' \mathrm{d}x_1 \ldots \mathrm{d}x_k \right|. \tag{4.64}$$

Now, we find l' the greatest integer and ordering $(e_1', \ldots, e_{l'}', e_{l'+1}', \ldots, e_l')$ such that $e_1', \ldots, e_{l'}'$ are pairwise different functions. Then

$$\rho^{k/2} \left| \int\limits_{K^k} e_1' \cdot \ldots \cdot e_l' \mathrm{d}x_1 \ldots \mathrm{d}x_k \right| \leqslant (\lambda \mathbb{E}\,|\Xi_0|_1)^{l-l'} \prod_{j \geqslant l'+1, n_j \geqslant 2} \|\gamma_{red}^{[n_j]}\|_{TV} \rho^{k/2} \left| \int\limits_{K^k} e_1' \cdot \ldots \cdot e_{l'}' \mathrm{d}x_1 \ldots \mathrm{d}x_k \right|. \tag{4.65}$$

Here, we used that

$$\int_{\mathbb{R}} w_q^{\cap}(p)\gamma^{[1]}(\mathrm{d}p) = \lambda \int_{\mathbb{R}} w_q^{\cap}(p)\mathrm{d}p \leqslant \lambda \mathbb{E}\,|\Xi_0|_1$$

and

$$\int_{\mathbb{R}^n} w_{q_1}^{\cap}(p_1)\cdot\dots\cdot w_{q_n}^{\cap}(p_n)\gamma^{[n]}(\mathrm{d}(p_1,\dots,p_n)) \leqslant \lambda \mathbb{E}\,|\Xi_0|_1 \|\gamma_{red}^{[n]}\|_{TV}$$

for all $n \in \mathbb{N}$, $q, q_1, \dots, q_n \subset \{\rho x_1, \dots, \rho x_k\}$. Define for $i = 1, \dots, l'$ functions e_i'' by

$$e_i'' = g^{(m_i)}(m_i, \{w_{q'^i_1}^{\cap}, \dots, w_{q'^i_{m_i}}^{\cap}\}).$$

Then $e_1'' \cdot \dots \cdot e_{l'}'' \in \mathbb{F}_{k,basic}$. By following the procedure in the proof of Lemma 4.8, we arrive at

$$\rho^{k/2} \int_{K^k} e_1'' \cdot \dots \cdot e_{l'}'' \mathrm{d}x_1 \dots \mathrm{d}x_k$$

$$\leqslant C_k^{(\rho)}(e_1'' \cdot \dots \cdot e_{l'}'', \lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) \prod_{i=1}^{l'} \|\gamma_{red}^{[m_i]}\|_{TV},$$

where $C_k^{(\rho)}(e_1'' \cdot \dots \cdot e_{l'}'', \lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) \to 0$ as $\rho \to \infty$. Since, $\mathbb{F}_{k,basic}$ is finite, we define

$$C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) := \max_{h \in \mathbb{F}_{k,basic}} C_k^{(\rho)}(h, \lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) < \infty.$$

Then,

$$\int_{K^k} e_1' \cdot \dots \cdot e_{l'}' \mathrm{d}x_1 \dots \mathrm{d}x_k \leqslant C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) \prod_{i=1,\dots,l':n_i \geqslant 2} \|\gamma_{red}^{[n_i]}\|_{TV}. \tag{4.66}$$

Finally, combining (4.64), (4.65) and (4.66), we arrive at

$$\left| \rho^{k/2} \int_{K^k} h(x_1, \dots, x_k)\mathrm{d}x_1 \dots \mathrm{d}x_k \right|$$

$$\leqslant (\lambda \mathbb{E}\,|\Xi_0|_1)^{l-l'} C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) \prod_{j=1}^{l} \|\gamma_{red}^{[n_j]}\|_{TV}$$

$$\leqslant (\max\{1, \lambda \mathbb{E}\,|\Xi_0|_1\})^l C_k^{(\rho)}(\lambda, K, \mathbb{E}\,|\Xi_0|_1, \dots, \mathbb{E}\,|\Xi_0|_1^k) \prod_{j=1}^{l} \|\gamma_{red}^{[n_j]}\|_{TV}.$$

□

Finally, we are ready to prove the main theorem.

