Limit theory for geometric statistics of clustering point processes

B. Błaszczyszyn [1], D. Yogeshwaran [†2], and J. E. Yukich [‡3]

[1]Inria/ENS 2 rue Simone Iff CS 42112 75589 Paris Cedex 12 France.
Bartek.Blaszczyszyn@ens.fr
[2]Statistics and Mathematics Unit, Indian Statistical Institute, Bangalore - 560059, India.
d.yogesh@isibang.ac.in
[3]Department of Mathematics, Lehigh University, Bethlehem PA 18015, USA.
joseph.yukich@lehigh.edu

July 12, 2016

Abstract

Let $\mathcal{P}$ be a simple, stationary, clustering point process on $\mathbb{R}^d$ in the sense that its correlation functions factorize up to an additive error decaying exponentially fast with the separation distance. Let $\mathcal{P}_n := \mathcal{P} \cap W_n$ be its restriction to windows $W_n := [-\frac{n^{1/d}}{2}, \frac{n^{1/d}}{2}]^d \subset \mathbb{R}^d$. We consider the statistic $H_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n)$ where $\xi(x, \mathcal{P}_n)$ denotes a score function representing the interaction of x with respect to $\mathcal{P}_n$. When ξ depends on local data in the sense that its radius of stabilization has an exponential tail, we establish expectation asymptotics, variance asymptotics, and central limit theorems for H_n^ξ and, more generally, for statistics of the random measures $\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n) \delta_{n^{-1/d}x}$, as $W_n \uparrow \mathbb{R}^d$. This gives the limit theory for non-linear geometric statistics (such as clique counts, the number of Morse critical points, intrinsic volumes of the Boolean model, and total edge length of the k-nearest neighbor graph) of determinantal point processes having fast decreasing kernels, including the β-Ginibre ensembles, extending the Gaussian fluctuation results of Soshnikov [68] to non-linear statistics. It also gives the limit theory for geometric U-statistics of α-permanental point processes (for $1/\alpha \in \mathbb{N}$), α-determinantal point processes (for $-1/\alpha \in \mathbb{N}$), as well as the zero set

[†]Research supported in part by DST-INSPIRE faculty award and TOPOSYS Grant.
[‡]Research supported in part by NSF grant DMS-1406410

of Gaussian entire functions, extending the central limit theorems of Nazarov and Sodin [51] and Shirai and Takahashi [67], which are also confined to linear statistics. The proof of the central limit theorem relies on a factorial moment expansion originating in [11, 12] to show clustering of mixed moments of ξ. Clustering extends the cumulant method to the setting of purely atomic random measures, yielding the asymptotic normality of μ_n^ξ.

Key words and phrases. Clustering point process, determinantal point process, permanental point process, Gaussian entire functions, Gibbs point process, U-statistics, stabilization, difference operators, cumulants, central limit theorem.

AMS 2010 Subject Classifications. Primary: 60F05 Central limit and other weak theorems, 60D05 Geometric probability and stochastic geometry; Secondary: 60G55 Point processes, 52A22 Random convex sets and integral geometry, 05C80 Random graphs.

Contents

1 Introduction and main results

Functionals of large geometric structures on finite point sets $\mathcal{X} \subset \mathbb{R}^d$ often consist of sums of spatially dependent terms admitting the representation

$$\sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}), \tag{1.1}$$

where the $\mathbb{R}$-valued *score function* ξ, defined on pairs $(x, \mathcal{X})$, $x \in \mathcal{X}$, represents the *interaction* of x with respect to $\mathcal{X}$, called the *input*. The sums (1.1) typically describe a global geometric feature of a structure on $\mathcal{X}$ in terms of local contributions $\xi(x, \mathcal{X})$.

It is frequently the case in stochastic geometry, statistical physics, and spatial statistics that one seeks the large n limit behavior of $\sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$, where ξ is an appropriately chosen score function and where $\mathcal{X}_n$ is a point process on $W_n := [-\frac{1}{2}n^{1/d}, \frac{1}{2}n^{1/d}]^d$. For example if $\mathcal{X}_n$ is either a Poisson or binomial point process and if ξ is either a local U-statistic or an exponentially stabilizing score function, then the limit theory for $\sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$ is established in [7, 22, 37, 40, 56, 58, 61, 62]. If $\mathcal{X}_n$ is a rarified Gibbs point process on W_n and ξ is exponentially stabilizing, then [65, 69] treat the limit theory for $\sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$.

It is natural to ask whether the limit theory of these papers extends to more general input $\mathcal{X}_n$ satisfying a notion of 'asymptotic independence' for point processes. Recall that if $\xi \equiv 1$ and if $\mathcal{X}_n$ is an α-determinantal point process on W_n with $\alpha = -1/m$ or an α-permanental point process on W_n with $\alpha = 2/m$ for some m in the set of positive integers $\mathbb{N}$ (respectively $\mathcal{X}_n$ is the restriction of the zero set of a Gaussian entire function to W_n), then remarkable results of Soshnikov [68], Shirai and Takahashi [67] (respectively Nazarov and Sodin [51]), show that the counting statistic $\mathcal{X}_n(W_n) := \sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$ is asymptotically normal. One may wonder whether asymptotic normality still holds when ξ is either a local U-statistic or an exponentially stabilizing score function. We answer these questions affirmatively. Loosely speaking, our approach shows that $\sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$ is asymptotically normal whenever $\mathcal{X}_n$ is a clustering point process.

Heuristically, when the score functions depend on 'local data' and when the input is 'asymptotically independent', one expects that the statistics $\sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$ obey a strong law and a central limit theorem. The notion of dependency on 'local data' for score functions is formalized via stabilization in [7, 22, 56, 58, 61] and here, adopting the notion of 'clustering' point processes as it arises in statistical physics [41, 43, 51], we formalize the notion of asymptotic independence in the setting of $\sum_{x \in \mathcal{X}_n} \xi(x, \mathcal{X}_n)$. Whereas mixing coefficients also formalize the notion of asymptotic independence [31]-[33], we find that the notion of clustering point processes aptly facilitates the generalization of the limit theory of the afore-mentioned papers. A point process $\mathcal{P}$ on $\mathbb{R}^d$ is clustering if for all $p, q \in \mathbb{N}$ and all $x_1, \ldots, x_{p+q} \in \mathbb{R}^d$, its correlation functions $\rho^{(p+q)}(x_1, \ldots, x_{p+q})$ factorize into $\rho^{(p)}(x_1, \ldots, x_p)\rho^{(q)}(x_{p+1}, \ldots, x_{p+q})$ up to an additive error decaying exponentially fast with the separation distance

$$s := d(\{x_1, \ldots, x_p\}, \{x_{p+1}, \ldots, x_{p+q}\}) := \inf_{i \in \{1,\ldots,p\}, j \in \{p+1,\ldots,p+q\}} |x_i - x_j| \tag{1.2}$$

as at (1.7) below. Roughly speaking, a clustering point process exhibits asymptotic independence at large distances. Examples of such point processes are given in Section 2.2. The terminology 'clustering point process' is perhaps not optimal, since, at least from the point of view of spatial statistics, it suggests that points of $\mathcal{P}$ clump or aggregate together, which is not necessarily the case. We have retained this terminology to maintain consistency with existing definitions in statistical physics [41, 43, 51].

If $\mathcal{P}_n := \mathcal{P} \cap W_n$, where $\mathcal{P}$ is a simple, stationary, clustering point process on $\mathbb{R}^d$ and if ξ is either a local U-statistic or an exponentially stabilizing score function, then our main results establish expectation and variance asymptotics, as well as central limit theorems for the (signed) random measures

$$\mu_n^\xi := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n) \delta_{n^{-1/d}x}, \tag{1.3}$$

as well as for their total mass given by the non-linear statistics

$$H_n^\xi := H_n^\xi(\mathcal{P}) := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}_n) \tag{1.4}$$

as $n \to \infty$. Here δ_x is the point mass at x. As shown in Theorems 1.11-1.14 this yields the limit theory for general non-linear statistics of determinantal and permanental point processes, the point process given by the zero set of a Gaussian entire function, as well as rarified Gibbsian input.

The benefit of the general approach taken here is three-fold: (i) we establish the asymptotic normality of the non-linear statistics μ_n^ξ, with $\mathcal{P}$ either an α-permanental point process (with $1/\alpha \in \mathbb{N}$), an α-determinantal point process (with $-1/\alpha \in \mathbb{N}$), or the zero set of a Gaussian entire function, thereby extending the work of Soshnikov [68], Shirai and Takahashi [67], and Nazarov and Sodin [51], who restrict to linear statistics, (ii) we extend the limit theory of [7, 40, 56, 58, 61], which is confined to Poisson and binomial input, to clustering point processes and (iii) we apply our general results to deduce asymptotic normality and variance asymptotics for statistics of simplicial complexes and germ-grain models, clique counts, Morse critical points, as well of statistics of random graphs on clustering input on expanding windows $W_n, n \to \infty$ (Section 2.3).

Given clustering input $\mathcal{P}$, an interesting feature of the measures μ_n^{ξ} is that their variances are at most of order $\mathrm{Vol}(W_n)$, the volume of the window W_n (Theorem 1.11). This holds also for the statistic $\hat{H}_n^{\xi} := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P})$, which involves summands having no boundary effects. An interesting feature of this statistic is that if its variance is $o(\mathrm{Vol}(W_n))$ then it has to be $O(\mathrm{Vol}_{d-1}(\partial W_n))$, where ∂W_n denotes the boundary of W_n and $\mathrm{Vol}_{d-1}(.)$ stands for the $(d-1)$th intrinsic volume (Theorem 1.14). In other words, if the fluctuations of $\hat{H}_n^{\xi}$ are not of volume order, then they are at most of surface order.

Our interest in these issues was stimulated by similarities in the methods of [41], [6, 7, 65] and [51]. These papers all use clustering of mixed moments of score functions (see (1.5) and (1.17)) and the classical cumulant method. The articles [7, 65] prove central limit theorems for stabilizing functionals of Poisson and rarified Gibbsian point processes, respectively, while [51] proves central limit theorems for linear statistics $\sum_{x \in \mathcal{P}_n} \xi(x)$ of clustering point processes. This paper unifies and extends the results of [6, 7, 51, 65] to more general input. The earlier work of [43] is not only a precursor to our paper, but has also stimulated our investigation of variance asymptotics. The idea of using clustering to show asymptotic normality via cumulants goes back to [41].

Coming back to our set-up, when a functional $H_n^{\xi}(\mathcal{P})$ is expressible as a sum of local U-statistics or, more generally, as a sum of exponentially stabilizing score functions ξ, then a key step towards proving the central limit theorem is to show that the mixed moments, defined via Palm expectations $\mathbb{E}_{x_1,\ldots,x_k}$ (cf Section 1.1) and given by

$$m^{(k_1,\ldots,k_{p+q})}(x_1,\ldots,x_{p+q};n) := \mathbb{E}_{x_1,\ldots,x_{p+q}}(\xi(x_1,\mathcal{P}_n)^{k_1}\ldots\xi(x_{p+q},\mathcal{P}_n)^{k_{p+q}})\rho^{(p+q)}(x_1,\ldots,x_{p+q}) \tag{1.5}$$

approximately factorize into $m^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p;n)m^{(k_{p+1},\ldots,k_{p+q})}(x_{p+1},\ldots,x_{p+q};n)$, up to an additive error decaying exponentially fast with $s := d(\{x_1,\ldots,x_p\},\{x_{p+1},\ldots,x_{p+q}\})$. Here $x_1,...,x_{p+q}$ are distinct points in W_n and $k_1,...,k_{p+q} \in \mathbb{N}$. This result, spelled out in Theorem 1.10, is at the heart of our approach. We then give two proofs of the central limit theorem (Theorem 1.12) for purely atomic random measures via the cumulant method and as a corollary derive the asymptotic normality of $H_n^{\xi}(\mathcal{P})$ and $\int f \mathrm{d}\mu_n^{\xi}$, f a test function, as $n \to \infty$. The proof of expectation and variance asymptotics (Theorem 1.11) mainly relies upon the refined Campbell theorem.

In contrast to the afore-mentioned works, our approach to clustering of mixed moments depends heavily on a factorial moment expansion for expected values of functionals of a general point process $\mathcal{P}$. This expansion, which originates in [11, 12], is expressed in terms of iterated difference operators of the considered functional on the null configuration of points and integrated against factorial moment measures of the point process. It is valid for general point processes, in contrast to the Fock space representation of Poisson functionals, which involves the same difference operators but is deeply related to chaos expansions [38]. Further connections with the literature are discussed in the remarks after Theorems 1.13 and 1.14.

Having described the goals and context of this paper, we now describe more precisely the assumptions on allowable score and input pairs $(\xi, \mathcal{P})$ as well as our main results. The generality of allowable pairs $(\xi, \mathcal{P})$ considered here necessitates several definitions which go as follows.

1.1 Admissible clustering point processes

Throughout $\mathcal{P} \subset \mathbb{R}^d$ denotes a simple point process. By a simple point process we mean a random element taking values in $\mathcal{N}$, the space of locally finite simple point sets in $\mathbb{R}^d$ (or equivalently Radon counting measures μ such that $\mu(\{x\}) \in \{0, 1\}$ for all $x \in \mathbb{R}^d$) and equipped with the canonical σ-algebra $\mathcal{B}$. Given a simple point process $\mathcal{P}$ we interchangeably use the following representations of $\mathcal{P}$:

$$\mathcal{P}(\cdot) := \sum_i \delta_{X_i}(\cdot) \text{ (random measure); } \mathcal{P} := \{X_i\}_{i\geq 1} \text{ (random set),}$$

where $X_i, i \geq 1$, are $\mathbb{R}^d$-valued random variables (given a measurable numbering of points, which is irrelevant for the results presented in this paper). Points of $\mathbb{R}^d$ are denoted by x or y whereas points of $\mathbb{R}^{d(k-1)}$ are denoted by $\mathbf{x}$ or $\mathbf{y}$. We let $\mathbf{0}$ denote a point at the origin of $\mathbb{R}^d$.

For a bounded function f on $\mathbb{R}^d$ and a measure μ, let $\mu(f) := \langle f, \mu \rangle$ denote the integral of f with respect to μ. For a bounded set $B \subset \mathbb{R}^d$ we let $\mu(B) = \mu(\mathbf{1}_B) = \text{card}(\mu \cap \text{B})$, with μ in the last expression interpreted as the set of its atoms.

For a simple Radon counting measure μ and $k \in \mathbb{N}$, the kth factorial power is

$$\mu^{(k)} := \begin{cases} \sum_{\text{distinct}\, x_1,\dots,x_k \in \mu} \delta_{(x_1,\dots,x_k)} & \text{when } \mu(\mathbb{R}^d) \geq k, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $\mu^{(k)}$ is a Radon counting measure on $(\mathbb{R}^d)^k$. Consistently, for a set $\mathcal{X} \subset \mathbb{R}^d$, we denote $\mathcal{X}^{(k)} := \{(x_1, \dots, x_k) \in (\mathbb{R}^d)^k : x_i \in \mathcal{X}, x_i \neq x_j \text{ for } i \neq j\}$. The *$k$th order factorial moment measure* of the (simple) point process $\mathcal{P}$ is defined as $\alpha^{(k)}(\cdot) := \mathbb{E}(\mathcal{P}^{(k)}(\cdot))$ on $(\mathbb{R}^d)^k$ i.e., $\alpha^{(k)}(\cdot)$ is the intensity measure of the point process $\mathcal{P}^{(k)}(\cdot)$. Its Radon-Nikodyn density $\rho^{(k)}(x_1, ..., x_k)$ (provided it exists) is the *k-point correlation* *function* and is characterized by the relation

$$\alpha^{(k)}(B_1 \times \cdots \times B_k) = \mathbb{E}\Big(\prod_{1\le i\le k} \mathcal{P}(B_i)\Big) = \int_{B_1\times\cdots\times B_k} \rho^{(k)}(x_1,...,x_k)\,\mathrm{d}x_1\ldots\mathrm{d}x_k,$$

where $B_1,...,B_k$ are mutually disjoint bounded Borel sets in $\mathbb{R}^d$. Since $\mathcal{P}$ is simple, we may put $\rho^{(k)}$ to be zero on the diagonals of $(\mathbb{R}^d)^k$, that is on the subsets of $(\mathbb{R}^d)^k$ where two or more coordinates coincide.

Heuristically, the *kth Palm measure* $P_{x_1,\ldots,x_k}$ of $\mathcal{P}$ is the probability distribution of $\mathcal{P}$ conditioned on $\{x_1,\ldots,x_k\} \subset \mathcal{P}$. More formally, if $\alpha^{(k)}$ is locally finite, there exists a family of probability distributions $P_{x_1,\ldots,x_k}$ on $(\mathcal{N},\mathcal{B})$, unique up to an $\alpha^{(k)}$-null set of $(\mathbb{R}^d)^k$, called the kth Palm measures of $\mathcal{P}$, and satisfying the disintegration formula

$$\mathbb{E}\Big(\sum_{(x_1,\ldots,x_k)\in\mathcal{P}^{(k)}} f(x_1,\ldots,x_k;\mathcal{P})\Big) = \int_{\mathbb{R}^{dk}}\int_{\mathcal{N}} f(x_1,\ldots,x_k;\mu)\,P_{x_1,\ldots,x_k}(\mathrm{d}\mu)\alpha^{(k)}(\mathrm{d}x_1,\ldots\mathrm{d}x_k) \tag{1.6}$$

for any (say non-negative) measurable function f on $(\mathbb{R}^d)^k \times \mathcal{N}$. Formula (1.6) is also known as the refined Campbell theorem.

To simplify notation, write $\int_{\mathcal{N}} f(x_1,\ldots,x_k;\mu)P_{x_1,\ldots,x_k}(\mathrm{d}\mu) = \mathbb{E}_{x_1,\ldots,x_k}(f(x_1,\ldots,x_k;\mathcal{P}))$, where $\mathbb{E}_{x_1,\ldots,x_k}$ is the expectation corresponding to the Palm probability $\mathbb{P}_{x_1,\ldots,x_k}$ on a canonical probability space on which $\mathcal{P}$ is defined. To further simplify notation, denote by $\mathbb{P}^!_{x_1,\ldots,x_k}$ the *reduced Palm probabilities* and their expectation by $\mathbb{E}^!_{x_1,\ldots,x_k}$, which satisfies $\mathbb{E}^!_{x_1,\ldots,x_k}(f(x_1,\ldots,x_k;\mathcal{P})) = \mathbb{E}_{x_1,\ldots,x_k}(f(x_1,\ldots,x_k;\mathcal{P}\setminus\{x_1,\ldots,x_k\}))$ [1].