Proof of Theorem 4.5. Take $k \geqslant 3$. Recall that

$$\rho^{-\frac{3k}{2}} \mathbf{Cum}_k(|\Xi \cap \rho K|_2) = \sum_{m=1}^{\infty} \frac{(-1)^m}{m!} \int_{K^k} T_{m,k}^{(\rho)}(x_1, \dots, x_k)\mathrm{d}x_1 \dots \mathrm{d}x_k,$$

where $\int_{K^k} T^{(\rho)}_{m,k}(x_1,\dots,x_k)\mathrm{d}x_1\dots\mathrm{d}x_k \to 0$ as $\rho\to\infty$ according to Lemma 4.10. We want to change the order of the summation and the limit.

Denote by $s_n(\rho) := \sum_{m=1}^{n} \frac{(-1)^m}{m!}\int_{K^k} T^{(\rho)}_{m,k}(x_1,\dots,x_k)\mathrm{d}x_1\dots\mathrm{d}x_k$ the partial sum and by $s(\rho) := \sum_{m=1}^{\infty} \frac{(-1)^m}{m!}\int_{K^k} T^{(\rho)}_{m,k}(x_1,\dots,x_k)\mathrm{d}x_1\dots\mathrm{d}x_k$ the infinite series. The goal is to show that $s_n(\rho)$ converges uniformly to $s(\rho)$ on some interval $[A,\infty)$. Then from the Moore–Osgood theorem, the assertion holds.

The uniform convergence shall be proved using the Weierstrass criterion for the absolute uniform convergence: Take $m\in\mathbb{N}$ and denote by

$$C_k := \sup_{\rho>0} C^{(\rho)}_k(\lambda, K, \mathbb{E}\,|\Xi_0|_1,\dots,\mathbb{E}\,|\Xi_0|_1^k),$$

where $C^{(\rho)}_k(\lambda, K, \mathbb{E}\,|\Xi_0|_1,\dots,\mathbb{E}\,|\Xi_0|_1^k)$ is the constant from Lemma 4.11. Note that $C_k<\infty$, since $C^{(\rho)}_k(\lambda, K, \mathbb{E}\,|\Xi_0|_1,\dots,\mathbb{E}\,|\Xi_0|_1^k)\to 0$ as $\rho\to\infty$, provided that $\mathbb{E}\,|\Xi_0|_1,\dots,\mathbb{E}\,|\Xi_0|_1^k<\infty$. Moreover, we will use the fact that there exist constants $a,b>0$ such that $\|\gamma^{[k]}_{red}\|_{TV}\leqslant ab^k$ for all $k\in\mathbb{N}$. Thus,

$$\begin{aligned}
&\left|\frac{(-1)^m}{m!}\int\limits_{K^k} T^{(\rho)}_{m,k}(x_1,\dots,x_k)\mathrm{d}x_1\dots\mathrm{d}x_k\right| \\
&= \left|\frac{(-1)^m}{m!}\rho^{\frac{k}{2}} \sum_{\substack{\cup K_i=\{1,\dots,m\}\\ L=\{K_1,\dots\}}} \sum_{e_1\cdot\ldots\cdot e_{\#L}\in\mathbb{F}_k} \int\limits_{K^k} e_1\cdot\ldots\cdot e_{\#L}\mathrm{d}x_1\dots\mathrm{d}x_k\right| \\
&\leqslant \rho^{\frac{k}{2}}\frac{1}{m!}\sum_{l=1}^{m}\sum_{\cup K_i=\{1,\dots,m\}} \sum_{\substack{e_1\cdot\ldots\cdot e_l\in\mathbb{F}_k\\ n_i=\#K_i, i=1,\dots,l}} \left|\int\limits_{K^k} e_1\cdot\ldots\cdot e_l\mathrm{d}x_1\dots\mathrm{d}x_k\right| \\
&\leqslant \frac{1}{m!}\sum_{l=1}^{m}\sum_{\cup K_i=\{1,\dots,m\}} \sum_{\substack{e_1\cdot\ldots\cdot e_l\in\mathbb{F}_k\\ n_i=\#K_i, i=1,\dots,l}} C_k(\max\{1,\lambda\mathbb{E}\,|\Xi_0|_1\})^l\prod_{i=1}^{l}\|\gamma^{[n_i]}_{red}\|_{TV} \\
&\leqslant C_k\frac{1}{m!}b^m\sum_{l=1}^{m}(a\max\{1,\lambda\mathbb{E}\,|\Xi_0|_1\})^l \sum_{\cup K_i=\{1,\dots,m\}} \sum_{\substack{e_1\cdot\ldots\cdot e_l\in\mathbb{F}_k\\ n_i=\#K_i, i=1,\dots,l}} 1 \\
&\leqslant C_k\frac{1}{m!}b^m\sum_{l=1}^{m}(a\max\{1,\lambda\mathbb{E}\,|\Xi_0|_1\})^l \sum_{\cup K_i=\{1,\dots,m\}} (2^k)^m \\
&\leqslant C_k\frac{1}{m!}b^m(2^k)^m(\max\{1,a\lambda\mathbb{E}\,|\Xi_0|_1\})^m\sum_{l=1}^{m}\left\{ m \atop l \right\} := Q_m.
\end{aligned}$$

We refer to $B_m := \sum_{l=1}^{m}\left\{ m \atop l \right\}$ as the m-th Bell number. Finally

$$\sum_{m=1}^{\infty} Q_m = C_k \exp\{e^{2^k b\max\{1,a\lambda\mathbb{E}\,|\Xi_0|_1\}}-1\}<\infty.$$

Hence, s_n converges uniformly absolutely to s. As a consequence, we have that $\rho^{-3k/2}\mathbf{Cum}_k(|\Xi\cap\rho K|_2)\to 0$. That, together with Theorem 4.4, concludes the proof.

□

Conclusion

We have investigated three approaches to study the asymptotic behaviour of geometrical structures frequently used in stochastic geometry. Numerous examples were presented to give the reader a rough idea on which situations are suitable for each individual method. We have also seen that it is not unusual that the methods are combined in order to achieve some asymptotic results.

In conclusion, the Malliavin-Stein approach is a very robust method that can be applied to a large scale of examples allowing the researcher to work on a general Polish space. It can be used for Poisson functionals as well as for functionals of Gibbs processes, DPP's or even for processes of particles. Especially, the U-statistics form an important class of functionals suitable for this analysis. The reason is that many interesting functionals can be expressed as U-statistics (e.g. the intrinsic volumes) while the U-statistics can be viewed as finite sums of multiple Wiener-Itô integrals. Hence, to study the asymptotic properties of a U-statistic is equivalent to study the individual Wiener-Itô integrals. The Malliavin operators in this case are usually easy to handle. However, an application of the Malliavin-Stein method in general usually results in abstract bounds for the normal approximation involving difference operators. To derive central limit theorems, one is sometimes obliged to use another technique such as the stabilization method. Either way, these bounds give an opportunity to compute explicit rates of convergence.

On the other hand, the stabilization method is a useful method for studying the limit behaviour of geometric structures mainly in $\mathbb{R}^d$ evincing local form of dependency. Those include random graphs, germ-grain models, weighted Voronoi tessellation etc. In order to obtain some information about the limit behaviour of these structures, we investigate a sum of spatially dependent terms called scores. We make use of the property that we can control the range of interactions. In other words, our score stabilizes if its behaviour at a given point is locally determined by a certain finite, possibly random, neighbourhood of this point. Those local effects appear mostly when the geometric structure is determined by the Poisson or binomial point process.

At last, we have acquainted ourselves with the method of cumulants based on a classical result from probability theory that normal distribution is the only one having only a finite number of non-zero cumulants. By the moment convergence theorem, the convergence of the higher-order cumulants to zero is equivalent to the convergence in distribution to a Gaussian random variable. One can apply the idea directly on a functional of a random structure in increasing observation window (e.g. the coverage volume). The advantage is that the geometrical structure can evince long-range dependencies. On the other hand, a very limited spectrum of geometrical structures is suitable for this analysis. Typically, those are the structures based on Poisson point process in $\mathbb{R}^d$ or very simple structures defined by a type of Brillinger-mixing point processes. Application of the cumulant method just by itself is accessible mainly if there exists a link between the cumulants of the random variable defined by the random structure and cumulant measures of the defining point process. In our experience, however, estimation of the cumulants could lead to a calculus with extremely unpleasant terms.