All Palm probabilities (expectations) are meaningfully defined only for $\alpha^{(k)}$ almost all $x_1,\ldots,x_k \in \mathbb{R}^d$. Consequently, all expressions involving these measures should be understood in the $\alpha^{(k)}$ a.e. sense. Similarly the considered suprema should be understood as *essential* suprema with respect to $\alpha^{(k)}$.

The following definition is reminiscent of the so-called weak exponential decrease of correlations introduced in [41] and subsequently used in [43, 51].

Definition 1.1 (Clustering of correlation functions)**.** *The correlation functions of the point process $\mathcal{P}$ are said to* cluster *if there exists a* fast decreasing clustering function *$\phi : \mathbb{R}^+ \to \mathbb{R}^+$ (i.e., $\phi \le 1$, ϕ decreasing, and $\lim_{x\to\infty} x^m\phi(x) = 0$ for all $m \ge 1$) such that for all $k \in \mathbb{N}$ there are strictly positive* clustering constants *c_k and C_k such that for all $p,q \in \mathbb{N}$ and all $(x_1,\ldots,x_{p+q}) \in \mathbb{R}^{d(p+q)}$ we have*

$$|\rho^{(p+q)}(x_1,\ldots,x_{p+q}) - \rho^{(p)}(x_1,\ldots,x_p)\rho^{(q)}(x_{p+1},\ldots,x_{p+q})| \le C_{p+q}\phi(c_{p+q}s), \tag{1.7}$$

where $s := d(\{x_1,\ldots,x_p\},\{x_{p+1},\ldots,x_{p+q}\})$ is as at (1.2). Without loss of generality, we assume that c_k is non-increasing in k, and that C_k is finite and non-decreasing in k.

[1] It can be shown that $\mathbb{P}_{x_1,\ldots,x_k}(x_1,\ldots,x_k \in \mathcal{P}) = 1$ for $\alpha^{(k)}$ a.e. $x_1,\ldots,x_k \in \mathbb{R}^d$.

Definition 1.2 (Admissible clustering point process)**.** *By an admissible point process $\mathcal{P}$ on $\mathbb{R}^d, d \geq 2$, we mean that $\mathcal{P}$ is simple, stationary (i.e., $\mathcal{P}+x \stackrel{d}{=} \mathcal{P}$ for all $x \in \mathbb{R}^d$, where $\mathcal{P}+x$ denotes the translation of $\mathcal{P}$ by the vector x), with non-null and finite intensity $\rho^{(1)}(\mathbf{0}) = \mathbb{E}(\mathcal{P}(W_1))$, and has k-point correlation functions of all orders $k \in \mathbb{N}$. If its correlation functions cluster as in Definition 1.1, then $\mathcal{P}$ is an* admissible clustering point process*.*

Admissible clustering point processes are ubiquitous and include certain determinantal, permanental, and Gibbs point processes, as explained in Section 2.2. The k-point correlation functions of an admissible clustering point process are bounded i.e.,

$$\sup_{(x_1,\ldots,x_k)\in\mathbb{R}^{dk}} \rho^{(k)}(x_1,\ldots,x_k) \leq \kappa_k < \infty, \tag{1.8}$$

for some constants κ_k, which without loss of generality are non-decreasing in k. For stationary $\mathcal{P}$ with intensity $\rho^{(1)}(\mathbf{0}) \in (0,\infty)$ we have that (1.7) implies (1.8) with

$$\kappa_k \leq \rho^{(1)}(\mathbf{0}) \sum_{i=2}^{k} C_i \leq (k-1)\rho^{(1)}(\mathbf{0}) C_k. \tag{1.9}$$

1.2 Admissible score functions

Throughout we restrict to translation-invariant score functions $\xi : \mathbb{R}^d \times \mathcal{N} \to \mathbb{R}$, i.e., those which are measurable in each coordinate, $\xi(x,\mathcal{X}) = 0$ if $x \notin \mathcal{X} \in \mathcal{N}$, and for all $y \in \mathbb{R}^d$, satisfy $\xi(\cdot + y, \cdot + y) = \xi(\cdot,\cdot)$.

We introduce classes (A1) and (A2) of *admissible* score and input pairs $(\xi,\mathcal{P})$. Specific examples of admissible input pairs of both the classes are provided in Sections 2.2 and 2.3. The first class allows for admissible input $\mathcal{P}$ as in Definition 1.2 whereas the second considers admissible input $\mathcal{P}$ satisfying clustering (1.7), subject to $c_k \equiv 1$ and growth conditions on the clustering constants C_k and the clustering function ϕ.

Definition 1.3 (Class (A1) of admissible score and input pairs $(\xi,\mathcal{P})$)**.** *Admissible input $\mathcal{P}$ consists of admissible clustering point processes as in Definition 1.2. Admissible score functions are of the form*

$$\xi(x,\mathcal{X}) := \frac{1}{k!} \sum_{\mathbf{x}\in\mathcal{X}^{(k-1)}} h(x,\mathbf{x}), \tag{1.10}$$

for some $k \in \mathbb{N}$ and a symmetric, translation-invariant function $h : \mathbb{R}^d \times (\mathbb{R}^d)^{k-1} \to \mathbb{R}$ such that $h(x_1,\ldots,x_k) = 0$ whenever either $\max_{2\leq i\leq k} |x_i - x_1| > r$ for some given $r > 0$ or when $x_i = x_j$ for some $i \neq j$. When $k = 1$, we set $\xi(x,\mathcal{X}) = h(x)$. Further, assume

$$\|h\|_\infty := \sup_{\mathbf{x}\in\mathbb{R}^{d(k-1)}} |h(\mathbf{0},\mathbf{x})| < \infty.$$

The interaction range for h is at most r, showing that the functionals H_n^ξ defined at (1.4) generated via scores (1.10) are local *U-statistics* of order k as in [62]. Before introducing a more general class of score functions, we recall [7, 40, 56, 58, 61] a few definitions formalizing the notion of the local dependence of ξ on its input. Let $B_r(x) := \{y : |y - x| \leq r\}$ denote the ball of radius r centered at x, and $B_r^c(x)$ its complement.

Definition 1.4 (Radius of stabilization). *Given a score function ξ, input $\mathcal{X}$, and $x \in \mathcal{X}$, define the radius of stabilization $R^\xi(x, \mathcal{X})$ to be the smallest $r \in \mathbb{N}$ such that*

$$\xi(x, \mathcal{X} \cap B_r(x)) = \xi(x, (\mathcal{X} \cap B_r(x)) \cup (\mathcal{A} \cap B_r^c(x)))$$

for all $\mathcal{A} \subset \mathbb{R}^d$ locally finite. If no such finite r exists, we set $R^\xi(x, \mathcal{X}) = \infty$.

If ξ is a translation invariant score function then so is $R^\xi(x, \mathcal{X})$. Score functions (1.10) of class (A1) have radius of stabilization upper-bounded by r.

Definition 1.5 (Stabilizing score function). *We say that ξ* is stabilizing on $\mathcal{P}$ *if for all $l \in \mathbb{N}$ there are constants $a_l > 0$, such that*

$$\sup_{1 \leq n \leq \infty} \sup_{x_1, \ldots, x_l \in W_n} \mathbb{P}_{x_1, \ldots, x_l}\big(R^\xi(x_1, \mathcal{P}_n) > t\big) \leq \varphi(a_l t) \tag{1.11}$$

with $\varphi(t) \downarrow 0$ as $t \to \infty$. Without loss of generality the a_l are non-increasing in l and $0 \leq \varphi \leq 1$. In (1.11) and elsewhere, we adopt the convention that $W_\infty := \mathbb{R}^d$ and $\mathcal{P}_\infty := \mathcal{P}$. The second sup in (1.11) is understood as ess sup *with respect to $\alpha^{(l)}$.*

Definition 1.6 (Exponentially stabilizing score function). *We say that ξ is* exponentially stabilizing *on $\mathcal{P}$ if ξ is stabilizing on $\mathcal{P}$ as in Definition 1.5 with φ satisfying*

$$\liminf_{t \to \infty} \frac{\log \varphi(t)}{t^c} < 0 \tag{1.12}$$

for some $c \in (0, \infty)$.

We define a general class of score functions exponentially stabilizing on their input.

Definition 1.7 (Class (A2) of admissible score and input pairs $(\xi, \mathcal{P})$). *Admissible input $\mathcal{P}$ consists of admissible clustering point processes as in Definition 1.2 with clustering constants satisfying $c_k \equiv 1$,*

$$C_k = O(k^{ak}), \tag{1.13}$$

for some $a \in [0, 1)$ and the clustering function ϕ satisfying the growth condition

$$\liminf_{t \to \infty} \frac{\log \phi(t)}{t^b} < 0 \tag{1.14}$$

for some constant $b \in (0, \infty)$. Admissible score functions ξ for this class are exponentially stabilizing on the input $\mathcal{P}$ and satisfy a power growth condition*, namely there* *exists* $\hat{c} \in [1, \infty)$ *such that for all* $r \in (0, \infty)$

$$|\xi(x, \mathcal{X} \cap B_r(x))| \mathbf{1}[\mathrm{card}(\mathcal{X} \cap B_r(x)) = n] \leq (\hat{c} \max(r, 1))^n. \tag{1.15}$$

The condition $c_k \equiv 1$ is equivalent to $c_* := \inf c_k > 0$. This follows since we may replace the fast decreasing function $\phi(.)$ by $\phi(c_* \times \cdot)$, with $c_k \equiv 1$ for this new fast decreasing function. Score functions of class (A1) also satisfy the power growth condition (1.15) since in this case the left hand side of (1.15) is at most $\|h\|_\infty n^{(k-1)}/k$. Thus the generalization from (A1) to (A2) consists in replacing local U-statistics by exponentially stabilizing score functions satisfying the power growth condition. This is done at the price of imposing stronger conditions on the input process, requiring in particular that it has finite exponential moments, as explained in Section 2.1.

1.3 Strong clustering of mixed moments

The following p-moment condition involves the score function ξ and the input $\mathcal{P}$. We shall describe in Section 2.1 ways to control the p-moments of input pairs of class (A1) and (A2).

Definition 1.8 (Moment condition)**.** *Given* $p \in [1, \infty)$*, say that the pair* $(\xi, \mathcal{P})$ *satisfies the* p*-moment condition if*

$$\sup_{1 \leq n \leq \infty} \sup_{1 \leq p' \leq \lfloor p \rfloor} \sup_{x_1, \ldots, x_{p'} \in W_n} \mathbb{E}_{x_1, \ldots, x_{p'}} \max\{|\xi(x_1, \mathcal{P}_n)|, 1\}^p \leq \tilde{M}_p < \infty \tag{1.16}$$

for some constant $\tilde{M}_p := \tilde{M}_p^\xi$*, where* $\sup$ *signifies* $\mathrm{ess\,sup}$ *with respect to* $\alpha^{(p)}$*. Without loss of generality we assume that* $\tilde{M}_p$ *is increasing in* p *for all* p *such that* (1.16) *holds.*

Recall the generalized mixed moments $m^{(k_1, \ldots, k_p)}(x_1, \ldots, x_p; n)$ as defined in (1.5). When $k_i \equiv 1$ for all $1 \leq i \leq p$, we write $m_{(p)}(x_1, ..., x_p; n)$ instead of $m^{(1, \ldots, 1)}(x_1, ..., x_p; n)$. Abbreviate $m^{(k_1, \ldots, k_p)}(x_1, \ldots, x_p; \infty)$ by $m^{(k_1, \ldots, k_p)}(x_1, \ldots, x_p)$. These generalized mixed moments exist whenever (1.16) is satisfied for p set to $k_1 + \ldots + k_p$ and provided the p-point correlation function $\rho^{(p)}$ exists.

Definition 1.9 (Strong clustering of mixed moments)**.** *Say that the mixed moments for* ξ *strongly cluster if there exists a fast decreasing function* $\tilde{\phi}$ *and clustering constants* $\tilde{C}_k < \infty$*,* $\tilde{c}_k < \infty, k \in \mathbb{N}$ *such that for all* $n \in \mathbb{N} \cup \{\infty\}$*,* $p, q \in \mathbb{N}$ *and any collection of positive integers* $k_1, \ldots, k_{p+q}$*, we have*

$$\Big| m^{(k_1, \ldots, k_{p+q})}(x_1, \ldots, x_{p+q}; n) - \\ m^{(k_1, \ldots, k_p)}(x_1, \ldots, x_p; n)\, m^{(k_{p+1}, \ldots, k_{p+q})}(x_{p+1}, \ldots, x_{p+q}; n) \Big| \leq \tilde{C}_K \tilde{\phi}(\tilde{c}_K s)\,, \tag{1.17}$$

where $K := \sum_{i=1}^{p+q} k_i$ *and* $s := d\big(\{x_1, \ldots, x_p\}, \{x_{p+1}, \ldots, x_{p+q}\}\big)$.

Our first theorem shows that strong clustering of generalized mixed moments holds for a wide class of score functions and input. This key result forms the starting point of our approach. We shall remark further on this result after Theorem 1.13.

Theorem 1.10. *Let $(\xi, \mathcal{P})$ be an admissible score and input pair of class (A1) or (A2) such that the p-moment condition (1.16) holds for all $p \in (1, \infty)$. Then the mixed moment functions for ξ strongly cluster as at (1.17).*

We prove this theorem in Section 3, where it is also shown that it subsumes more specialized clustering results of [7, 65].

1.4 Main results

We give the limit theory for the measures $\mu_n^\xi, n \geq 1$, and the non-linear statistics $H_n^\xi(.), n \geq 1$, defined at (1.3) and (1.4), respectively. Given a score function ξ on admissible input $\mathcal{P}$ we set [2]

$$\sigma^2(\xi) := \mathbb{E}_{\mathbf{0}} \xi^2(\mathbf{0}, \mathcal{P}) \rho^{(1)}(\mathbf{0}) + \int_{\mathbb{R}^d} (m_{(2)}(\mathbf{0}, x) - m_{(1)}(\mathbf{0})^2) \, \mathrm{d}x. \tag{1.18}$$

The following result provides expectation and variance asymptotics for $\mu_n^\xi(f)$, with f belonging to the space $\mathcal{B}(W_1)$ of bounded measurable functions on W_1.

Theorem 1.11. *Let $\mathcal{P}$ be an admissible point process on $\mathbb{R}^d$.*
(i) If ξ satisfies exponential stabilization (1.12) and the p-moment condition (1.16) for some $p \in (1, \infty)$ then for all $f \in \mathcal{B}(W_1)$

$$\left| n^{-1} \mathbb{E} \mu_n^\xi(f) - \mathbb{E}_{\mathbf{0}} \xi(\mathbf{0}, \mathcal{P}) \rho^{(1)}(\mathbf{0}) \int_{W_1} f(x) \, \mathrm{d}x \right| = O(n^{-1/d}). \tag{1.19}$$

If ξ only satisfies stabilization (1.11) and the p-moment condition (1.16) for some $p \in (1, \infty)$, then the right hand side of (1.19) is $o(1)$.
(ii) Assume that the second correlation function $\rho^{(2)}$ of $\mathcal{P}$ exists and is bounded as in (1.8), that ξ satisfies (1.11), and that $(\xi, \mathcal{P})$ satisfies the p-moment condition (1.16) for some $p \in (2, \infty)$. If the second mixed moment $m^{(1,1)}$ for ξ strongly clusters, i.e. satisfies (1.17) with $p = q = k_1 = k_2 = 1$ and all $n \in \mathbb{N} \cup \{\infty\}$, then for all $f \in \mathcal{B}(W_1)$

$$\lim_{n \to \infty} n^{-1} \mathrm{Var} \mu_n^\xi(f) = \sigma^2(\xi) \int_{W_1} f(x)^2 \, \mathrm{d}x \in [0, \infty), \tag{1.20}$$

whereas for all $f, g \in \mathcal{B}(W_1)$

$$\lim_{n \to \infty} n^{-1} \mathrm{Cov}(\mu_n^\xi(f), \mu_n^\xi(g)) = \sigma^2(\xi) \int_{W_1} f(x) g(x) \, \mathrm{d}x. \tag{1.21}$$

[2] For a stationary point process $\mathcal{P}$, its Palm expectation $\mathbb{E}_{\mathbf{0}}$ (and consequently $m_{(1)}(\mathbf{0})$, $m_{(2)}(\mathbf{0}, x)\mathrm{d}x$) is meaningfully defined e.g. via the Palm-Matthes approach.

We remark that (1.19) and (1.20) together show convergence in probability

$$n^{-1}\mu_n^{\xi}(f) \xrightarrow{P} \mathbb{E}_{\mathbf{0}}\xi(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})\int_{W_1} f(x)\,\mathrm{d}x$$

as $n \to \infty$.

The proof of variance asymptotics (1.20) requires strong clustering of the second mixed moment. Strong clustering of *all* mixed moments yields Gaussian fluctuations of the purely atomic random measure μ_n^{ξ} under moment conditions on the atom sizes (i.e. under moment conditions on ξ) and a variance lower bound. Let $N(0,\sigma^2)$ denote a mean zero normal random variable with variance σ^2. Following Knuth's definition, in what follows we write $f(n) = \Omega(g(n))$ when $g(n) = O(f(n))$; i.e., when $\liminf_{n\to\infty} |f(n)/g(n)| > 0$.

Theorem 1.12. *Let $\mathcal{P}$ be an admissible point process on $\mathbb{R}^d$ and let the pair $(\xi,\mathcal{P})$ satisfy the p-moment condition (1.16) for all $p \in [1,\infty)$. If the mixed moments for ξ strongly cluster as at (1.17) and if $f \in \mathcal{B}(W_1)$ satisfies*

$$\mathrm{Var}\mu_n^{\xi}(f) = \Omega(n^{\nu}) \tag{1.22}$$

for some $\nu \in [0,\infty)$, then as $n \to \infty$

$$\frac{\mu_n^{\xi}(f) - \mathbb{E}\mu_n^{\xi}(f)}{\sqrt{\mathrm{Var}\mu_n^{\xi}(f)}} \xrightarrow{\mathcal{D}} N(0,1). \tag{1.23}$$

Combining Theorem 1.10 and Theorem 1.12 yields the following theorem, which is well-suited for off-the-shelf use in applications, as seen in Section 2.3.

Theorem 1.13. *Let $(\xi,\mathcal{P})$ be an admissible pair of class (A1) or (A2) such that the p-moment condition (1.16) holds for all $p \in (1,\infty)$. If $f \in \mathcal{B}(W_1)$ satisfies condition (1.22) for some $\nu \in (0,\infty)$, then $\mu_n^{\xi}(f)$ is asymptotically normal as in (1.23), as $n \to \infty$.*

Theorems 1.11 and 1.12 are proved in Section 4. We next compare our results with those in the literature. Definitions of point processes mentioned below are in Section 2.2.