Bibliography

F. Avram and D. Bertsimas. On central limit theorems in geometrical probability. *Annals of Applied Probability*, 3:1033–1046, 1993.

E. Azmoodeh, G. Peccati, and X. Yang. Malliavin—Stein method: a survey of some recent developments. *Modern Stochastics: Theory and Applications*, 8: 141–177, 2021.

F. Baccelli, B. Błaszczyszyn, and M. Karray. Random Measures, Point Processes and Stochastic Geometry. 2020. hal-02460214.

A. Baddeley. *Spatial Point Processes and Their Applications. In: A. Baddeley, I. Bárány, R. Schneider, W. Weil (eds.). Stochastic Geometry. Lecture Notes in Mathematics*, volume 1892. Springer, Berlin, 2007. ISBN 9783540381747.

A. J. Baddeley. *Spatial Sampling and Censoring. In: O. E. Barndorff-Nielsen, W. S. Kendall and M. N. M. van Lieshout (eds.). Stochastic Geometry: Likelihood and Computation.* Chapman and Hall, London, 1999. ISBN 9780203738276.

A. Barbour and L. H.-Y. Chen. *An Introduction to Stein's Method.* Lecture Notes Series. Singapore University Press, Singapore, and World Scientific, London, 2005a. ISBN 981-256-280-X.

A. Barbour and L. H.-Y. Chen. *Stein's Method and Applications.* Lecture Notes Series. Singapore University Press, Singapore, and World Scientific, London, 2005b. ISBN 981-256-281-8.

A. D. Barbour and A. Xia. Normal approximation for random sums. *Advances in Applied Probability*, 38:693–728, 2006.

Y. Baryshnikov and J. E. Yukich. Gaussian limits for random measures in geometric probability. *Annals of Applied Probability*, 15:213–253, 2005.

J. Beardwood, J. H. Halton, and J. M. Hammersley. The shortest path through many points. *Mathematical Proceedings of the Cambridge Philosophical Society*, 55:299–327, 1959.

V. Beneš, J. Večeřa, and M. Pultar. Planar segment processes with reference mark distributions, modeling and estimation. *Methodology and Computing in Applied Probability*, 21:683–698, 2019.

C. A. N. Biscio and F. Lavancier. Brillinger mixing of determinantal point processes and statistical applications. *Electronic Journal of Statistics*, 10:582–607, 2016.

B. Błaszczyszyn, D. Yogeshwaran, and J. E. Yukich. Limit theory for geometric statistics of point processes having fast decay of correlations. *Annals of Probability*, 47:835–895, 2019.

V. I. Bogachev. *Measure Theory. Volume I.* Springer-Verlag, Berlin, 2007. ISBN 9783540345138.

S. Bourguin and G. Peccati. *The Malliavin—Stein Method on the Poisson Space. In: Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry.* Bocconi & Springer Series. Springer, Cham, 2016. ISBN 978-3-319-05232-8.

D. R. Brillinger. *Time Series: Data Analysis and Theory.* Holt, Rinehart and Winston, 1975. ISBN 9780030769757.

A. Bulinski and A. Shashkin. *Limit Theorems for Associated Random Fields and Related Systems.* Advanced series on statistical science and applied probability. World Scientific, Singapore, 2007. ISBN 978-981-270-940-0.

A. Bulinski, E. Spodarev, and F. Timmetmann. Central limit theorems for the excursion set volumes of weakly dependent random fields. *Bernoulli*, 18:100–118, 2021.

S. Chatterjee. Fluctuations of eigenvalues and second order Poincaré inequalities. *Probability Theory and Related Fields*, 143:1–40, 2009.

L. H.-Y. Chen. Poisson approximation for dependent trials. *The Annals of Probability*, 3(3):534–545, 1975.

L. H.-Y. Chen and Q.-M. Shao. *Stein's Method for Normal Approximation. In: A. D. Barbour and L. H.-Y. Chen. An Introduction to Stein's Method.* Singapore University Press, Singapore, and World Scientific, London, 2005. ISBN 981-256-280-X.

L. H.-Y. Chen, L. Goldstein, and Q.-M. Shao. *Normal Approximation by Stein's Method.* Springer-Verlag, Berlin, 2011. ISBN 978-3-642-15006-7.

S. N Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications.* Wiley, Hoboken, New Jersey, 2013. ISBN 978-0-470-66481-0.

F. Cuevas-Pacheco and J. Møller. Log Gaussian Cox processes on the sphere. *Spatial Statistics*, 26:69–82, 2018.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods. 2nd Edition.* Probability and Its Applications. Springer, New York, 2003. ISBN 0-387-95541-0.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure. 2nd Edition.* Probability and Its Applications. Springer, New York, 2008. ISBN 978-0-387-21337-8.

D. Dereudre. Introduction to the theory of Gibbs point processes. *arXiv preprint arXiv:1701.08105*, 2017.

D. Dereudre, R. Drouilhet, and H.-O. Georgii. Existence of Gibbsian point processes with geometry-dependent interactions. *Probability Theory and Related Fields*, 153:643–670, 2010.

P. L. Dobruschin. The description of a random field by means of conditional probabilities and conditions of its regularity. *Theory of Probability and Its Applications*, 13(2):197–224, 1968.

R. M. Dudley. *Real Analysis and Probability*, volume 74. Cambridge University Press, Cambridge, 2002. ISBN 978-0521007542.

P. Erdös and A. Rényi. On random graphs i. *Publicationes Mathematicae Debrecen*, 6:290–297, 1959.

R. Fernández, P. Ferrari, and N. Garcia. Measures on contour, polymer or animal models, a probabilistic approach. *Markov Processes and Related Fields*, 4:479–497, 1998.

D. Flimmel. *Consequences and applications of the Fock space representation theorem*. Author's master thesis. Charles University, 2017.

D. Flimmel and V. Beneš. Gaussian approximation for functionals of Gibbs particle processes. *Kybernetika*, 54:765–777, 2018.

D. Flimmel and L. Heinrich. On the variance of the area of planar cylinder processes driven by Brillinger-mixing point processes. *submitted*, 2021+. URL https://arxiv.org/abs/2104.10224.

D. Flimmel, Z. Pawlas, and J. E. Yukich. Limit theory for unbiased and consistent estimators of statistics of random tessellations. *Journal of Applied Probability*, 57:679–702, 2020.

S. Friedli and Y. Velenik. *Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction*. Cambridge University Press, Cambridge, 2017. ISBN 978-1-107-18482-4.

J. Galambos and I. Simonelli. *Bonferroni-type Inequalities with Applications*. Probability and Its Applications. Springer-Verlag, New York, 1996. ISBN 978-0-387-94776-1.

H.-O. Georgii. Canonical and grand canonical Gibbs states for continuum systems. *Communications of Mathematical Physics*, 48:31–51, 1976.

H.-O. Georgii. *Gibbs Measures and Phase Transitions. 2nd edition.*, volume 9 of *Studies in Mathematics*. De Gruyter, Berlin, 2011. ISBN 978-3-11-025029-9.

H.-O. Georgii and H. J. Yoo. Conditional intensity and Gibbsianness of determinantal point processes. *Journal of Statistical Physics*, 118:617–666, 2005.

B. V. Gnedenko and A. N. Kolmogorov. *Limit Distributions for Sums of Independent Random Variables*. Addison and Wesley, Reading, 1954. ISBN 9780030769757.

L. Heinrich. Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes. *Statistics: A Journal of Theoretical and Applied Statistics*, 19:87–106, 1988.

L. Heinrich. Large deviations of the empirical volume fraction for stationary Poisson grain models. *Annals of Applied Probability*, 15:392–420, 2005.

L. Heinrich. *Asymptotic Methods in Statistics of Random Point Processes. In: Stochastic Geometry, Spatial Statistics and Random Fields. Asymptotic Methods. E. Spodarev (eds.). Lecture Notes in Mathematics*, volume 2068. Springer, Berlin, 2013. ISBN 978-3-642-33304-0.

L. Heinrich. Brillinger-mixing point processes need not to be ergodic. *Statistics and Probability Letters*, 138:31–35, 2018.