Remarks:

(i) *Theorem 1.11.* In the case of Poisson and binomial input $\mathcal{P}$, the limits (1.19) and (1.20) are shown in [60] and [7, 56], respectively. In the case of Gibbsian input, the limits (1.19) and (1.20) are established in [65]. Theorem 1.11 shows these limits hold for general stationary input. For general stationary input, the paper [70] gives a weaker version of Theorem 1.11 for specific ξ and for $f = \mathbf{1}[x \in W_1]$. In full generality, the convergence rate (1.19) is new for any point process $\mathcal{P}$.

(ii) *Theorems 1.12 and 1.13.* Under condition (1.22),Theorems 1.12 and 1.13 provide a central limit theorem for non-linear statistics of either determinantal and permanental input with a fast-decaying kernel as at (2.7), the zero set $\mathcal{P}_{GEF}$ of a Gaussian entire function, or rarified Gibbsian input. When $\xi \equiv 1$, then $\mu_n^\xi(f)$ reduces to the *linear statistic* $\sum_{x \in \mathcal{P}_n} f(x)$. These theorems extend the central limit theorem for linear statistics of $\mathcal{P}_{GEF}$ as established in [51]. In the case that the input is determinantal with a fast decaying kernel as at (2.7), then Theorems 1.12 and 1.13 also extend the main result of Soshnikov [68], whose pathbreaking paper gives a central limit theorem for linear statistics for any determinantal input, provided the variance grows as least as fast as a power of the expectation. The generality of the score functionals considered here necessitates assumptions on the determinantal kernel which are more restrictive than those required by [68]. Proposition 5.7 of [67] shows central limit theorems for linear statistics of α-determinantal point processes with $\alpha = -1/m$ or α-permanental point processes with $\alpha = 2/m$ for some $m \in \mathbb{N}$. Theorems 1.12 and 1.13 extend these results in the case $|\alpha| = 1/m$.

(iii) *Variance lower bounds.* To prove asymptotic normality it is customary to require variance lower bounds as at (1.22); [51] and [68] both require assumptions of this kind. Showing condition (1.22) is a separate problem and it fails in general; recall that the variance of the point count of some determinantal point processes, including the GUE point process, grows at most logarithmically. This phenomena is especially pronounced in dimensions $d = 1, 2$. On the other hand, if $\xi \equiv 1$, and if the kernel K for a determinantal point process satisfies $\int_{\mathbb{R}^d} |K(\mathbf{0}, x)|^2 \mathrm{d}x < K(\mathbf{0}, \mathbf{0}) = \rho^{(1)}(\mathbf{0})$, then recalling the definition of $\sigma^2(\xi)$ at (1.18), we have $\sigma^2(\xi) = \sigma^2(1) = \rho^{(1)}(\mathbf{0}) - \int_{\mathbb{R}^d} |K(\mathbf{0}, x)|^2 \mathrm{d}x > 0$. In the case of rarified Gibbsian input, the bound (1.22) holds with $\nu = 1$, as shown in of [69, Theorem 1.1]. Theorem 1.13 allows for surface-order variance growth, which arises for linear statistics $\sum_{x \in \mathcal{P}_n} \xi(x)$ of determinantal point processes; see [24, (4.15)].

(iv) *Poisson, binomial, and Gibbs input.* When $\mathcal{P}$ is Poisson or binomial input and when ξ is a functional which stabilizes exponentially fast as at (1.12), then μ_n^ξ is asymptotically normal (1.23) under moment conditions on ξ; see the survey [72]. When $\mathcal{P}$ is a rarified Gibbs point process with 'ancestor clans' which decay exponentially fast, and when ξ is an exponentially stabilizing functional, then μ_n^ξ satisfies normal convergence (1.23) as established in [65, 69].

(v) *Mixing and clustering.* Central limit theorems for geometric functionals of mixing point processes (random fields) are established in [3, 15, 31, 33, 32]. The geometric functionals considered in these papers are different than the ones considered here; furthermore the relation between the mixing conditions (in these papers) and clustering (1.7) is unclear. Though correlation functions are simpler than mixing coefficients, which depend on σ-algebras generated by the point processes, our decay rates appear more restrictive than those needed in [3, 15, 31, 33, 32].

(vi) *Multivariate central limit theorem.* We may prove a multivariate central limit theorem via Theorems 1.11 and 1.13 and the Cramér-Wold device. This goes as follows. Let $(\xi, \mathcal{P})$ be a pair satisfying the hypotheses of Theorems 1.11 and 1.13. If $f_i \in \mathcal{B}_0(W_1), 1 \le i \le k$, satisfy the variance limit (1.20) with $\sigma^2(\xi) > 0$, then as $n \to \infty$ the fidis

$$\left(\frac{\mu_n^\xi(f_1) - \mathbb{E}\mu_n^\xi(f_1)}{\sqrt{n}}, \dots, \frac{\mu_n^\xi(f_k) - \mathbb{E}\mu_n^\xi(f_k)}{\sqrt{n}} \right)$$

converge to the fidis of a mean zero Gaussian field having covariance kernel $f, g \mapsto \sigma^2(\xi) \int_{W_1} f(x)g(x)\mathrm{d}x$.

(vii) *Deterministic radius of stabilization.* It may be shown that our main results go through without the condition (1.14) if the radius of stabilization $R^\xi(x, \mathcal{P})$ is bounded by a non-random (deterministic) constant, and if (1.13) and (1.15) are satisfied. However we are unable to find any interesting examples of point processes satisfying (1.7) but not (1.14).

(viii) *Clustering of mixed moments; Theorem 1.10.* Though the cumulant method is common to [7, 65, 51] and this article, a distinguishing and novel feature of our approach is the proof of strong clustering of mixed moment functions for a wide class of functionals and point processes. As mentioned in the introduction, the proof of this result is via factorial moment expansions, which differs from the approach of [7, 65, 51] (see the discussion at the beginning of Section 3). Strong clustering (1.17) appears to be of independent interest. It features in the proofs of moderate deviation principles and laws of the iterated logarithms for stabilizing functionals of Poisson point process, see [5], [21]. Strong clustering (1.17) yields cumulant bounds, useful in establishing concentration inequalities as well as moderate deviations, as explained in [27, Lemma 4.2].

(ix) *Normal approximation.* Difference operators (which appear in our factorial moment expansions) are also a key tool in the Malliavin-Stein method [52, 53]. This method has been highly successful in obtaining presumably 'optimal' rates of normal convergence for various statistics (including those considered in Section 2.3) in stochastic geometric problems [37, 40, 62]. However, these methods currently apply only to functionals defined on Poisson and binomial point processes. It is an open question whether a refined use of these methods would yield rates of convergence in our central limit theorems.

(x) *Cumulant bounds.* Our approach shows that the kth order cumulants for $\langle f, \mu_n^\xi \rangle$ grow *at most linearly* in n for $k \ge 1$. Thus, under ssumption (1.22), the cumulant C_n^k

for $(\mathrm{Var}\langle f, \mu_n^\xi \rangle)^{-1/2} \langle f, \mu_n^\xi \rangle$ satisfies $C_n^k \le D(k) n^{1-(\nu k/2)}$, with $D(k)$ depending only on k. For $k = 3, 4, ...$ and $\nu > 2/3$, we have $C_n^k \le D(k)/(\Delta(n))^{k-2}$, where $\Delta(n) := n^{(3\nu-2)/2}$. When $D(k)$ satisfies $D(k) \le (k!)^{1+\gamma}$, γ a constant, then we obtain the Berry-Esseen bound (cf. [27, Lemma 4.2])

$$\sup_{t \in \mathbb{R}} \left| \mathbb{P}\left(\frac{\mu_n^\xi(f) - \mathbb{E}\mu_n^\xi(f)}{\sqrt{\mathrm{Var}\mu_n^\xi(f)}} \le t \right) - \mathbb{P}(N(0,1) \le t) \right| = O(\Delta(n)^{-1/(1+2\gamma)}).$$

Determining conditions on input pairs $(\xi, \mathcal{P})$ insuring the bounds $\nu > 2/3$ and $D(k) \le (k!)^{1+\gamma}$, γ a constant, is beyond the scope of this paper. When $\mathcal{P}$ is Poisson input, this issue is addressed by [21].

We next consider the case when the fluctuations of $H_n^\xi(\mathcal{P})$ are not of volume order, that is to say $\sigma^2(\xi) = 0$. Though this may appear to be a degenerate condition, interesting examples involving determinantal point processes or zeros of GEF in fact satisfy $\sigma^2(1) = 0$. Such point processes are termed 'super-homogeneous point processes' [51, Remark 5.1]. Put

$$\widehat{H}_n^\xi(\mathcal{P}) := \sum_{x \in \mathcal{P}_n} \xi(x, \mathcal{P}). \tag{1.24}$$

The summands in $\widehat{H}_n^\xi(\mathcal{P})$, in contrast to those appearing in $H_n^\xi(\mathcal{P})$, are not sensitive to boundary effects. We shall show that under volume order scaling the asymptotic variance of $\hat{H}_n^\xi(\mathcal{P})$ also equals $\sigma^2(\xi)$. However, when $\sigma^2(\xi) = 0$ we derive surface order variance asymptotics for $\hat{H}_n^\xi(\mathcal{P})$. Though a similar result should plausibly hold for $H_n^\xi(\mathcal{P})$, a proof seems beyond the scope of the current paper. For $y \in \mathbb{R}^d$ and $W \subset \mathbb{R}^d$, put

$$\gamma_W(y) := \mathrm{Vol}(W \cap (\mathbb{R}^d \setminus W - y)) \tag{1.25}$$

and

$$\gamma(y) := \lim_{n \to \infty} \frac{\gamma_{W_n}(y)}{n^{d-1/d}}.$$

For a proof of existence of the function γ, see [43, Lemma 1(a)].

Theorem 1.14. *Under the assumptions of Theorem 1.11(ii) suppose also that the pair $(\xi, \mathcal{P})$ exponentially stabilizes as in (1.12). Then*

$$\lim_{n \to \infty} n^{-1} \mathrm{Var} H_n^\xi(\mathcal{P}) = \sigma^2(\xi). \tag{1.26}$$

If moreover $\sigma^2(\xi) = 0$ in (1.20) then

$$\lim_{n \to \infty} \frac{\mathrm{Var}\widehat{H}_n^\xi(\mathcal{P})}{n^{(d-1)/d}} = \sigma^2(\xi, \gamma) := \int_{\mathbb{R}^d} (m_{(1)}(\mathbf{0})^2 - m_{(2)}(\mathbf{0}, x)) \gamma(x) \, \mathrm{d}x \in [0, \infty). \tag{1.27}$$

Remarks:
(i) Checking positivity of $\sigma^2(\xi,\gamma) > 0$ is not always straightforward, though we note if ξ has the form (1.10), then the disintegration formula (1.6) yields

$$\begin{aligned}\sigma^2(\xi,\gamma) &= \sum_{j=0}^{k} \frac{1}{j!(k-j-1)!(k-j-1)!} \int_{\mathbb{R}^d} \gamma(x) \\ &\quad \Big(\int_{(B_r(\mathbf{0})\cap B_r(x))^j \times B_r(\mathbf{0})^{k-j-1} \times B_r(x)^{k-j-1}} h(\mathbf{0},\mathbf{y},\mathbf{z}) h(x,\mathbf{x},\mathbf{z}) \\ &\quad \big[\rho^{(k)}(\mathbf{0},\mathbf{y},\mathbf{z})\rho^{(k)}(x,\mathbf{x},\mathbf{z}) - \rho^{(2k-j)}(\mathbf{0},\mathbf{y},\mathbf{z},x,\mathbf{x})\big] \,\mathrm{d}\mathbf{z}\mathrm{d}\mathbf{y}\mathrm{d}\mathbf{x}\Big) \,\mathrm{d}x. \qquad (1.28)\end{aligned}$$

(ii) Theorem 1.11 and Theorem 1.14 extend [43, Propositions 1 and 2], which are valid only for $\xi \equiv 1$, to general functionals. If an admissible pair $(\xi,\mathcal{P})$ of type (A1) or (A2) is such that $\widehat{H}_n^{\xi}(\mathcal{P})$ does not have volume-order variance growth, then Theorems 1.11 and 1.14 show that $\widehat{H}_n^{\xi}(\mathcal{P})$ has at most surface-order variance growth.

2 Examples and applications

Before providing examples and applications of our general results, we briefly discuss the moment assumptions involved in our main theorems.

2.1 Moments of clustering point processes

We say that $\mathcal{P}$ has exponential moments if for all bounded Borel $B \subset \mathbb{R}^d$ and all $t \in \mathbb{R}^+$ we have

$$\mathbb{E}[t^{\mathcal{P}(B)}] < \infty\,. \qquad (2.1)$$

Similarly, say that $\mathcal{P}$ has all moments if for all bounded Borel $B \subset \mathbb{R}^d$ and all $k \in \mathbb{N}$, we have

$$\mathbb{E}[\mathcal{P}(B)^k] < \infty\,. \qquad (2.2)$$

Remarks:
(i) The point process $\mathcal{P}$ has exponential moments whenever $\sum_{k=1}^{\infty} \kappa_k t^k/k! < \infty$ for all $t \in \mathbb{R}^+$ with κ_k as in (1.8) (cf. the expansion of the probability generating function of a random variable in terms of factorial moments [17, Proposition 5.2.III.]). By (1.9) an admissible clustering point process has exponential moments provided

$$\sum_{k=1}^{\infty} \frac{C_k t^k}{k!} < \infty,\ \ t \in \mathbb{R}^+. \qquad (2.3)$$

Note that input of type (A2) has exponential moments since by (1.13), we have $C_k = O(k^{ak}), a \in [0,1)$, making (2.3) summable. For pairs $(\xi, \mathcal{P})$ of type (A2) with radius of stabilization bounded by $r_0 \in [1,\infty)$, by (1.15) the p-moment in (1.16) is consequently controlled by a finite exponential moment, i.e.,

$$\begin{aligned}
&\sup_{1\le n\le\infty}\ \sup_{1\le p'\le\lfloor p\rfloor}\ \sup_{x_1,\dots,x_{p'}\in W_n} \mathbb{E}_{x_1,\dots,x_{p'}} \max\{|\xi(x_1,\mathcal{P}_n)|,1\}^p \\
&\le \mathbb{E}_{x_1,\dots,x_{p'}}(\max\{\hat{c}r_0,1\}^{p\mathcal{P}(B_{r_0}(x_1))}). \qquad (2.4)
\end{aligned}$$

Finally, if $\mathcal{P}$ has exponential moments under its stationary probability $\mathbb{P}$, the same is true under $\mathbb{P}_{x_1,\dots,x_k}$ for $\alpha^{(k)}$ almost all $x_1,\dots,x_k$. [3]

(ii) For pairs $(\xi, \mathcal{P})$ of type (A1), the p-moment (1.16) satisfies

$$\begin{aligned}
&\sup_{1\le n\le\infty}\ \sup_{1\le p'\le\lfloor p\rfloor}\ \sup_{x_1,\dots,x_{p'}\in W_n} \mathbb{E}_{x_1,\dots,x_{p'}} \max\{|\xi(x_1,\mathcal{P}_n)|,1\}^p \\
&\le \Big(\frac{\|h\|_\infty}{k}\Big)^p \mathbb{E}_{x_1,\dots,x_{p'}}[(\mathcal{P}(B_r(x_1)))^{(k-1)p}]. \qquad (2.5)
\end{aligned}$$

We next show that (2.5) may be controlled by moments of Poisson random variables.

From the definition of factorial moment measures, we have for any Borel subset B that $\alpha^{(k)}(B) \le \kappa_k \mathrm{Vol}(B)$ where Vol(.) denotes the Lebesgue volume of a set. Since moments may be expressed as a linear combination of factorial moments, for $k \in \mathbb{N}$ and a bounded Borel subset $B \subset \mathbb{R}^d$ we have

$$\mathbb{E}[(\mathcal{P}(B))^k] = \sum_{j=0}^{k} \left\{ {k \atop j} \right\} \alpha^{(j)}(B^j) \le \kappa_k \sum_{j=0}^{k} \left\{ {k \atop j} \right\} \mathrm{Vol}(B)^j = \kappa_k \mathbb{E}(\mathrm{Po}(\mathrm{Vol}(B))^k), \qquad (2.6)$$

where $\left\{ {k \atop j} \right\}$ stand for the *Stirling numbers of the second kind*, $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with mean λ and where κ_j's are non-decreasing in j. Thus by (1.9), an admissible clustering point process has all moments, as in (2.2). If $\mathcal{P}$ has all moments under its stationary probability $\mathbb{P}$, the same is true under $\mathbb{P}_{x_1,\dots,x_k}$ for $\alpha^{(k)}$ almost all $x_1,\dots,x_k$ (by the same arguments as in Footnote 3).

2.2 Examples of clustering point processes

The notion of a stabilizing functional is well established in the stochastic geometry literature but since the notion of clustering is less well studied, we shall first convince

[3] Indeed, if $\mathbb{E}_{x_1,\dots,x_k}[\rho^{\mathcal{P}(B_r(x_1))}] = \infty$ for $x_1,\dots,x_k \in B'$ for some bounded $B' \in \mathbb{R}^d$ such that $\alpha^{(k)}(B'^k) > 0$ then $\mathbb{E}_{x_1,\dots,x_k}[\rho^{\mathcal{P}(B_r(x_1))}] \le \mathbb{E}_{x_1,\dots,x_k}[\rho^{\mathcal{P}(B'_r)}] = \infty$ with $B'_r = B' \oplus B_r(0) = \{y'+y : y' \in B', y \in B_r(0)\}$ the r-parallel set of B'. Integrating with respect to $\alpha^{(k)}$ in B'^k, by the Campbell formula $\mathbb{E}[(\mathcal{P}(B'_r))^k \rho^{\mathcal{P}(B'_r)}] = \infty$, which contradicts the existence of exponential moments under $\mathbb{P}$.

the reader that there are many interesting examples of admissible clustering point processes. For more details on the first five examples, we refer to [9].