L. Heinrich. On the Brillinger-mixing property of stationary point processes. *to be submitted for publication, 20 pages*, 2021+.

L. Heinrich and Z. Pawlas. Absolute regularity and Brillinger-mixing of stationary point processes. *Lithuanian Mathematical Journal*, 53:293–310, 2013.

L. Heinrich and V. Schmidt. Normal convergence of multidimensional shot noise and rates of this convergence. *Advances in Applied Probability*, 17:709–730, 1985.

L. Heinrich and M. Spiess. Berry–Esseen bounds and Cramér-type large deviations for the volume distribution of Poisson cylinder processes. *Lithuanian Mathematical Journal*, 49:381–398, 2009.

L. Heinrich and M. Spiess. Central limit theorems for volume and surface content of stationary Poisson cylinder processes in expanding domains. *Advances in Applied Probability*, 45:312–331, 2013.

C. Hofer-Temmel and P. Houdebert. Disagreement percolation for Gibbs ball models. *Stochastic Processes and their Applications*, 129(10):3922–3940, 2019.

J. B. Hough, M. Krishnapur, Y. Peres, and B. Virag. *Zeros of Gaussian Analytic Functions and Determinantal Point Processes.* University Lecture Series. American Mathematical Society, Providence, 2009. ISBN 978-0-8218-4373-4.

K. Itô. Spectral type of the shift transformation of differential processes with stationary increments. *Transactions of the American Mathematical Society*, 81:253–263, 1956.

S. Jansen. Cluster expansions for Gibbs point processes. *Advances in Applied Probability*, 51(4):1129–1178, 2019.

H. Kesten and S. Lee. The central limit theorem for weighted minimal spanning trees on random points. *Annals of Applied Probability*, 6:495–527, 1996.

R. Lachièze-Rey and G. Peccati. Fine Gaussian fluctuations on the Poisson space II: rescaled kernels, marked processes and geometric U-statistics. *Stochastic Processes and Their Applications*, 123:4186–4218, 2013.

R. Lachièze-Rey and G. Peccati. New Berry–Esseen bounds for functionals of binomial point processes. *Annals of Applied Probability*, 27:1992–2031, 2017.

R. Lachièze-Rey and M. Reitzner. *U-statistics in Stochastic Geometry. In: Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry.* Bocconi & Springer Series. Springer, Cham, 2016. ISBN 978-3-319-05232-8.

R. Lachièze-Rey, M. Schulte, and J. E. Yukich. Normal approximation for stabilizing functionals. *Annals of Applied Probability*, 29:931–993, 2019.

G. Last. *Stochastic Analysis for Poisson Processes. In: Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry.* Bocconi & Springer Series. Springer, Cham, 2016. ISBN 978-3-319-05232-8.

G. Last and M. Penrose. Fock space representation, chaos expansion and covariance inequalities for general Poisson processes. *Probability Theory and Related Fields*, 150:663–690, 2011.

G. Last and M. Penrose. *Lectures on the Poisson Process.* Cambridge University Press, Cambridge, 2017.

G. Last, G. Peccati, and M. Schulte. Normal approximation on Poisson spaces: Mehler's formula, second order Poincaré inequalities and stabilization. *Probability Theory and Related Fields*, 165:667–723, 2016.

C. Lautensack. *Random Laguerre Tessellations.* PhD thesis, Universität Karlsruhe, 2007.

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

F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference. *Journal of Royal Statistical Society: Series B (Statistical Methodology)*, 77:853–877, 2015.

S. Lee. The central limit theorem for euclidean minimal spanning trees I. *Annals of Applied Probability*, 7:996–1020, 1997.

S. Lee. The central limit theorem for euclidean minimal spanning trees II. *Advances in Applied Probability*, 31:969–984, 1999.

E. Lukacs. *Characteristic Functions.* Charles Griffin & Co., Ltd.; 2nd Edition, London, 1970. ISBN 978-0852641705.

O. Macchi. The coincidence approach to stochastic point processes. *Advances in Applied Probability*, 7:83–122, 1975.

P. Malliavin. *Stochastic Calculus of Variations and Hypoelliptic Operators. In: Proceedings of the International Symposium on Stochastic Differential Equations, Kyoto.* Wiley, New York, 1978. ISBN 978-0471053750.