2.2.1 Class A1 input

Permanental input. The point process $\mathcal{P}$ is permanental if its correlation functions are defined by $\rho^{(k)}(x_1, ..., x_k) := \mathrm{per}(K(x_i, x_j))_{1 \le i,j \le k}$, where the permanent of an $n \times n$ matrix M is $\mathrm{per}(M) := \sum_{\pi \in S_n} \Pi_{i=1}^{n} M_{i,\pi(i)}$, with S_n denoting the permutation group of the first n integers and $K(\cdot,\cdot)$ is the Hermitian kernel of a locally trace class integral operator $\mathcal{K} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$ [9, Assumption 4.2.3]. A kernel K is *fast-decreasing* if

$$|K(x, y)| \le \omega(|x - y|), \quad x, y \in \mathbb{R}^d, \tag{2.7}$$

for some fast decreasing $\omega : \mathbb{R}^+ \to \mathbb{R}^+$. Lemma 5.5 in Section 5 shows that if a stationary permanental point process has a fast-decreasing kernel as at (2.7), then it is an admissible clustering point process with clustering function $\phi = \omega$ and with clustering constants satisfying

$$C_k := kk!||K||^{k-1}, c_k \equiv 1, \tag{2.8}$$

where $||K|| := \sup_{x,y} |K(x, y)|$ and we can choose $\kappa_k = k!\|K\|^k$. However, a trace class permanental point process in general does not have exponential moments, i.e., the right-hand side of (2.1) might be infinite for some bounded B and ρ large enough. [4]

A useful property of the permanental point process with kernel K is that it can be represented as a Cox point process (see Section 2.2.3) with intensity field $\lambda(x) := Z_1(x)^2 + Z_2(x)^2$ where Z_1, Z_2 are i.i.d. Gaussian random fields with zero mean and covariance function $K/2$ [67, Thm 6.13]. Thus mean zero Gaussian random fields with a fast decaying covariance function $K/2$ yield a fast decaying clustering permanental (Cox) point process with kernel K.

α-Permanental point processes. See [9, Section 4.10], [44], and [67] for more details on this class of point processes which generalize permanental point processes. Given $\alpha \ge 0$ and a kernel K which is Hermitian, non-negative definite and locally trace class, a point process $\mathcal{P}$ is said to be *α-permanental* [5] if its correlation functions satisfy

$$\rho^{(k)}(x_1, \dots, x_k) = \sum_{\pi \in S_k} \alpha^{k-\nu(\pi)} \prod_{i=1}^{k} K(x_i, x_{\pi(i)}) \tag{2.9}$$

[4]This is because, the number of points of a (trace-class) permanental p.p. in a compact set B is a sum of independent geometric random variables $\mathrm{Geo}(1/(1+\lambda))$ where λ runs over all eigenvalues of the integral operator defining the process truncated to B.

[5]In contrast to terminology in [9, 67], here we distinguish the two cases (i) $\alpha \ge 0$ (α-permanental) and (ii) $\alpha \le 0$ (α-determinantal)

where S_k stands for the usual symmetric group and $\nu(.)$ denotes the number of cycles in a permutation. The right hand side is the α-permanent of the matrix $((K(x_i,x_j))_{i,j\le k}$. The special cases $\alpha=0$ and $\alpha=1$ respectively give the Poisson point process with intensity $K(\mathbf{0},\mathbf{0}))$ and the permanental point process with kernel K. In what follows, we assume $\alpha=1/m$ for $m\in\mathbb{N}$, i.e. $1/\alpha$ is a positive integer. Existence of such α-permanental point processes is guaranteed by [67, Theorem 1.2]. The property of these point processes most important to us is that an α-permanental point process with kernel K is a superposition of $1/\alpha$ i.i.d. copies of a permanental point process with kernel αK (see [9, Section 4.10]). Also from definition (2.9), we obtain

$$\rho^{(k)}(x_1,\dots,x_k)\le\|K\|^k\alpha^k\sum_{\pi\in S_k}(\alpha^{-1})^{\nu(\pi)},$$

and so we can take $\kappa_k=\prod_{i=0}^{k-1}(j\alpha+1)\|K\|^k$ for an α-permanental point process. The following result is a consequence of the upcoming Proposition 2.3 and the identity (2.8) for clustering constants of a permanental point process with kernel αK.

Proposition 2.1. *Let $\alpha=1/m$ for some $m\in\mathbb{N}$ and let $\mathcal{P}_\alpha$ be the stationary α-permanental point process with a kernel K which is Hermitian, non-negative definite and locally trace class. Assume also that $|K(x,y)|\le\omega(|x-y|)$ for some fast decreasing ω. Then $\mathcal{P}_\alpha$ is an admissible clustering point process with clustering function $\phi=\omega$ and clustering constants $C_k=km^{1-k(m-1)}m!(k!)^m\|K\|^{km-1}, c_k=1$.*

Zero set of Gaussian entire function (GEF). A Gaussian entire function $f(z)$ is the sum $\sum_{j\ge0}X_j\frac{z^j}{\sqrt{j!}}$ with independent standard complex Gaussian coefficients X_j, that is the X_j are i.i.d. with the normal density on the complex plane. The zero set $f^{-1}(\{0\})$ gives rise to the point process $\mathcal{P}_{GEF}:=\sum_{x\in f^{-1}(\{0\})}\delta_x$ on $\mathbb{R}^2$. The point process $\mathcal{P}_{GEF}$ is an admissible clustering point process [51, Theorem 1.4], exhibiting local repulsion of points. Though $\mathcal{P}_{GEF}$ satisfies condition (1.14), it is unclear whether (1.13) holds. Further, by [36, Theorem 1], $\mathcal{P}_{GEF}(B_r(\mathbf{0}))$ has exponential moments.

Moment conditions. For $p\in[1,\infty)$, we show that the p-moment condition (1.16) holds when ξ is such that the pair $(\xi,\mathcal{P}_{GEF})$ is of class (A1). By [51, Theorem 1.3], given $\mathcal{P}:=\mathcal{P}_{GEF}$, there exists constants $\tilde{D}_k$ such that

$$\tilde{D}_k^{-1}\prod_{i<j}\min\{|y_i-y_j|^2,1\}\le\rho^{(k)}(y_1,\dots,y_k)\le\tilde{D}_k\prod_{i<j}\min\{|y_i-y_j|^2,1\}. \tag{2.10}$$

Recall from [67, Lemma 6.4] (see also [30, Theorem 1], [11, Proposition 2.5]), that the existence of correlation functions of any point process implies existence of reduced Palm correlation functions $\rho^{(k)}_{x_1,\dots,x_p}(y_1,\dots,y_k)$, which satisfy the following useful relation: For

Lebesgue a.e. $(x_1, \ldots, x_p)$ and $(y_1, \ldots, y_k)$, all distinct,

$$\rho^{(p)}(x_1, \ldots, x_p)\rho^{(k)}_{x_1,\ldots,x_p}(y_1, \ldots, y_k) = \rho^{(p+k)}(x_1, \ldots, x_p, y_1, \ldots, y_k). \tag{2.11}$$

Combining (2.10) and (2.11), we get for Lebesgue a.e. $(x_1, \ldots, x_p)$ and $(y_1, \ldots, y_k)$, that

$$\rho^{(k)}_{x_1,\ldots,x_p}(y_1, \ldots, y_k) \leq D_{p+k}\rho^{(k)}(y_1, \ldots, y_k), \tag{2.12}$$

where $D_{p+k} := \tilde{D}_{p+k}\tilde{D}_p\tilde{D}_k$. Thus we have shown there exists constants D_j's such that for any bounded Borel subset B, $k \in \mathbb{N}$ and Lebesgue a.e. $(x_1, \ldots, x_p) \in \mathbb{R}^{dp}$, we have

$$\mathbb{E}^!_{x_1,\ldots,x_p}(\mathcal{P}^{(k)}(B^k)) \leq D_{p+k}\mathbb{E}(\mathcal{P}^{(k)}(B^k)). \tag{2.13}$$

By (2.5), (2.13), and (2.6) in this order, along with stationarity of $\mathcal{P}_{GEF}$, we have for any $p \in [1, \infty)$,

$$\begin{aligned} &\sup_{1\leq n\leq\infty} \sup_{1\leq p'\leq\lfloor p\rfloor} \sup_{x_1,\ldots,x_{p'}\in W_n} \mathbb{E}_{x_1,\ldots,x_{p'}} \max\{|\xi(x_1, \mathcal{P}_n)|, 1\}^p \\ \leq \;& \left(\frac{\|h\|_\infty}{k}\right)^p \kappa_{(k-1)p} D_{kp}\mathbb{E}[(\mathrm{Po}(\mathrm{Vol}(B_r(\mathbf{0}))) + p)^{(k-1)p}] < \infty, \end{aligned} \tag{2.14}$$

where as before $\mathrm{Po}(\lambda)$ denotes a Poisson random variable with mean λ. Thus the p-moment condition (1.16) holds for pairs $(\xi, \mathcal{P}_{GEF})$ of class (A1) for all $p \in [1, \infty)$.

2.2.2 Class A2 input

Determinantal input. The point process $\mathcal{P}$ is determinantal if its correlation functions are defined by $\rho^{(k)}(x_1, \ldots, x_k) = \det(K(x_i, x_j))_{1\leq i,j\leq k}$, where $K(\cdot, \cdot)$ is again the Hermitian kernel of a locally trace class integral operator $\mathcal{K} : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d)$. Determinantal point processes exhibit local repulsivity. Also, determinantal structure is preserved when restricting determinantal input to subsets of $\mathbb{R}^d$ and as well as when considering their reduced Palm versions. These facts facilitate our analysis of determinantal input; the Appendix (Section 5) provides lemmas further illustrating the tractability of determinantal point processes. If a stationary determinantal point process has a fast-decreasing kernel as at (2.7), then Lemma 5.3 in Section 5 shows that it is an admissible clustering point process satisfying (1.13) with clustering function $\phi = \omega$ and clustering constants

$$C_k := k^{1+(k/2)}||K||^{k-1}, c_k \equiv 1. \tag{2.15}$$

Consequently, it satisfies the inequality (1.14) provided ω itself satisfies (1.14). Also, by Hadamard's inequality we can take $\kappa_k = K(\mathbf{0}, \mathbf{0})^k$.

The Ginibre ensemble of eigenvalues of $N \times N$ matrices with independent standard complex Gaussian entries is a leading example of a determinantal point process. The limit of the Ginibre ensemble as $N \to \infty$ is the Ginibre point process (or the infinite

Ginibre ensemble), here denoted $\mathcal{P}_{GIN}$. It is the prototype of a stationary determinantal point process and has kernel

$$K(z_1, z_2) := \exp(z_1\bar{z}_2)\exp(-(|z_1|^2+|z_2|^2)/2) = \exp(i\mathrm{Im}(z_1\bar{z}_2)-|z_1-z_2|^2/2), \ \ z_1, \ z_2 \in \mathbb{C}.$$

More generally, for $0 < \beta \leq 1$, the β-Ginibre (determinantal) point process (see [28]) has kernel

$$K_\beta(z_1, z_2) := \exp(\frac{1}{\beta}z_1\bar{z}_2)\exp(-(|z_1|^2+|z_2|^2)/(2\beta)), \quad z_1, \ z_2 \in \mathbb{C}.$$

When $\beta = 1$, we obtain $\mathcal{P}_{GIN}$ and as $\beta \to 0$ we obtain the Poisson point process. Thus the β-Ginibre point process interpolates between the Poisson and Ginibre point processes. All β-Ginibre point processes are admissible clustering point processes satisfying (1.13) and (1.14).

Moment Conditions. Let $p \in [1, \infty)$ and let $\mathcal{P}$ be a stationary determinantal point process with a continuous and fast-decreasing kernel. We now show that the p-moment condition (1.16) holds for pairs $(\xi, \mathcal{P})$ of class (A1) or (A2), provided ξ has a deterministic radius of stabilization, say $r_0 \in [1, \infty)$. First, for all $(x_1, \ldots, x_p) \in (\mathbb{R}^d)^p$, all increasing $F : \mathbb{N} \to \mathbb{R}^+$ and all bounded Borel sets B we have [28, Theorem 2]

$$\mathbb{E}^!_{x_1,\ldots,x_p}(F(\mathcal{P}(B))) \leq \mathbb{E}(F(\mathcal{P}(B))).$$

Thus using (2.4), the above inequality and stationarity of $\mathcal{P}$, we get that for any bounded stabilizing score function ξ of class (A2),

$$\sup_{1\leq n\leq\infty} \sup_{1\leq p'\leq\lfloor p\rfloor} \sup_{x_1,\ldots,x_{p'}\in W_n} \mathbb{E}_{x_1,\ldots,x_{p'}} \max\{|\xi(x_1, \mathcal{P}_n)|, 1\}^p \leq \mathbb{E}(\max\{\hat{c}r_0, 1\}^{p\mathcal{P}(B_{r_0}(\mathbf{0}))+p^2}) < \infty. \tag{2.16}$$

The finiteness of the last term follows from the fact that determinantal input considered here is of class (A2) and, by Remark (i) at the beginning of Section 2.1 such input has finite exponential moments.

α-Determinantal point processes. Similar to permanental point processes, we generalize determinantal point processes to include the α-determinantal point processes, by requiring that the correlation functions satisfy (2.9) for some $\alpha \leq 0$. In what follows, we shall assume that $\alpha = -1/m, m \in \mathbb{N}$. Existence of such α-determinantal point processes again follows from [67, Theorem 1.2]. Likewise, an α-determinantal point process with kernel K is a superposition of $-1/\alpha$ i.i.d. copies of a determinantal point process with kernel $-\alpha K$ (see [9, Section 4.10]). By [67, Proposition 4.3], we can take $\kappa_k = K(\mathbf{0}, \mathbf{0})^k$ for an α-determinantal point process. Analogously to Proposition 2.1, the next result follows from Proposition 2.3 below and the identity (2.15) for clustering constants of a determinantal point process with kernel $-\alpha K$.

Proposition 2.2. *Let $\alpha = -1/m$ for some $m \in \mathbb{N}$ and $\mathcal{P}_\alpha$ be the stationary α-determinantal point process with a kernel K which is Hermitian, non-negative definite and locally trace class. Assume also that $|K(x,y)| \le \omega(|x-y|)$ for some fast decreasing function ω. Then $\mathcal{P}_\alpha$ is an admissible clustering point process with clustering function $\phi = \omega$ and clustering constants $C_k = m^{1-k(m-1)} m! K(\mathbf{0},\mathbf{0})^{k(m-1)} k^{1+(k/2)} \|K\|^{k-1}, c_k = 1$. Further if w satisfies (1.14), then $\mathcal{P}_\alpha$ is admissible clustering input of type (A2).*

Rarified Gibbsian input. Consider the class $\mathbf{\Psi}$ of Hamiltonians consisting of pair potentials without negative part, area interaction Hamiltonians, hard core Hamiltonians, and potentials generating a truncated Poisson point process (see [65] for further details of such potentials). For $\Psi \in \mathbf{\Psi}$ and $\beta \in (0,\infty)$, let $\mathcal{P}^{\beta\Psi}$ be the Gibbs point process having Radon-Nikodym derivative $\exp(-\beta\Psi(\cdot))$ with respect to a reference homogeneous Poisson point process on $\mathbb{R}^d$ of intensity $\tau \in (0,\infty)$. There is a range of inverse temperature and activity parameters (β and τ) such that $\mathcal{P}^{\beta\Psi}$ strongly clusters; see the introduction to Section 3 and [65] for further details. These rarified Gibbsian point processes are admissible clustering point processes satisfying the input conditions (1.13) and (1.14) of class (A2). Setting $\xi(.,.) \equiv 1$ in Lemma 3.4 of [65] shows that (1.7) holds with C_k a scalar multiple of k and c_k a constant.

2.2.3 Additional input examples

Here we provide a non-exhaustive list of examples of admissible clustering input.

Cox point processes. A point process $\mathcal{P}$ is said to be a Cox point process with (random) *intensity measure* $\Lambda(.)$ if conditioned on $\Lambda(.)$, $\mathcal{P}$ is a Poisson point process with intensity measure $\Lambda(.)$. We shall assume that the measure $\Lambda(.)$ has a (random) density $\lambda(.)$ with respect to the Lebesgue measure, called the *intensity field* i.e., $\Lambda(B) = \int_B \lambda(x)\mathrm{d}x$ for all Borel sets B. In such a case, the correlation functions of the Cox point process $\mathcal{P}$ are given by $\rho^{(k)}(x_1,\ldots,x_k) = \mathbb{E}(\prod_{i=1}^k \lambda(x_i))$ for $x_1,\ldots,x_k$ distinct. Hence, clustering properties of the random field $\{\lambda(x)\}_{x\in\mathbb{R}^d}$ translate in a straightforward manner to clustering properties of the point process $\mathcal{P}$. Also, if $\lambda(.)$ is a stationary random field, then $\mathcal{P}$ is a stationary point process. We have already seen one class of clustering Cox point processes in the permanental point process and we shall see below another class in thinned Poisson point processes. Another tractable class of Cox point processes, called the *shot-noise Cox point process* and studied in [48], includes examples of admissible clustering point processes.

Finite-range dependent point process. Correlation functions of these processes, when assumed locally finite, (trivially) cluster with the clustering function $\phi(s) = 0$ for all $s \in (r_0,\infty)$, for some $r_0 \in (0,\infty)$, where r_0 may be thought of as the range of dependence, and with constants $c_k = 1$. Whether clustering constants C_k satisfy condition (1.13) depends on the local properties of the correlation functions. [6] Examples of finite range dependent point processes include perturbed lattices [13], Matérn cluster point processes or Matérn hard-core point processes with finite dependence radius.

Thinned Poisson point processes. Suppose $\mathcal{P}$ is a Poisson point process, $\xi(.,\mathcal{P}) \in \{0,1\}$, and consider the thinned point process $\tilde{\mathcal{P}} := \sum_{x\in\mathcal{P}} \xi(x,\mathcal{P})\delta_x$. If ξ stabilizes exponentially fast then Lemma 5.2 of [7] shows that $\tilde{\mathcal{P}}$ is an admissible clustering point process. $\tilde{\mathcal{P}}$ is a Cox point process with intensity field $\lambda(x) = \xi(x,\mathcal{P})$. This set-up includes finite-range dependent point processes $\mathcal{P}$ as well as Matérn cluster and Matérn hard-core point process with exponentially decaying dependence radii. Tractable procedures generating thinnings of Poisson point processes are in [2].

Thinned general point processes. Suppose $(\xi,\mathcal{P})$ is an admissible pair of class A2 and suppose further that $\xi(.,\mathcal{P}) \in \{0,1\}$. Then μ_n^ξ is a thinned point process and the correlation functions of μ_n^ξ coincide with the mixed moment functions $m_{(k)}^{(1,\dots,1)}(x_1,\dots,x_k;\infty)$ in (1.5). In view of Theorem 1.10 these functions (strongly) cluster and hence μ_n^ξ is an admissible clustering point process. For similar examples and generalizations, termed *generalized shot-noise Cox point process*, see [49].

Superpositions of i.i.d. point processes. Apart from thinning another natural operation on point processes generating new point processes consists of independent superposition. We show that this operation preserves clustering.