P. Malliavin. *Stochastic Analysis.* Springer-Verlag, Berlin, 1997. ISBN 978-3-540-57024-0.

J. Marcinkiewicz. Sur une propriété de la loi de Gauss. *Mathematische Zeitschrift*, 44:612–618, 1939.

B. Matérn. *Spatial Variation. In: Lecture Notes in Statistics*, volume 36. Springer, Springer, Berlin, 1986. ISBN 9783540963653.

K. McGivney and J. E. Yukich. Asymptotics for Voronoi tessellations on random samples. *Stochastic Processes and Their Applications*, 83:273–288, 1999.

D. L. McLeish. Dependent central limit theorems and invariance principles. *Annals of Probability*, 2:620–628, 1974.

P.-A. Meyer. *Quantum Probability for Probabilists. 2nd Edition.* Lecture Notes in Mathematics. Springer-Verlag, Berlin, 1995. ISBN 3-540-60270-4.

R. E. Miles. On the homogeneous planar Poisson point process. *Mathematical Biosciences*, 6:85–127, 1970.

R. E. Miles. *On the Elimination of Edge Effects in Planar Sampling. In: E. F. Harding and D. G. Kendall (eds.). Stochastic Geometry: A Tribute to the Memory of Rollo Davidson.* John Wiley and Sons, London, 1974. ISBN 9780471351412.

I. Molchanov. *Theory of Random Sets.* Springer-Verlag, London, 2005. ISBN 978-1-84996-949-9.

J. Møller. Random Johnson–Mehl tessellations. *Advances in Applied Probability*, 24(4):814–844, 1992.

J. Møller. *Lectures on Random Voronoi Tessellations.* Lecture Notes in Statistics. Springer, New York, 1994. ISBN 978-0-387-94264-3.

J. Møller and E. Rubak. Functional summary statistics on the sphere with an application to determinantal point processes. *Spatial Statistics*, 18:4–23, 2016.

J. Møller, M. Nielsen, E. Porcu, and E. Rubak. Determinantal point process models on the sphere. *Bernoulli*, 24:1171–1201, 2018.

X. Nguyen and H. Zessin. Integral and differential characterizations of Gibbs processes. *Mathematische Nachrichten*, 88:105–115, 1979.

I. Nourdin and G. Peccati. Stein's method on Wiener chaos. *Probability Theory and Related Fields*, 145:75–118, 2009a.

I. Nourdin and G. Peccati. Stein's method and exact Berry–Esseen asymptotics for functionals of gaussian fields. *Annals of Probability*, 37:2231–2261, 2009b.

I. Nourdin and G. Peccati. *Normal Approximations with Malliavin Calculus. From Stein's Method to Universality.* Cambridge University Press, Cambridge, 2012. ISBN 978-1-107-01777-1.

D. Nualart. *The Malliavin Calculus and Related Topics. 2nd Edition.* Probability and Its Applications. Springer-Verlag, Berlin, 2006. ISBN 978-3-540-28328-7.

D. Nualart and G. Peccati. Central limit theorems for sequences of multiple stochastic integrals. *Annals of Probability*, 33:177–193, 2005.

D. Nualart and J. Vives. Anticipative calculus for the Poisson process based on the Fock space. *Séminaire de probabilités de Strasbourg*, 24:154–165, 1990.

A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams. 2nd Edition.* Wiley, New Chichester, 2000. ISBN 978-0-471-98635-5.

C. Palm. Intensitätsschwankungen im Fernsprechverkehr. *Ericssons Techniks*, 44:1–189, 1943.

F. Papangelou. The conditional intensity of general point processes and an application to line processes. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 28:207–226, 1974.

G. Peccati and M. Reitzner. *Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry.* Bocconi & Springer Series. Springer, Cham, 2016. ISBN 978-3-319-05232-8.

G. Peccati and M. Taqqu. *Moments, cumulants and diagram formulae for non-linear functionals of random measures.* hal-00338132f, 2008.

G. Peccati, J. L. Solé, M. S. Taqqu, and F. Utzet. Stein's method and normal approximation of Poisson functionals. *Annals of Probability*, 38:443–478, 2010.

M. Penrose. *Random Geometric Graphs.* Oxford University Press, New York, 2003. ISBN 978-0198506263.

M. Penrose. Laws of large numbers in stochastic geometry with statistical applications. *Bernoulli*, 13:1124–1150, 2007a.

M. Penrose and J. E. Yukich. *Normal Approximation in Geometric Probability. In: A. D. Barbour and L. H.-Y. Chen. Stein's Method and Applications.* Singapore University Press, Singapore, 2005. ISBN 978-981-256-281-4.

M. D. Penrose. Gaussian limits for random geometric measures. *Electronic Journal of Probability*, 12:989–1035, 2007b.

M. D. Penrose and J. E. Yukich. Central limit theorems for some graphs in computational geometry. *Annals of Applied Probability*, 11:1005–1041, 2001.

M. D. Penrose and J. E. Yukich. Limit theory for random sequential packing and deposition. *Annals of Applied Probability*, 12:272–301, 2002.

M. D. Penrose and J. E. Yukich. Weak laws of large numbers in geometric probability. *Annals of Applied Probability*, 13:277–303, 2003.

J. Rataj. *Bodové Procesy.* Karolinum, Praha, 2006. ISBN 8024611821.

D. Ruelle. *Statistical Mechanics: Rigorous Results.* W.A. Benjamin, Massachusetts, 1969.

L. Saulis and V. A. Statulevičius. *Limit Theorems for Large Deviations.* Mathematics and its Applications. Springer, Kluwer, Dordrecht, 1991. ISBN 9780792314752.

R. Schneider and W. Weil. *Stochastic and Integral Geometry.* Probability and Its Applications. Springer, Berlin, 2008. ISBN 9783540788584.

T. Schreiber. *Limit Theorems in Stochastic Geometry. In: New perspectives in Stochastic Geometry.* Oxford University Press, Oxford, 2010. ISBN 978-0199232574.

T. Schreiber and J. E. Yukich. Limit theorems for geometric functionals of Gibbs point processes. *Annales de l'Institut Henri Poincaré*, 49:1158–1182, 2013.

M. Schulte and C. Thäle. *Poisson Point Process Convergence and Extreme Values in Stochastic Geometry. In: Stochastic Analysis for Poisson Point Processes. Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry.* Bocconi & Springer Series. Springer, Cham, 2016. ISBN 978-3-319-05232-8.

A. N. Shiryayev. *Probability.* Springer-Verlag, Berlin Heidelberg New York, 1984. ISBN 978-1-4899-0018-0.

A. Soshnikov. Determinantal random point fields. *Russian Mathematical Surveys*, 55:923–975, 2000.

M. Spiess and E. Spodarev. Anisotropic Poisson processes of cylinders. *Methodology and Computing in Applied Probability*, 13:801–819, 2011.

Ch. Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 2:583–602, 1972.

J. Štěpán. *Teorie Pravděpodobnosti.* Academia, Praha, 1987.

G. L. Torrisi. Probability approximation of point processes with Papangelou conditional intensity. *Bernoulli*, 23:2210–2256, 2017.

G. T. Toussaint. *Computational Geometric Problems in Pattern Recognition. In: Pattern Recognition Theory and Applications.* Reidel, Dordrecht, 1982. ISBN 978-94-009-7774-7.

I. S. Tyurin. New estimates of the convergence rate in the Lyapunov theorem. 2009. URL http://arxiv.org/abs/0912.0726.

R. Vitale. Covariances of symmetric statistics. *Journal of Multivariate Analysis*, 41:14–26, 1992.

W. Weil. Point processes of cylinders, particles and flats. *Acta Applicandae Mathematicae*, 6:103–136, 1987.

N. Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60: 897–936, 1938.

J. E. Yukich. *Probability Theory of Classical Optimization Problems.* Lecture Notes in Mathematics. Springer-Verlag, Berlin Heidelberg, 1998. ISBN 3-540-63666-8.

S. Zuyev. *Stochastic Geometry and Telecommunications Networks. In: Kendall, W. S., Molchanov, I. (eds.) New Perspectives in Stochastic Geometry.* Oxford University Press, Oxford, 2010. ISBN 978-0-19-923257-4.

List of Figures

Index of Notation