Let $\mathcal{P}_1,\dots,\mathcal{P}_m, m \in \mathbb{N}$, be i.i.d. copies of an admissible clustering point process $\mathcal{P}$ with correlation functions ρ. Let ρ_0 denote the correlation functions of the point process $\mathcal{P}_0 := \cup_{i=1}^m \mathcal{P}_i$. Notice that for any $k \geq 1$ and distinct $x_1,\dots,x_k \in \mathbb{R}^d$ the following relation holds

$$\rho_0^{(k)}(x_1,\dots,x_k) = \sum_{\sqcup_{i=1}^m S_i=[k]} \prod_{i=1}^m \rho(S_i), \tag{2.17}$$

where $\sqcup$ stands for disjoint union and we abbreviate $\rho^{(|S_i|)}(x_j : j \in S_i)$ by $\rho(S_i)$. Here S_i may be empty, in which case we set $\rho(\emptyset) = 1$. It follows from (2.17) that $\mathcal{P}_0$ is an admissible point process with intensity $m\rho^{(1)}(\mathbf{0})$. Further, we can take $\kappa_k(\mathcal{P}_0) = (\kappa_k)^m m^k$. The proof of the proposition below, which shows that $\mathcal{P}_0$ clusters, is in the Appendix.

Proposition 2.3. *Let $m \in \mathbb{N}$ and $\mathcal{P}_1,\dots,\mathcal{P}_m$ be i.i.d. copies of an admissible clustering point process $\mathcal{P}$ with clustering function ϕ and clustering constants C_k and c_k. Then*

[6] A point process consisting of sufficiently heavy-tailed random number of points distributed independently (and say uniformly) in each hard-ball of a hard-core (say Matérn) model will not satisfy (1.13).

the point process $\mathcal{P}_0 := \cup_{i=1}^m \mathcal{P}_i$ is an admissible clustering point process with clustering function ϕ and clustering constants $m^k m!(\kappa_k)^{m-1} C_k$ and c_k. Further, if $\mathcal{P}$ is admissible clustering input of type (A2) with $\kappa_k \leq \lambda^k$ for some $\lambda \in (0, \infty)$, then $\mathcal{P}_0$ is also admissible clustering input of type (A2).

We have already used this proposition in the context of clustering of α-determinantal point processes.

2.3 Applications

Having provided examples of admissible point processes, we shall now establish the limit theory for geometric and topological statistics of these point processes. We rely heavily on Theorems 1.11 and 1.13. Our examples include statistics arising in (i) combinatorial topology, (ii) differential topology, (iii) integral geometry, and (iv) computational geometry, respectively. When the underlying point process is a Gibbs point process as described in Section 2.2, then these results can be deduced from [7] or [65] respectively. The examples are not exhaustive and indeed include many other functionals in stochastic geometry already discussed in e.g. [7, 61]. Indeed there are further applications to (i) random packing models on clustering input (extending [59]), (ii) statistics of percolation models (extending e.g. [39, 58]), and (iii) statistics of extreme points of clustering input (extending [4, 69]). Details are left to the reader. We shall need to assume moment bounds and variance lower bounds in our applications. In view of (2.14) and (2.16), the moment conditions are valid for determinantal point processes and zeros of Gaussian entire functions in the applications considered in Sections 2.3.1-2.3.3.

2.3.1 Statistics of simplicial complexes

A nonempty family Δ of finite subsets of a set V is *an abstract simplicial* complex if $\mathcal{Y} \in \Delta$ and $\mathcal{Y}_0 \subset \mathcal{Y}$ implies that $\mathcal{Y}_0 \in \Delta$. Elements of Δ are called *faces/simplices* and the *dimension of a face* is one less than its cardinality. The 0-dimensional faces are *vertices*. The collection of all faces of Δ with dimension less than k is a sub-complex called the *k-skeleton* of Δ and denoted by $\Delta^{\leq k}$. The 1-skeleton of a simplicial complex is a graph whose vertices are 0-dimensional faces and whose edges are 1-dimensional faces. The simplicial complex, or 'complex' for short, represents a combinatorial generalization of a graph, as seen in some of the examples below and is a fundamental object in combinatorial as well as computational topology [20, 50].

Given a finite point set $\mathcal{X}$ in $\mathbb{R}^d$ (or generally, in a metric space) there are various ways to define a complex that captures some of the geometry/topology of $\mathcal{X}$. One such complex is the Čech complex. Recall that if $\mathcal{X} = \{x_i\}_{i=1}^n \subset \mathbb{R}^d$ is a finite set of points and $r \in (0, \infty)$, then the Čech complex of radius r is the abstract complex

$$\mathcal{C}(\mathcal{X}, r) := \{\sigma \subset \mathcal{X} : \bigcap_{x \in \sigma} B_r(x) \neq \emptyset\}.$$

By the nerve theorem [10, Theorem 10.7], the Čech complex is homotopy equivalent (in particular, same topological invariants) to the classical germ-grain model

$$\mathcal{C}_B(\mathcal{X}, r) := \bigcup_{x \in \mathcal{X}} B_x(r), \tag{2.18}$$

The 1-skeleton of the Čech complex, $\mathcal{C}(\mathcal{X}, r)^{\leq 1}$, $\mathcal{X}$ random, is the well-known *random geometric graph* [55], denoted by $G(\mathcal{X}, r)$. One can study many geometric or topological statistics similar to those described below for the Čech complex for other geometric complexes (for example, see [73, Section 3.2]) or geometric graphs. Indeed, motivated by problems in topological data analysis, random geometric complexes on Poisson or binomial point processes were studied in [35] and later were extended to stationary point processes in [70].

We next establish the limit theory for statistics of random Čech complexes. *The central limit theorems are applicable whenever the input $\mathcal{P}$ is either α-determinantal ($|\alpha| = \frac{1}{m}, m \in \mathbb{N}$) with kernel as at (2.7), $\mathcal{P}_{GEF}$, or rarified Gibbsian input.* In all that follows we fix $r \in (0, \infty)$.

Simplex counts or clique counts. Let Γ be a complex on k-vertices such that $\Gamma^{\leq 1}$ is a connected graph. For $x \in \mathbb{R}^d$ and $\mathbf{x} := (x_1, \dots, x_{k-1}) \in (\mathbb{R}^d)^{k-1}$, let

$$h^\Gamma(x, \mathbf{x}) := \mathbf{1}[\mathcal{C}(\{x, x_1, \dots, x_{k-1}\}, r) \cong \Gamma],$$

where $\cong$ stands for simplicial isomorphism. For an admissible point process $\mathcal{P}$ as in Definition 1.2, we put

$$\gamma^{(k)}(x, \mathcal{P}) := \frac{1}{k!} \sum_{\mathbf{x} \in (\mathcal{P} \cap B_r(x))^{k-1}} h^\Gamma(x, \mathbf{x}),$$

that is $(\gamma^{(k)}, \mathcal{P})$ is an admissible pair of type (A1). If Γ denotes the $(k-1)$-simplex, then $H_n^{\gamma^{(k)}}(\mathcal{P})$ is the number of $(k-1)$-simplices in $\mathcal{C}(\mathcal{P}_n, r)$ and for $k = 2$, $H_n^{\gamma^{(k)}}(\mathcal{P})$ is the *edge count* in the random geometric graph $G(\mathcal{P}_n, r)$. Theorem 3.4 of [70] establishes expectation asymptotics for $n^{-1} H_n^{\gamma^{(k)}}(\mathcal{P})$ for stationary input. The next result establishes variance asymptotics and asymptotic normality of $n^{-1} H_n^{\gamma^{(k)}}(\mathcal{P})$. It is an immediate consequence of Theorem 1.11(ii) and Theorem 1.13. Let $\sigma^2(\gamma^{(k)})$ be as at (1.18), with ξ put to be $\gamma^{(k)}$.

Theorem 2.4. *Let $k \in \mathbb{N}$. If $\mathcal{P}$ is an admissible clustering point process as in Definition 1.2 and the pair $(\gamma^{(k)}, \mathcal{P})$ satisfies the moment condition (1.16) for all $p \in (1, \infty)$, then*

$$\lim_{n\to\infty} n^{-1} \mathrm{Var} H_n^{\gamma^{(k)}}(\mathcal{P}) = \sigma^2(\gamma^{(k)}).$$

Additionally, if $\mathrm{Var} H_n^{\gamma^{(k)}}(\mathcal{P}) = \Omega(n^\nu)$ *for some $\nu \in (0, \infty)$, then as $n \to \infty$*

$$\frac{H_n^{\gamma^{(k)}}(\mathcal{P}) - \mathbb{E} H_n^{\gamma^{(k)}}(\mathcal{P})}{\sqrt{\mathrm{Var} H_n^{\gamma^{(k)}}(\mathcal{P})}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.19}$$

Up to now, the central limit theorem theory for clique counts has been restricted to binomial or Poisson input, cf. [19, 22, 37, 55, 62]. Theorem 2.4 shows that asymptotic normality holds for more general input.

Edge lengths. For $x, y \in \mathbb{R}^d$, let

$$h(x, y) := |x - y| \mathbf{1}[|x - y| \le r].$$

The U-statistic

$$\xi^L(x, \mathcal{X}) := \frac{1}{2} \sum_{y \in \mathcal{X} \cap B_r(x)} h(x, y),$$

is of generic type (1.10) and $H_n^{\xi^L}(\mathcal{P})$ is the *total edge length* of the geometric graph $G(\mathcal{P}_n, r)$. The following is an immediate consequence of Theorems 1.11 and 1.13.

Theorem 2.5. *For any admissible clustering point process $\mathcal{P}$ as in Definition 1.2 with the pair $(\xi^L, \mathcal{P})$ satisfying the moment condition (1.16) for all $p \in (1, \infty)$, we have*

$$|n^{-1} \mathbb{E} H_n^{\xi^L}(\mathcal{P}) - \mathbb{E}_{\mathbf{0}} \xi^L(\mathbf{0}, \mathcal{P}) \rho^{(1)}(\mathbf{0})| = O(n^{-1/d}),$$

and

$$\lim_{n\to\infty} n^{-1} \mathrm{Var} H_n^{\xi^L}(\mathcal{P}) = \sigma^2(\xi^L).$$

Moreover, if $\mathrm{Var} H_n^{\xi^L}(\mathcal{P}) = \Omega(n^\nu)$ *for some $\nu \in (0, \infty)$ then as $n \to \infty$*

$$\frac{H_n^{\xi^L}(\mathcal{P}) - \mathbb{E} H_n^{\xi^L}(\mathcal{P}))}{\sqrt{\mathrm{Var} H_n^{\xi^L}(\mathcal{P})}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.20}$$

The central limit theory for $H_n^{\xi^L}(\cdot)$ for Poisson or binomial input is a consequence of [19, 22, 37, 55, 62]. Theorem 2.5 shows that $H_n^{\xi^L}(\cdot)$ still satisfies a central limit theorem when Poisson and binomial input is replaced by more general clustering input.

Degree counts. Define the (down) degree of a k-simplex to be the number of k-simplices with which the given simplex has a common $(k-1)$-simplex. For $x \in \mathbb{R}^d$

and $\mathbf{x} := (x_1, \ldots, x_{k+1}) \in (\mathbb{R}^d)^{k+1}$, define the indicator that $(x_1, \ldots, x_k)$ is the common $(k-1)$ simplex between two k-simplices:

$$h(x, \mathbf{x}) := \mathbf{1}[C(x, \ldots, x_k) \text{ is a k} - \text{simplex}]\ \mathbf{1}[C(x_1, \ldots, x_{k+1}) \text{ is a k} - \text{simplex}].$$

The total (down) degree of order k of a complex is the sum of the degrees of the constituent k-simplices. Consider the U-statistic

$$\xi^{(k)}(x, \mathcal{X}) := \frac{1}{(k+2)!} \sum_{\mathbf{x} \in (\mathcal{X} \cap B_r(x))^{k+1}} h(x, \mathbf{x})$$

which is of generic type (1.10). Then $H_n^{\xi^{(k)}}(\mathcal{P})$ is the total down degree (of order k) of the geometric complex $\mathcal{C}(\mathcal{P}_n, r)$. Note that $(\xi^{(k)}, \mathcal{P})$ is of type (A1) whenever $\mathcal{P}$ is admissible in the sense of Definition 1.2. Theorems 1.11 and 1.13 yield the following limit theory for $H_n^{\xi^{(k)}}(\mathcal{P})$.

Theorem 2.6. *Let $k \in \mathbb{N}$. For any admissible clustering point process $\mathcal{P}$ as in Definition 1.2 with the pair $(\xi^{(k)}, \mathcal{P})$ satisfying the moment condition (1.16) for all $p \in (1, \infty)$, we have*

$$|n^{-1}\mathbb{E}H_n^{\xi^{(k)}}(\mathcal{P}) - \mathbb{E}_{\mathbf{0}}\xi^{(k)}(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0})| = O(n^{-1/d}),$$

and

$$\lim_{n \to \infty} n^{-1}\mathrm{Var}H_n^{\xi^{(k)}}(\mathcal{P}) = \sigma^2(\xi^{(k)}).$$

Moreover if $\mathrm{Var}H_n^{\xi^{(k)}}(\mathcal{P}) = \Omega(n^\nu)$ for some $\nu \in (0, \infty)$ then as $n \to \infty$

$$\frac{H_n^{\xi^{(k)}}(\mathcal{P}) - \mathbb{E}H_n^{\xi^{(k)}}(\mathcal{P})}{\sqrt{\mathrm{Var}H_n^{\xi^{(k)}}(\mathcal{P})}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.21}$$

2.3.2 Morse critical points

Understanding the topology of a manifold via smooth functions on the manifold is a classical topic in differential topology known as Morse theory [45]. Among the various extensions of Morse theory to non-smooth functions, the one of interest to us is the 'min-type' Morse theory developed in [26]. This theory was exploited to study the topology of random Čech complexes on Poisson and binomial point processes by [14] and later on stationary point processes by [70].

As above $\mathcal{X} \subset \mathbb{R}^d$ denotes a locally finite point set and $k \in \mathbb{N}$. Given $\mathbf{z} \in \mathbb{R}^{d(k+1)}$, let $C(\mathbf{z})$ denote the center of the unique $k-1$ dimensional sphere (if it exists) containing the points of $\mathbf{z}$ and let $R(\mathbf{z})$ be the radius of this unique ball. The set of points $\mathbf{z}$ of cardinality $k+1$ in general position generates *an index k critical point* iff

$$C(\mathbf{z}) \in \mathring{\mathrm{co}}(\mathbf{z}) \ \text{ and } \ \mathcal{X}(B_{R(\mathbf{z})}(C(\mathbf{z}))) = z,$$

where $\text{co}(\mathbf{z})$ is the convex hull of the points comprising $\mathbf{z}$ and $\mathring{A}$ stands for the interior of a Euclidean set A. We are interested in critical points $C(\mathbf{z})$ distant at most r from $\mathcal{X}$ i.e, $R(\mathbf{z}) \in (0, r]$. To this end, for $(x, \mathbf{x}) \in \mathbb{R}^{d(k+1)}$ define

$$g_r(x, \mathbf{x}) := \mathbf{1}[C(x, \mathbf{x}) \in \mathring{\text{co}}(x, \mathbf{x})]\mathbf{1}[R(x, \mathbf{x}) \leq r]\,; \; Q(x, \mathbf{x}) := B_{R(\mathbf{x})}(C(\mathbf{x}))\mathbf{1}[R(x, \mathbf{x}) \leq r].$$

Thus $Q(x, \mathbf{x}) \subset B_{2r}(x)$. Now, for $\mathbf{x} \in \mathbb{R}^{dk}$, we set $h(x, \mathbf{x}) := (k+1)^{-1} g_r(x, \mathbf{x})$, and, in keeping with (1.10), define the Morse score function

$$\xi^M(x, \mathcal{X}) := \frac{1}{k!} \sum_{\mathbf{x} \in (\mathcal{X} \cap B_{2r}(x))^k} h(x, \mathbf{x}) \mathbf{1}[\mathcal{X}(Q(x, \mathbf{x}) \setminus \{x, \mathbf{x}\}) = 0]. \tag{2.22}$$

Then ξ^M satisfies the power growth condition (1.15) and is of type (A2) (and nearly of type (A1)), with the understanding that $h(x, x_1, \ldots, x_k) = 0$ whenever $\max_{1 \leq i \leq k} |x_i - x| > 2r$. The statistic $H^{\xi^M}(\mathcal{P})$ is simply the number $N_k(\mathcal{P}_n, r)$ of index k Morse critical points generated by $\mathcal{P}_n$ which are within a distance r of $\mathcal{P}_n$, whereas $\mu_n^{\xi^M}$ is the random measure generated by index k critical points. Index 0 Morse critical points are trivially the points of $\mathcal{X}$ and so in this case $N_0(\mathcal{X}) = \text{card}(\mathcal{X})$. Thus we shall be interested in asymptotics for only Morse critical points of higher indices.

The next result establishes variance asymptotics and asymptotic normality of $n^{-1} N_k(\mathcal{P}_n, r)$ valid for class (A2) input $\mathcal{P}$. It is an immediate consequence of Theorem 1.11(ii), Theorem 1.13, and the fact that for input $\mathcal{P}$ of class (A2), $(\xi^M, \mathcal{P})$ is an input pair of class (A2). Let $\sigma^2(\xi^M)$ be as at (1.18), with ξ put to be ξ^M.

Theorem 2.7. *For all $k \in \{1, \ldots, d\}$ and class (A2) input $\mathcal{P}$ with the pair $(\xi^M, \mathcal{P})$ satisfying the moment condition (1.16) for all $p \in (1, \infty)$, we have*

$$\lim_{n \to \infty} n^{-1} \text{Var} N_k(\mathcal{P}_n, r) = \sigma^2(\xi^M).$$

Moreover if $\text{Var} N_k(\mathcal{P}_n, r) = \Omega(n^\nu)$ for some $\nu \in (0, \infty)$, then as $n \to \infty$

$$\frac{N_k(\mathcal{P}_n, r) - \mathbb{E} N_k(\mathcal{P}_n, r)}{\sqrt{\text{Var} N_k(\mathcal{P}_n, r)}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.23}$$

Remarks:

(i) Theorem 5.2 of [70] establishes expectation asymptotics for $n^{-1} N_k(\mathcal{P}_n, r)$ for stationary input, though without a rate of convergence and [14] estabilishes a central limit theorem but only for the case of Poisson or binomial point processes.

(ii) The Morse inequalities relate the Morse critical points (local functionals) to the Betti numbers (global functionals) of the Boolean model and in particular, imply that the changes in the homology of the Boolean model $\mathcal{C}_B(\mathcal{P}, r)$ occurs at radii $r = R(\mathbf{x})$ whenever $C(\mathbf{x})$ is a Morse critical point. A trivial consequence is that the kth Betti number of $\mathcal{C}(\mathcal{P}_n, r)$ is upper bounded by $N_k(\mathcal{P}_n, r)$.

(iii) Other examples of similar score functions satisfying a modified version of (A1) (similar to (2.22)) include component counts of random geometric graphs [55, Chapter 3], number of simplices of degree k in the Čech complex, simplicial counts in an alpha complex [73, Sec 3.2] or an appropriate discrete Morse complex on $\mathcal{P}_n$ (see [23]).

2.3.3 Statistics of germ-grain models

We furnish two more applications of Theorem 1.13 when ξ has a deterministic radius of stabilization. The two applications concern the germ-grain model, a classic model in stochastic geometry [64].

k-covered region of the germ-grain model. The following is a statistic of interest in coverage processes [29]. For locally-finite $\mathcal{X} \subset \mathbb{R}^d$ and $x \in \mathcal{X}$, define the score function

$$\beta^{(k)}(x, \mathcal{X}) := \int_{y \in B_r(x)} \frac{\mathbf{1}[\mathcal{X}(B_r(y)) \geq k]}{\mathcal{X}(B_r(y))} \mathrm{d}y.$$

Clearly, $\beta^{(k)}$ is an exponentially stabilizing score function as in Definition 1.1 with stabilization radius $2r$. Define the k-covered region of the germ-grain model $\mathcal{C}_B(\mathcal{P}_n, r)$ at (2.18) by $\mathcal{C}_B^k(\mathcal{P}_n, r) = \{y : \mathcal{P}_n(B_r(y)) \geq k\}$. Thus $H_n^{\beta^{(k)}}(\mathcal{P})$ is the volume of $\mathcal{C}_B^k(\mathcal{P}_n, r)$. When $k = 1$, $H_n^{\beta^{(k)}}(\mathcal{P})$ is the volume of the germ-grain model having germs in $\mathcal{P}_n$. Clearly $\beta^{(k)}$ is bounded by the volume of a radius r ball and so ξ satisfies the power growth condition (1.15). The following is an immediate consequence of Theorems 1.11 and 1.13 and the fact that if $\mathcal{P}$ is of class (A2) then the input pair $(\beta^{(k)}, \mathcal{P})$ is also of class (A2).

Theorem 2.8. *For all $k \in \mathbb{N}$ and any point process $\mathcal{P}$ of class (A2) with the pair $(\beta^{(k)}, \mathcal{P})$ satisfying the moment condition (1.16) for all $p \in (1, \infty)$, we have*

$$|n^{-1}\mathbb{E}\mathrm{Vol}(\mathcal{C}_B^k(\mathcal{P}_n, r)) - \mathbb{E}_{\mathbf{0}}\beta^{(k)}(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0})| = O(n^{-1/d}),$$

and

$$\lim_{n \to \infty} n^{-1}\mathrm{Var}\mathrm{Vol}(\mathcal{C}_B^k(\mathcal{P}_n, r)) = \sigma^2(\beta^{(k)}).$$

Moreover, if $\mathrm{Var}\mathrm{Vol}(\mathcal{C}_B^k(\mathcal{P}_n, r)) = \Omega(n^\nu)$ *for some $\nu \in (0, \infty)$, then as $n \to \infty$*

$$\frac{\mathrm{Vol}(\mathcal{C}_B^k(\mathcal{P}_n, r)) - \mathbb{E}\mathrm{Vol}(\mathcal{C}_B^k(\mathcal{P}_n, r))}{\sqrt{\mathrm{Var}\mathrm{Vol}(\mathcal{C}_B^k(\mathcal{P}_n, r))}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.24}$$

The central limit theorem (2.24) is valid for input which is either a α-determinantal point process ($\alpha = \frac{-1}{m}, m \in \mathbb{N}$) with a fast decreasing kernel or a rarified Gibbsian point process. In the case of Poisson input and $k = 1$, [29] establishes a central limit theorem for $\mathcal{C}_B^1(\mathcal{P}_n, r)$. For general k, the central limit theorem can be deduced from the general results in [58, 7] with presumably optimal bounds following from [40, Proposition 1.4].

Intrinsic volumes of the germ-grain model. Let $\mathcal{K}$ denote the set of *convex bodies* i.e., compact, convex subsets of $\mathbb{R}^d$. The *intrinsic volumes* $V_0, \ldots, V_d$ are non-negative functionals on $\mathcal{K}$ which satisfy *Steiner's formula*

$$V_d(K \oplus B_r(\mathbf{0})) = \sum_{j=0}^{d} r^{d-j}\theta_{d-j}V_j(K), \ K \in \mathcal{K}$$

where $r > 0$, $\oplus$ is the Minkowski sum, V_d denotes the d-dimensional Lebesgue measure, and $\theta_j := \pi^{j/2}/\Gamma(j/2+1)$ is the volume of the unit ball in $\mathbb{R}^j$. Intrinsic volumes satisfy translation invariance and additivity i.e., for $K_1, \ldots, K_m \in \mathcal{K}$,

$$V_j(\cup_{i=1}^m K_i) = \sum_{k=1}^{m}(-1)^{k+1} \sum_{1 \le i_1 < \ldots < i_k \le m} V_j(\cap_{l=1}^k K_{i_l}), \quad j \in \{0, \ldots, d\}.$$

This identity allows an extension of intrinsic volumes to real-valued functionals on the family of finite unions of convex bodies (see [64, Ch. 14]). Intrinsic volumes coincide with *quer-mass integrals* or *Minkowski functional* up to a normalization. The V_j's define certain j-dimensional volumes of K, independently of the ambient space. V_d is the d-dimensional volume, $2V_{d-1}$ is the surface measure, and V_0 is the *Euler-Poincaré characteristic* which may also be expressed as an alternating sum of simplex counts, Morse critical points, or Betti numbers [20, Sections IV.2, VI.2]. Save for V_d and V_{d-1}, the remaining V_j's may assume negative values on unions of convex bodies.

For finite $\mathcal{X}$, $m = \mathrm{card}(\mathcal{X})$, and $r > 0$, we express $V_j(\mathcal{C}_B(\mathcal{X}, r)), j \in \{0, \ldots, d\}$ as a sum of bounded stabilizing scores, which goes as follows. For $x_1 \in \mathcal{X}$, define the score

$$\xi_j(x_1, \mathcal{X}) := \sum_{k=1}^{m}(-1)^{k+1} \sum_{\{x_2,\ldots,x_k\} \subset \mathcal{X} \cap B_{2r}(x_1)} \frac{V_j(B_r(x_1) \cap \ldots \cap B_r(x_k))}{k!}.$$

The score ξ_j is translation invariant with radius of stabilization $R^{\xi_j}(x_1, \mathcal{P}) \le \lceil 3r \rceil$. By additivity, we have $V_j(\mathcal{C}_B(\mathcal{P}_n, r)) = H_n^{\xi_j}(\mathcal{P})$ for $j \in \{0, \ldots, d\}$. The homogeneity relation $V_j(B_r(\mathbf{0})) = r^j V_j(B_1(\mathbf{0})) = r^j \binom{d}{j} \frac{\theta_d}{\theta_{d-j}}$ and the monotonicity of V_j on $\mathcal{K}$ yield

$$|\xi_j(x_1, \mathcal{X})| \mathbf{1}[\mathcal{X}(B_{2r}(x_1)) = l] \le r^j \binom{d}{j} \frac{\theta_d}{\theta_{d-j}} \sum_{k=1}^{l} \binom{l}{k} \le 2^l r^j \binom{d}{j} \frac{\theta_d}{\theta_{d-j}}.$$

In other words, $\xi_j, 0 \le j \le d$, satisfy the power growth condition (1.15) and are exponentially stabilizing. Theorems 1.11 and 1.13 yield the following limit theorems, where we note that $(\xi_j, \mathcal{P})$ is of class (A2) whenever $\mathcal{P}$ is of class (A2).

Theorem 2.9. *Fix $r \in (0, \infty)$ as above. For all $j \in \{0, \ldots, d\}$ and any point process $\mathcal{P}$ of class (A2) with the pair $(\xi_j, \mathcal{P})$ satisfying the p-moment condition (1.16) for all*

$p \in (1, \infty)$, we have

$$|n^{-1}\mathbb{E}V_j(\mathcal{C}_B(\mathcal{P}_n, r)) - \mathbb{E}_{\mathbf{0}}\xi_j(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0})| = O(n^{-1/d}),$$

and

$$\lim_{n\to\infty} n^{-1}\mathrm{Var}V_j(\mathcal{C}_B(\mathcal{P}_n, r)) = \sigma^2(\xi_j).$$

Moreover, if $\mathrm{Var}V_j(\mathcal{C}_B(\mathcal{P}_n, r)) = \Omega(n^\nu)$ for some $\nu \in (0, \infty)$ then as $n \to \infty$

$$\frac{V_j(\mathcal{C}_B(\mathcal{P}_n, r)) - \mathbb{E}V_j(\mathcal{C}_B(\mathcal{P}_n, r))}{\sqrt{\mathrm{Var}V_j(\mathcal{C}_B(\mathcal{P}_n, r))}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.25}$$

Remarks:

(i) Theorem 2.9 extends the analogous central limit theorems of [34], which are confined to Poisson input, to any point process of class (A2).

(ii) We may likewise prove central limit theorems for other functionals of germ-grain models, including mixed volumes, integrals of surface area measures [63, Chapters 4 and 5], and total measures of translative integral geometry [64, Section 6.4]. These functionals, like intrinsic volumes, are expressed as sums of bounded stabilizing scores and thus, under suitable assumptions, the limit theory for these functionals follows from Theorems 1.11 and 1.13.

2.3.4 Edge-lengths in k-nearest neighbor graphs

We now use the full force of Theorems 1.11 and 1.13, applying them to sums of score functions whose radius of stabilization has an exponentially decaying tail.

Statistics of the Voronoi tessellation as well as of graphs in computational geometry such as the k-nearest neighbor graph and sphere of influence graph may be expressed as sums of exponentially stabilizing score functionals [58] and hence via Theorems 1.11 and 1.13, we may deduce the limit theory for these statistics. To illustrate, we establish a weak law of large numbers, variance asymptotics, and a central limit theorem for the total edge-length of the k-nearest neighbor graph on a determinantal point process $\mathcal{P}$ with a fast-decreasing kernel as in (2.7). As noted in Section 2.2, such a determinantal point process is of class (A2) as in Definition 1.7.

As shown in Lemma 5.6, we may explicitly upper bound void probabilities for $\mathcal{P}$, allowing us to deduce exponential stabilization for score functions on $\mathcal{P}$. This is a recurring phenomena, and it is often the case that to show exponential stabilization of statistics, it suffices to control the Palm probability content of large Euclidean balls. This opens the way towards showing that other relevant statistics of random graphs exhibit exponential stabilization on $\mathcal{P}$. This includes intrinsic volumes of faces of Voronoi tessellations [64, Section 10.2], edge-lengths in a radial spanning tree [66, Lemma 3.2], proximity graphs including the Gabriel graph, and global Morse critical points i.e., critical points as defined in Section 2.3.2 but without the restriction $1[R(x, \mathbf{x}) \leq r]$.

Given locally finite $\mathcal{X} \subset \mathbb{R}^d$ and $k \in \mathbb{N}$, the (undirected) k-nearest neighbor graph $NG(\mathcal{X})$ is the graph with vertex set $\mathcal{X}$ obtained by including an edge $\{x, y\}$ if y is one of the k nearest neighbors of x and/or x is one of the k nearest neighbors of y. In the case of a tie we may break the tie via some pre-defined total order (say lexicographic order) on $\mathbb{R}^d$. For any finite $\mathcal{X} \subset \mathbb{R}^d$ and $x \in \mathcal{X}$, we let $\mathcal{E}(x)$ be the edges e in $NG(\mathcal{X})$ which are incident to x. Defining

$$\xi_L(x, \mathcal{X}) := \frac{1}{2} \sum_{e \in \mathcal{E}(x)} |e|,$$

we write the total edge length of $NG(\mathcal{X})$ as $L(NG(\mathcal{X})) = \sum_{x \in \mathcal{X}} \xi_L(x, \mathcal{X})$. Let $\sigma^2(\xi_L)$ be as at (1.18), with ξ put to be ξ_L.

Theorem 2.10. *Let $\mathcal{P}$ be a stationary determinantal point process on $\mathbb{R}^d$ with intensity $\lambda = K(\mathbf{0}, \mathbf{0})$ and a kernel satisfying $K(x, y) \leq \omega(|x - y|)$, with ω fast-decreasing as at (2.7). We have*

$$|n^{-1}\mathbb{E}L(NG(\mathcal{P}_n)) - \mathbb{E}_{\mathbf{0}}\xi_L(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0})| = O(n^{-1/d}),$$

whereas

$$\lim_{n \to \infty} n^{-1}\mathrm{Var}L(NG(\mathcal{P}_n)) = \sigma^2(\xi_L).$$

If $\mathrm{Var}L(NG(\mathcal{P}_n)) = \Omega(n^\nu)$ for some $\nu \in (0, \infty)$ then as $n \to \infty$

$$\frac{L(NG(\mathcal{P}_n)) - \mathbb{E}L(NG(\mathcal{P}_n))}{\sqrt{\mathrm{Var}L(NG(\mathcal{P}_n))}} \xrightarrow{\mathcal{D}} N(0, 1). \tag{2.26}$$

Remark. Theorem 2.10 extends Theorem 6.4 of [56] which is confined to Poisson input. In this context, the work [40] provides a rate of normal approximation.

Proof. We want to show that $(\xi_L, \mathcal{P})$ is an admissible score and input pair of type (A2) and then apply Theorem 1.13. Note that $\mathcal{P}$ is an admissible clustering point process satisfying (1.13) and (1.14). Thus we only need to show that ξ_L is exponentially stabilizing, that ξ_L satisfies the power growth condition (1.15), and the p-moment condition (1.16). When $d = 2$, we show exponential stabilization of ξ_L by closely following the proof of Lemma 6.1 of [60]. This goes as follows. For each $t > 0$, construct six disjoint equal triangles $T_j(t), 1 \leq j \leq 6$, such that x is a vertex of each triangle and each edge has length t. Let the random variable R be the minimum t such that $\mathcal{P}_n(T_j) \geq k + 1$ for all $1 \leq j \leq 6$. Notice that $R \in [r, \infty)$ implies that there is a ball inscribed in some $T_j(t)$ with center c_j of radius γr which does not contain $k + 1$ points. By Lemma 5.6

in the appendix, the probability of this event satisfies

$$\mathbb{P}_{x_1,\ldots,x_p}[R > r] \leq 6\mathbb{P}_{x_1,\ldots,x_p}[\mathcal{P}(B_{\gamma r}(c_1)) \leq k-1] \leq 6\mathbb{P}^!_{x_1,\ldots,x_p}[\mathcal{P}(B_{\gamma r}(c_1)) \leq k-1]$$
$$\leq 6e^{(2k+p-2)/8}e^{-\lambda\pi\gamma^2 r^2/8},$$

that is to say that R has exponentially decaying tails. As in Lemma 6.1 of [60], we find that $R^\xi(x, \mathcal{P}_n) := 4R$ is a radius of stabilization for ξ_L, showing that (1.12) holds with $c = 2$. For $d > 2$, we may extend these geometric arguments (cf. the proof of Theorem 6.4 of [56]) to define a random variable R serving as a radius of of stabilization. Mimicking the above arguments we may likewise show that R has exponentially decaying tails.

For all $r \in (0, \infty)$ and $l \in \mathbb{N}$ we notice that

$$|\xi_L(x, \mathcal{X} \cap B_r(x))|\mathbf{1}[\mathcal{X}(B_r(x)) = l] \leq r \cdot \min(l, 6) \leq (cr)^l,$$

and so (1.15) holds. Since vertices in the k-nearest neighbor graph have degree bounded by $kC(d)$ as in Lemma 8.4 of [71], and since each edge incident to x has length at most $4R$, it follows that $|\xi_L(x, \mathcal{P}_n)| \leq k \cdot C(d) \cdot 4R$. Since R has moments of all orders, $(\xi_L, \mathcal{P})$ satisfies the p-moment condition (1.16) for all $p \geq 1$. Thus ξ_L satisfies all conditions of Theorem 1.13 and we deduce Theorem 2.10 as desired. $\square$

3 Proof of strong clustering of mixed moments

We show strong clustering (1.17) via a factorial moment expansion for the expectation of functionals of point processes. Notice that (1.17) holds for any exponentially stabilizing score function ξ satisfying the p-moment condition (1.16) for all $p \in [1, \infty)$ on a Poisson point process $\mathcal{P}$. Indeed if $x, y \in \mathbb{R}^d$ and $r_1, r_2 > 0$ satisfy $r_1 + r_2 < |x - y|$ then the random variables $\xi(x, \mathcal{P})\mathbf{1}[R^\xi(x, \mathcal{P}) \leq r_1]$ and $\xi(y, \mathcal{P})\mathbf{1}[R^\xi(y, \mathcal{P}) \leq r_2]$ are independent. This yields clustering (1.17) with $k_1 = \ldots, k_{p+q} = 1$ and $\tilde{C}_n \leq c_1^n$ with c_1 a constant, as in [7, Lemma 5.2]. On the other hand, if $\mathcal{P}$ is rarified Gibbsian input and ξ is exponentially stabilizing, then [65, Lemma 4.1] shows mixed moment clustering (1.17) with $k_1 = \ldots, k_{p+q} = 1$. These methods depend on quantifying the region of spatial dependencies of Gibbsian points via exponentially decaying diameters of their ancestor clans. Such methods apparently neither extend to determinantal input nor to the zero set $\mathcal{P}_{GEF}$ of a Gaussian entire function. On the other hand, for $\mathcal{P}_{GEF}$ and for $\xi \equiv 1$, the paper [51] uses the Kac-Rice-Hammersley formula and complex analysis tools to show clustering (1.17) with $k_1 = \ldots, k_{p+q} = 1$. All three proofs are specific to the underlying point process or the specific score function ξ. Our more general and considerably different approach, includes these results as special cases.

3.1 Difference operators and factorial moment expansions

We shall now introduce some notations and collect some auxilliary results required for an application of the much-needed factorial moment expansions for general point processes from [11, 12]. Equip $\mathbb{R}^d$ with a total order $\prec$ defined using the lexicographical ordering of the polar co-ordinates. For $\mu \in \mathcal{N}$ and $x \in \mathbb{R}^d$, define the measure $\mu_{|x}(.) := \mu(. \cap \{y : y \prec x\})$. Note that since μ is a locally finite measure and the ordering is defined via polar co-ordinates $\mu_{|x}$ is a finite measure for all $x \in \mathbb{R}^d$. Let o denote the null-measure i.e., $o(B) = 0$ for all Borel subsets B of $\mathbb{R}^d$. For a measurable function $\psi : \mathcal{N} \to \mathbb{R}$, $l \in \mathbb{N} \cup \{0\}$, and $x_1, ..., x_l \in \mathbb{R}^d$, we define the factorial moment expansion (FME) kernels [11, 12] as follows. For $l = 0$, define $D^0\psi(\mu) := \psi(o)$. For $l \geq 1$,

$$D^l_{x_1,\ldots,x_l}\psi(\mu) = \sum_{i=0}^{l}(-1)^{l-i}\sum_{J\subset\binom{[l]}{i}}\psi(\mu_{|x_*}+\sum_{j\in J}\delta_{x_j}) = \sum_{J\subset[l]}(-1)^{l-|J|}\psi(\mu_{|x_*}+\sum_{j\in J}\delta_{x_j}), \quad (3.1)$$

where $\binom{[l]}{j}$ denotes the collection of all subsets of $[l] = \{1, \ldots, l\}$ with cardinality j and $x_* := \min\{x_1, \ldots, x_l\}$, with the minimum taken with respect to the order $\prec$. Note that $D^{(l)}_{x_1,\ldots,x_l}\psi(\mu)$ is a symmetric function of $x_1, \ldots, x_l$.[7]

We say that ψ is $\prec$-continuous at ∞ if for all $\mu \in \mathcal{N}$ we have

$$\lim_{x\uparrow\infty}\psi(\mu_{|x}) = \psi(\mu).$$

We first recall the FME expansion proved in [11, cf. Theorem 3.2] for dimension one and then extended to higher-dimensions in [12, cf. Theorem 3.1]. Recall that $\mathbb{E}^!_{y_1,\ldots,y_l}$ denote expectations with respect to reduced Palm probabilities.

Theorem 3.1. *Let $\mathcal{P}$ be a simple point process and $\psi : \mathcal{N} \to \mathbb{R}$ be $\prec$-continuous at ∞ and assume that for all $l \geq 1$*

$$\int_{\mathbb{R}^{dl}} \mathbb{E}^!_{y_1,\ldots,y_l}[|D^l_{y_1,\ldots,y_l}\psi(\mathcal{P})|]\rho^{(l)}(y_1, \ldots, y_l)\,\mathrm{d}y_1 \ldots \mathrm{d}y_l < \infty \quad (3.2)$$

and

$$\frac{1}{l!}\int_{\mathbb{R}^{dl}} \mathbb{E}^!_{y_1,\ldots,y_l}[D^l_{y_1,\ldots,y_l}\psi(\mathcal{P})]\rho^{(l)}(y_1, \ldots, y_l)\,\mathrm{d}y_1 \ldots \mathrm{d}y_l \to 0 \text{ as } l \to \infty. \quad (3.3)$$

Then $\mathbb{E}[\psi(\mathcal{P})]$ has the following factorial moment expansion

$$\mathbb{E}[\psi(\mathcal{P})] = \psi(o) + \sum_{l=1}^{\infty}\frac{1}{l!}\int_{\mathbb{R}^{dl}} D^l_{y_1,\ldots,y_l}\psi(o)\rho^{(l)}(y_1, \ldots, y_l)\,\mathrm{d}y_1 \ldots \mathrm{d}y_l. \quad (3.4)$$

Consider now admissible pairs $(\xi, \mathcal{P})$ of type (A1) or (A2) and $x_1, \ldots, x_p \in \mathbb{R}^d$.

[7]For $x_l \prec x_{l-1} \prec \ldots \prec x_1$ the functional $D^l_{x_1,\ldots,x_l}\psi(\mu)$ is equal to the iterated difference operator: $D^1_{x_1}\psi(\mu) = \psi(\mu_{|x_1} + \delta_{x_1}) - \psi(\mu_{|x_1})$, $D^l_{x_1,\ldots,x_l}\psi(\mu) = D^1_{x_l}(D^{l-1}_{x_1,\ldots,x_{l-1}}\psi(\mu))$.

The proof of (1.17) given in the next sub-section is based on the FME expansion for $\mathbb{E}_{x_1,\dots,x_p}[\psi(\mathcal{P}_n)]$, where $\psi(\mu)$ is the following product of the score functions

$$\psi(\mu) = \psi_{k_1,\dots,k_p}(x_1,\dots,x_p;\mu) := \prod_{i=1}^{p} \xi(x_i,\mu)^{k_i} \tag{3.5}$$

with $k_1,\dots,k_p \geq 1$. However, under $\mathbb{P}_{x_1,\dots,x_p}$ the point process $\mathcal{P}_n$ has fixed atoms at $x_1,\dots,x_p$, which complicates the form of its factorial moment measures. It is more handy to consider these points as parameters of the following modified functional

$$\psi^!(\mu) = \psi^!_{k_1,\dots,k_p}(x_1,\dots,x_p;\mu) := \prod_{i=1}^{p} \xi\Big(x_i,\mu+\textstyle\sum_{j=1}^{p}\delta_{x_j}\Big)^{k_i} \tag{3.6}$$

and to not count points $x_1,\dots,x_p$ in $\mathcal{P}$, i.e., consider $\mathcal{P}$ under the reduced Palm probabilities $\mathbb{P}^!_{x_1,\dots,x_p}$. Obviously $\mathbb{E}_{x_1,\dots,x_p}[\psi(\mathcal{P}_n)] = \mathbb{E}^!_{x_1,\dots,x_p}[\psi^!(\mathcal{P}_n)]$ and the latter expectation is more suitable for FME expansion with respect to the correlation functions $\rho^{(l)}_{x_1,\dots,x_p}(y_1,\dots,y_l)$ of $\mathcal{P}$ with respect to the Palm probabilities $\mathbb{P}^!_{x_1,\dots,x_p}$.

The following crucial consequence of Theorem 3.1 will allow us to use the FME expansion to prove (1.17).

Lemma 3.2. *Assume that either (i) $(\xi,\mathcal{P})$ is an admissible score and input pair of type (A1) or (ii) $(\xi,\mathcal{P})$ satisfies the power growth condition (1.15), with ξ having a radius of stabilization satisfying $\sup_{x\in\mathcal{P}} R^{\xi}(x,\mathcal{P}) \leq r$ a.s. for some $r \in (1,\infty)$ and $\mathcal{P}$ has exponential moments. Then for distinct $x_1,\dots,x_p \in \mathbb{R}^d$, non-negative integers $k_1,\dots,k_p$ and $n \leq \infty$ the functional ψ' at (3.6) admits the FME*

$$\mathbb{E}_{x_1,\dots,x_p}[\psi_{k_1,\dots,k_p}(x_1,\dots,x_p;\mathcal{P}_n)] = \mathbb{E}^!_{x_1,\dots,x_p}[\psi^!_{k_1,\dots,k_p}(x_1,\dots,x_p;\mathcal{P}_n)]$$

$$= \psi^!_{k_1,\dots,k_p}(x_1,\dots,x_p;o)$$

$$+ \sum_{l=1}^{\infty} \frac{1}{l!} \int_{\mathbb{R}^{dl}} D^l_{y_1,\dots,y_l} \psi^!_{k_1,\dots,k_p}(x_1,\dots,x_p;o)\rho^{(l)}_{x_1,\dots,x_p}(y_1,\dots,y_l)\,\mathrm{d}y_1\dots\mathrm{d}y_l, \tag{3.7}$$

When $(\xi,\mathcal{P})$ is of type (A1), the series (3.7) has at most $(k-1)\sum_{i=1}^{p} k_i$ non-zero terms.

Proof. Throughout the proof we fix non-negative integers $k_1,\dots,k_p$ and suppress them when writing $\psi^!$; i.e., $\psi^!(x_1,\dots,x_p;\mathcal{P}_n) := \psi^!_{k_1,\dots,k_p}(x_1,\dots,x_p;\mathcal{P}_n)$. We put $K_p := \sum_{i=1}^{p} k_i$. The bounded radius of stabilization for ξ implies $\psi^!$ is $\prec$-continuous at ∞.

Consider first $\psi^!$ at (3.6) with ξ as in case (ii); later we consider the simpler case (i). We show the validity of the expansion (3.7) as follows. For $y_1,\dots,y_l \in \mathbb{R}^d$ and $y_k \notin \cup_{i=1}^{p} B_r(x_p)$ for some $k \in \{1,\dots,l\}$, we have

$$D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mu) = 0. \tag{3.8}$$

To prove this, set $\mu_J = \mu|_{y_*} + \sum_{j\in J}\delta_{y_j}$ for $J \subset [l]$ and $y_* := \min\{y_1,\dots,y_l\}$, with the minimum taken with respect to $\prec$ order. From (3.1) we obtain

$$
\begin{aligned}
& D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mu) \\
&= \sum_{J\subset[l],k\notin J}(-1)^{l-|J|}\psi^!(x_1,\dots,x_p;\mu_J) + \sum_{J\subset[l],k\notin J}(-1)^{l-|J|-1}\psi^!(x_1,\dots,x_p;\mu_{J\cup\{k\}}) = 0,
\end{aligned}
$$

where the last equality follows by noting that for $J\subset[l]$ with $k\notin J$, $\psi^!(x_1,\dots,x_p;\mu_J) = \psi^!(x_1,\dots,x_p;\mu_{J\cup\{k\}})$ because $R^\xi(x,\mathcal{P})\in[1,r]$ by assumption.

Consider now $y_1,\dots,y_l\in\cup_{i=1}^p B_r(x_p)$. For $J\subset[l]$, the inequality $1\le R^\xi(x,\mathcal{P})\le r$ and (1.15) yield

$$
\psi^!(x_1,\dots,x_p;\mu_J) \le (\hat{c}r)^{K_p|J|+pK_p+\sum_{i=1}^p k_i\mu(B_r(x_i))}. \tag{3.9}
$$

The term pK_p in the exponent of (3.9) is due to $\sum_{j=1}^p\delta_{x_j}$ in the argument of ξ in (3.6). Substituting this bound in (3.1) yields

$$
\begin{aligned}
|D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mu)| &\le (\hat{c}r)^{pK_p+\sum_{i=1}^p k_i\mu(B_r(x_i))}\sum_{J\subset[l]}(\hat{c}r)^{K_p|J|} &(3.10)\\
&= (\hat{c}r)^{pK_p+\sum_{i=1}^p k_i\mu(B_r(x_i))}(1+(\hat{c}r)^{K_p})^l. &(3.11)
\end{aligned}
$$

We now consider $\psi^!(x_1,\dots,x_p;\mathcal{P}_n)$, with $\mathcal{P}_n := \mathcal{P}\cap W_n$ and $\psi^!$ defined as above. By the bound (3.10) we have

$$
\begin{aligned}
& \frac{1}{l!}\int_{\mathbb{R}^{dl}}(\mathbb{E}^!_{x_1,\dots,x_p})^!_{y_1,\dots,y_l}[|D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mathcal{P}_n)|]\rho^{(l)}_{x_1,\dots,x_p}(y_1,\dots,y_l)\,\mathrm{d}y_1\dots\mathrm{d}y_l \\
=& \frac{1}{l!}\int_{\mathbb{R}^{dl}}\mathbb{E}^!_{x_1,\dots,x_p,y_1,\dots,y_l}[|D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mathcal{P}_n)|]\rho^{(l)}_{x_1,\dots,x_p}(y_1,\dots,y_l)\,\mathrm{d}y_1\dots\mathrm{d}y_l \\
\le& \frac{(1+(\hat{c}r)^{K_p})^l(\hat{c}r)^{pK_p}}{l!}\mathbb{E}^!_{x_1,\dots,x_p}\left[(\mathcal{P}_n(\cup_{i=1}B_r(x_i))^l(\hat{c}r)^{\sum_{i=1}^p k_i\mathcal{P}_n(B_r(x_i))}\right] \\
\le& \frac{(1+(\hat{c}r)^{K_p})^l(\hat{c}r)^{pK_p}}{l!}\mathbb{E}^!_{x_1,\dots,x_p}\left[(\mathcal{P}_n(\cup_{i=1}^p B_r(x_i))^l(\hat{c}r)^{K_p\mathcal{P}_n(\cup_{i=1}^p B_r(x_i))}\right] \\
\le& \frac{(1+(\hat{c}r)^{K_p})^l}{l!}\mathbb{E}_{x_1,\dots,x_p}\left[(\mathcal{P}_n(\cup_{i=1}^p B_r(x_i))^l(\hat{c}r)^{K_p\mathcal{P}_n(\cup_{i=1}^p B_r(x_i))}\right], \qquad (3.12)
\end{aligned}
$$

where the last equality follows since the distribution of $\mathcal{P}$ under $\mathbb{P}_{x_1,\dots,x_p}$ is equal to that of $\mathcal{P}+\sum_{i=1}^p\delta_{x_i}$ under $\mathbb{P}^!_{x_1,\dots,x_p}$. Defining $N := \mathcal{P}_n(\cup_{i=1}^p B_r(x_i))$, we bound (3.12) by

$$
\mathbb{E}_{x_1,\dots,x_p}\left[(\hat{c}r)^{K_pN}\sum_{m=l}^\infty\frac{(1+(\hat{c}r)^{K_p})^l}{l!}N^l\right] \le \mathbb{E}_{x_1,\dots,x_p}\left[(\hat{c}r)^{(1+(\hat{c}r)^{K_p}+K_p)N}\right] < \infty\,,
$$

where the last inequality follows since $\mathcal{P}$ has exponential moments under the Palm measure as well (see Remark (i) at the beginning of Section 2.1). Consequently, by the Lebesgue dominated convergence theorem, the expression (3.12) converges to 0 as $l \to \infty$. Thus conditions (3.2) and (3.3) hold and (3.7) follows by Theorem 3.1.

Now we consider case (i), that is to say $\psi^!$ is as at (3.6) with ξ a U-statistic of type (A1). By Lemma 5.1, $\psi^!$ is a sum of U-statistics of orders not larger than $K_p(k-1)$. Consequently, for $l > K_p(k-1)$ we have

$$D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mu) = 0 \;\; \forall y_1,\dots,y_l \in \mathbb{R}^d, \tag{3.13}$$

as shown in [62, Lemma 3.3] for Poisson point processes (the proof for general simple counting measures μ is identical). This implies that conditions (3.2) for $l > K_p(k-1)$ and (3.3) are trivially satisfied for $\psi^!$ as at (3.6). Now, we need to verify the condition (3.2) for $l \le K_p(k-1)$. For $y_1,\dots,y_l \in \mathbb{R}^d$, set as before $\mu_J = \mu|_{y_*} + \sum_{j\in J}\delta_{y_j}$ for $J \subset [l]$ and $y_* := \min\{y_1,\dots,y_l\}$, with the minimum taken with respect to the order $\prec$. Since ξ has a bounded stabilization radius, by (3.8) and (2.5), we have

$$\begin{aligned}\psi^!(x_1,\dots,x_p;\mu_J) &\le \prod_{i=1}^p \|h\|_\infty^{k_i}(\mu(\cup_{i=1}^p B_r(x_i)) + |J| + p)^{k_i(k-1)} \\ &\le \|h\|_\infty^{K_p}(\mu(\cup_{i=1}^p B_r(x_i)) + |J| + p)^{K_p(k-1)},\end{aligned} \tag{3.14}$$

and so by (3.1), we derive that

$$\begin{aligned}|D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mu)| &\le \|h\|_\infty^{K_p}\sum_{J\subset[l]}(\mu(\cup_{i=1}^p B_r(x_i)) + |J| + p)^{K_p(k-1)} \\ &\le \|h\|_\infty^{K_p}2^l(\mu(\cup_{i=1}^p B_r(x_i)) + l + p)^{K_p(k-1)}.\end{aligned} \tag{3.15}$$

Consider $\psi^!(x_1,\dots,x_p;\mathcal{P}_n)$ with $\psi^!$ defined as above. Using the refined Campbell theorem (1.6), the bound (3.15) and following the calculations as in (3.12), we obtain

$$\begin{aligned}&\frac{1}{l!}\int_{\mathbb{R}^{dl}}(\mathbb{E}^!_{x_1,\dots,x_p})^!_{y_1,\dots,y_l}[|D^l_{y_1,\dots,y_l}\psi^!(x_1,\dots,x_p;\mathcal{P}_n)|]\rho^{(l)}_{x_1,\dots,x_p}(y_1,\dots,y_l)\,\mathrm{d}y_1\dots\mathrm{d}y_l \\ \le\; &\|h\|_\infty^{K_p}2^l\mathbb{E}_{x_1,\dots,x_p}[\mathcal{P}(\cup_{i=1}^p B_r(x_i))^l(\mathcal{P}(\cup_{i=1}^p B_r(x_i)) + l + p)^{K_p(k-1)}].\end{aligned}$$

Since $\mathcal{P}$ has all moments under the Palm measure (see Remark (ii) at the beginning of Section 2.1), the finiteness of the last term and hence the validity of the condition (3.2) for $l \le K_p(k-1)$ follows. This justifies the FME expansion (3.7), with finitely many non-zero terms, when $\psi^!$ is the product of score functions of type (A1). □

3.2 Proof of Theorem 1.10

First assume that $(\xi,\mathcal{P})$ is of type (A2). Later we consider pairs of type (A1). For fixed positive integers $p, q, k_1,\dots,k_{p+q}$ consider mixed moments $m^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n)$,

$m^{(k_{p+1},\ldots,k_{p+q})}(x_{p+1},\ldots,x_{p+q};n)$ and $m^{(k_1,\ldots,k_{p+q})}(x_1,\ldots,x_{p+q};n)$. Denote $K_p := \sum_{i=1}^p k_i$, $K_q := \sum_{i=p+1}^q k_i$ and $K := \sum_{i=1}^{p+q} k_i$. We abbreviate $\psi_{k_1,\ldots,k_p}(x_1,\ldots,x_p;\mu)$ by $\psi(x_1,\ldots,x_p;\mu)$ as at (3.5), and similarly for $\psi(x_{p+1},\ldots,x_{p+q};\mu)$ and $\psi(x_1,\ldots,x_{p+q};\mu)$.

Given $x_1,\ldots,x_{p+q}\in W_n$ we recall $s := d(\{x_1,\ldots,x_p\},\{x_{p+1},\ldots,x_{p+q}\})$. Without loss of generality we assume $s\in(4,\infty)$. Recalling the definition of b at (1.14) we may assume without loss of generality that $b\in(0,d)$. Put

$$t := t(s) := (\frac{s}{4})^{b(1-a)/(2(K+d))}, \tag{3.16}$$

where a is defined at (1.13). Since $s\in(4,\infty)$ and $K\geq 2$, we easily have $t\in(1,s/4)$. Given stabilization radii $R^\xi(x_i,\mathcal{P}_n), 1\leq i\leq p+q$, we put

$$\tilde{\xi}(x_i,\mathcal{P}_n) := \xi(x_i,\mathcal{P}_n\cap B_{R^\xi(x_i,\mathcal{P}_n)}(x))\mathbf{1}[R^\xi(x_i,\mathcal{P}_n)\leq t]$$

considered under $\mathbb{E}_{x_1,\ldots,x_p}$. We denote by $\tilde{m}^{(k_1,\ldots,k_p)}$ the p-mixed moments induced by $\tilde{\xi}$ and $\mathcal{P}_n$, that is

$$\tilde{m}^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p;n) := \mathbb{E}_{x_1,\ldots,x_p}[\tilde{\xi}(x_1,\mathcal{P}_n)^{k_1}\ldots\tilde{\xi}(x_p,\mathcal{P}_n)^{k_p}]\rho^{(p)}(x_1,\ldots,x_p).$$

Similarly we consider $\tilde{m}^{(k_{p+1},\ldots,k_{p+q})}$ and $\tilde{m}^{(k_1,\ldots,k_{p+q})}$. Put

$$\tilde{\psi}(x_1,\ldots,x_p;\mathcal{P}_n) = \prod_{i=1}^p \tilde{\xi}(x_i,\mathcal{P}_n)^{k_i} \tag{3.17}$$

and write $\tilde{\psi}(x_1,\ldots,x_p;\mathcal{P}_n) = \psi(x_1,\ldots,x_p;\mathcal{P}_n)\mathbf{1}[\max_{i\leq p} R^\xi(x_i,\mathcal{P}_n)\leq t]$.

Next, write $\mathbb{E}_{x_1,\ldots,x_p}\psi(x_1,\ldots,x_p;\mathcal{P}_n)$ as a sum of

$$\mathbb{E}_{x_1,\ldots,x_p}[\psi(x_1,\ldots,x_p;\mathcal{P}_n)\mathbf{1}[\max_{i\leq p} R^\xi(x_i,\mathcal{P}_n)\leq t]]$$

and

$$\mathbb{E}_{x_1,\ldots,x_p}[\psi(x_1,\ldots,x_p;\mathcal{P}_n)\mathbf{1}[\max_{i\leq p} R^\xi(x_i,\mathcal{P}_n)> t]].$$

The bounds (1.8), (1.11), the moment condition (1.16), Hölder's inequality, and $p\leq\sum_{i=1}^p k_i = K_p$ give for Lebesgue almost all $x_1,\ldots,x_p$

$$\begin{aligned}
&\Big|\mathbb{E}_{x_1,\ldots,x_p}\psi(x_1,\ldots,x_p;\mathcal{P}_n) - \mathbb{E}_{x_1,\ldots,x_p}\tilde{\psi}(x_1,\ldots,x_p;\mathcal{P}_n)\Big|\rho^{(p)}(x_1,\ldots,x_p)\\
&\leq \kappa_p(\tilde{M}_{K_p+1})^{K_p/(K_p+1)}\varphi(a_p t)^{1/(K_p+1)}\\
&\leq \kappa_{K_p}(\tilde{M}_{K_p+1})^{K_p/(K_p+1)}\varphi(a_{K_p} t)^{1/(K_p+1)}\\
&\leq c_1(K_p)\varphi(a_{K_p}t)^{1/(K_p+1)}.
\end{aligned} \tag{3.18}$$

Here $c_1(m) := \kappa_m\tilde{M}_{m+1}\geq\kappa_m(\tilde{M}_{m+1})^{m/(m+1)}$, as $\tilde{M}_m\geq 1$ by assumption. Similarly, condition (1.16) yields $|\mathbb{E}_{x_1,\ldots,x_p}\psi(x_1,\ldots,x_p;\mathcal{P}_n)|\rho^{(p)}(x_1,\ldots,x_p)\leq c_1(K_p)$. Using (3.18) with p replaced by $p+q$, we find $m^{(k_1,\ldots,k_{p+q})}(x_1,\ldots,x_{p+q};n)$ differs from $\tilde{m}^{(k_1,\ldots,k_{p+q})}(x_1,\ldots,x_{p+q};n)$ by $c_1(K)\varphi(a_K t)^{1/(K+1)}$, which is fast decreasing by (1.12).

For any reals $A, B, \tilde{A}, \tilde{B}$, with $|\tilde{B}| \leq |B|$ we have $|AB - \tilde{A}\tilde{B}| \leq |A(B-\tilde{B})| + |(A - \tilde{A})\tilde{B}| \leq (|A|+|B|)(|B-\tilde{B}| + |A-\tilde{A}|)$. Hence, it follows that

$$\begin{aligned}
&|m^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n)m^{(k_{p+1},\dots,k_q)}(x_{p+1},\dots,x_{p+q};n)\\
&- \tilde{m}^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n)\tilde{m}^{(k_{p+1},\dots,k_q)}(x_{p+1},\dots,x_{p+q};n)|\\
&\leq (c_1(K_p)+c_1(K_q))\Big(c_1(K_p)\varphi(a_{K_p}t)^{1/(K_p+1)} + c_1(K_q)\varphi(a_{K_q}t)^{1/(K_q+1)}\Big)\\
&\leq c_2(K)\varphi(a_K t)^{1/(K+1)},
\end{aligned}$$

with $c_2(m) := 4(c_1(m))^2$ and where we note that $\varphi(a_m t)^{1/(m+1)}$ is also fast decreasing by (1.12). The difference of mixed moments is thus bounded by

$$\begin{aligned}
&\Big|m^{(k_1,\dots,k_{p+q})}(x_1,\dots,x_{p+q};n) - m^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n)\, m^{(k_{p+1},\dots,k_{p+q})}(x_{p+1},\dots,x_{p+q};n)\Big|\\
\leq\ &(c_1(K)+c_2(K))\varphi(a_k t)^{1/(K+1)}\\
&+|\tilde{m}^{(k_1,\dots,k_{p+q})}(x_1,\dots,x_{p+q};n) - \tilde{m}^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n)\tilde{m}^{(k_{p+1},\dots,k_{p+q})}(x_{p+1},\dots,x_{p+q};n)|
\end{aligned} \tag{3.19}$$

The rest of the proof consists of bounding $|\tilde{m}^{(k_1,\dots,k_{p+q})} - \tilde{m}^{(k_1,\dots,k_p)}\tilde{m}^{(k_{p+1},\dots,k_{p+q})}|$ by a fast decreasing function of s. In this regard we will consider the expansion (3.7) with $\psi(x_1,\dots,x_p;\mathcal{P}_n)$ replaced by $\tilde{\psi}(x_1,\dots,x_p;\mathcal{P}_n)$ as at (3.17) and similarly for $\tilde{\psi}(x_{p+1},\dots,x_q;\mathcal{P}_n)$ and $\tilde{\psi}(x_1,\dots,x_{p+q};\mathcal{P}_n)$. By Lemma 5.2 in the Appendix, $\tilde{\xi}(x_i,\mathcal{P}_n), 1 \leq i \leq p$, have radii of stabilization bounded above by t and also satisfy the power-growth condition (1.15) since $\tilde{\xi}(.,.) \leq \xi(.,.)$. Thus the pair $(\tilde{\xi},\mathcal{P})$ satisfies the assumptions of Lemma 3.2. The corresponding version of $\tilde{\psi}$, which accounts for the fixed atoms of $\mathcal{P}_n$ is now

$$\tilde{\psi}^!(x_1,\dots,x_p;\mu) := \prod_{i=1}^{p} \tilde{\xi}(x_i, \mu + \textstyle\sum_{i=1}^p \delta_{x_i})^{k_i}$$

and similarly for $\tilde{\psi}^!(x_{p+1},\dots,x_q;\mathcal{P}_n)$ and $\tilde{\psi}^!(x_1,\dots,x_{p+q};\mathcal{P}_n)$.

Applying (3.7), (2.11) and (3.8), we obtain

$$\begin{aligned}
&\tilde{m}^{(k_1,\dots,k_{p+q})}(x_1,\dots,x_{p+q})\\
&= \mathbb{E}^!_{x_1,\dots,x_{p+q}}[\tilde{\psi}^!(x_1,\dots,x_{p+q};\mathcal{P}_n)]\rho^{(p+q)}(x_1,\dots,x_{p+q})\\
&= \sum_{l=0}^{\infty}\frac{1}{l!}\int_{(W_n)^l} D^l_{y_1,\dots,y_l}\tilde{\psi}^!(o)\rho^{(l+p+q)}(x_1,\dots,x_{p+q},y_1,\dots,y_l)\,\mathrm{d}y_1\dots\mathrm{d}y_l\\
&= \sum_{l=0}^{\infty}\frac{1}{l!}\int_{(\cup_{i=1}^{p+q} W_n\cap B_t(x_i))^l} D^l_{y_1,\dots,y_l}\tilde{\psi}^!(o)\rho^{(l+p+q)}(x_1,\dots,x_{p+q},y_1,\dots,y_l)\,\mathrm{d}y_1\dots\mathrm{d}y_l\,.
\end{aligned}$$

Put $B_{t,n}(x_i) := B_t(x_i)\cap W_n$. Applying (3.1) when μ is the null measure, this gives for

$\alpha^{(p+q)}$ almost all $x_1, \ldots, x_{p+q}$

$$\tilde{m}^{(k_1,\ldots,k_{p+q})}(x_1, \ldots, x_{p+q}) \tag{3.20}$$
$$= \sum_{l=0}^{\infty} \frac{1}{l!} \sum_{j=0}^{l} \frac{l!}{j!(l-j)!} \int_{(\cup_{i=1}^{p} B_{t,n}(x_i))^j \times (\cup_{i=1}^{q} B_{t,n}(x_{p+i}))^{l-j}} D^l_{y_1,\ldots,y_l} \tilde{\psi}^!(x_1, \ldots, x_{p+q}; o)$$
$$\times \rho^{(l+p+q)}(x_1, \ldots, x_{p+q}, y_1, \ldots, y_l) \mathrm{d}y_1 \ldots \mathrm{d}y_l$$

$$= \sum_{l=0}^{\infty} \sum_{j=0}^{l} \frac{1}{j!(l-j)!} \int_{(\cup_{i=1}^{p} B_{t,n}(x_i))^j \times (\cup_{i=1}^{q} B_{t,n}(x_{p+i}))^{l-j}}$$
$$\times \sum_{J \subset [l]} (-1)^{l-|J|} \tilde{\psi}^!(x_1, \ldots, x_{p+q}; \textstyle\sum_{j \in J} \delta_{y_j}) \times \rho^{(l+p+q)}(x_1, \ldots, x_{p+q}, y_1, \ldots, y_l) \mathrm{d}y_1 \ldots \mathrm{d}y_l. \tag{3.21}$$

To compare the $(p+q)$th mixed moments with the product of p, q-mixed moments, we shall use the fact that $R^{\tilde{\xi}}(x_i, \mathcal{P}_n) \in (0, t]$ (cf. Lemma 5.2) implies the following factorization, which holds for $y_1, \ldots, y_j \in \cup_{i=1}^{p} B_t(x_i)$ and $y_{j+1}, \ldots, y_l \in \cup_{i=1}^{q} B_t(x_{p+i})$, with $t \in (1, s/4)$ (making $\cup_{i=1}^{p} B_t(x_i)$ and $\cup_{i=1}^{q} B_t(x_{p+i})$ disjoint):

$$\tilde{\psi}^!(x_1, \ldots, x_{p+q}; \textstyle\sum_{i=1}^{l} \delta_{y_i}) = \tilde{\psi}^!(x_1, \ldots, x_p; \sum_{i=1}^{j} \delta_{y_i}) \tilde{\psi}^!(x_{p+1}, \ldots, x_{p+q}; \sum_{i=j+1}^{l} \delta_{y_i}) \,. \tag{3.22}$$

Using the expansion (3.7) as above along with (3.22), we next derive an expansion for the product of pth and qth mixed moments. Recalling (2.11) and that by (3.6) we have $\mathbb{E}_{x_1,\ldots,x_p}[\psi(\mathcal{P}_n)] = \mathbb{E}^!_{x_1,\ldots,x_p}[\psi^!(\mathcal{P}_n)]$, we obtain

$$\tilde{m}^{(k_1,\ldots,k_p)}(x_1, \ldots, x_p) \tilde{m}^{(k_{p+1},\ldots,k_q)}(x_{p+1}, \ldots, x_{p+q})$$
$$= \mathbb{E}^!_{x_1,\ldots,x_p}[\tilde{\psi}^!(x_1, \ldots, x_p; \mathcal{P}_n)] \mathbb{E}^!_{x_{p+1},\ldots,x_{p+q}}[\tilde{\psi}^!(x_{p+1}, \ldots, x_{p+q}; \mathcal{P}_n)]$$
$$\times \rho^{(p)}(x_1, \ldots, x_p) \rho^{(q)}(x_{p+1}, \ldots, x_{p+q})$$

$$= \sum_{l_1,l_2=0}^{\infty} \frac{1}{l_1! l_2!} \int_{(\cup_{i=1}^{p} B_{t,n}(x_i))^{l_1} \times (\cup_{i=1}^{q} B_{t,n}(x_{p+i}))^{l_2}} D^{l_1}_{y_1,\ldots,y_{l_1}} \tilde{\psi}^!(x_1, \ldots, x_p; o) \; D^{l_2}_{z_1,\ldots,z_{l_2}} \tilde{\psi}^!(x_{p+1}, \ldots, x_{p+q}; o)$$
$$\times \rho^{(l_1+p)}(x_1, \ldots, x_p, y_1, \ldots, y_{l_1}) \rho^{(l_2+q)}(x_{p+1}, \ldots, x_{p+q}, z_1, \ldots, z_{l_2}) \, \mathrm{d}y_1 \ldots \mathrm{d}y_{l_1} \mathrm{d}z_1 \ldots \mathrm{d}z_{l_2} \,. \tag{3.23}$$

Applying (3.1) once more for μ the null measure, this gives

$$\begin{aligned}
&\tilde{m}^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p)\tilde{m}^{(k_{p+1},\ldots,k_q)}(x_{p+1},\ldots,x_{p+q})\\
&=\sum_{l_1,l_2=0}^{\infty}\frac{1}{l_1!l_2!}\int_{(\cup_{i=1}^{p}B_{t,n}(x_i))^{l_1}\times(\cup_{i=1}^{q}B_{t,n}(x_{p+i}))^{l_2}}\sum_{J_1\subset[l_1],J_2\subset[l_2]}(-1)^{l_1+l_2-|J_1|-|J_2|}\\
&\quad\times\tilde{\psi}^!(x_1,\ldots,x_p;\textstyle\sum_{i\in J_1}\delta_{y_i})\tilde{\psi}^!(x_{p+1},\ldots,x_{p+q};\sum_{i\in J_2}\delta_{z_i})\\
&\quad\times\rho^{(l_1+p)}(x_1,\ldots,x_p,y_1,\ldots,y_{l_1})\rho^{(l_2+q)}(x_{p+1},\ldots,x_{p+q},z_1,\ldots,z_{l_2})\,\mathrm{d}y_1\ldots\mathrm{d}y_{l_1}\mathrm{d}z_1\ldots\mathrm{d}z_{l_2}\\
&=\sum_{l=0}^{\infty}\sum_{j=0}^{l}\frac{1}{j!(l-j)!}\int_{(\cup_{i=1}^{p}B_{t,n}(x_i))^{j}\times(\cup_{i=1}^{q}B_{t,n}(x_{p+i}))^{l-j}}\sum_{J_1\subset[j],J_2\subset[l]\setminus[j]}(-1)^{l-|J_1|-|J_2|}\\
&\quad\times\tilde{\psi}^!(x_1,\ldots,x_p;\textstyle\sum_{i\in J_1}\delta_{y_i})\tilde{\psi}^!(x_{p+1},\ldots,x_{p+q};\sum_{i\in J_2}\delta_{y_i})\\
&\quad\times\rho^{(j+p)}(x_1,\ldots,x_p,y_1,\ldots,y_j)\rho^{(l-j+q)}(x_{p+1},\ldots,x_{p+q},y_{j+1},\ldots,y_l)\,\mathrm{d}y_1\ldots\mathrm{d}y_l\\
&=\sum_{l=0}^{\infty}\sum_{j=0}^{l}\frac{1}{j!(l-j)!}\int_{(\cup_{i=1}^{p}B_{t,n}(x_i))^{j}\times(\cup_{i=1}^{q}B_{t,n}(x_{p+i}))^{l-j}}\sum_{J\subset[l]}(-1)^{l-|J|}\tilde{\psi}^!(x_1,\ldots,x_{p+q};\textstyle\sum_{i\in J}\delta_{y_i})\\
&\quad\times\rho^{(j+p)}(x_1,\ldots,x_p,y_1,\ldots,y_j)\rho^{(l-j+q)}(x_{p+1},\ldots,x_{p+q},y_{j+1},\ldots,y_l)\,\mathrm{d}y_1\ldots\mathrm{d}y_l,
\end{aligned}\tag{3.24}$$

where we have used (3.22) in the last equality.

Now we estimate the difference of (3.21) and (3.24). Applying the clustering bounds (1.7) and replacing $B_{t,n}(x_i)$ with $B_t(x_i)$, we obtain

$$\begin{aligned}
&|\tilde{m}^{(k_1,\ldots,k_{p+q})}(x_1,\ldots,x_{p+q})-\tilde{m}^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p)\tilde{m}^{(k_{p+1},\ldots,k_q)}(x_{p+1},\ldots,x_{p+q})|\\
&\leq\phi(\frac{s}{2})\sum_{l=0}^{\infty}\sum_{j=0}^{l}\frac{C_{l+p+q}}{j!(l-j)!}\\
&\qquad\times\int_{(\cup_{i=1}^{p}B_t(x_i))^{j}\times(\cup_{i=1}^{q}B_t(x_{p+i}))^{l-j}}\sum_{J\subset[l]}|\tilde{\psi}^!(x_1,\ldots,x_{p+q};\textstyle\sum_{i\in J}\delta_{y_i})|\,\mathrm{d}y_1\ldots\mathrm{d}y_l.
\end{aligned}\tag{3.25}$$

Recalling (3.22) and (3.9), we bound $\sum_{J\subset[l]}|\tilde{\psi}^!(x_1,\ldots,x_{p+q};\sum_{i\in J}\delta_{y_i})|$ by $2^l(\hat{c}t)^{jK_p+(l-j)K_q+K}$, where $\hat{c}t\geq 1$ holds since $\hat{c}\geq 1$ in (1.15). This gives

$$\begin{aligned}
&|\tilde{m}^{(k_1,\ldots,k_{p+q})}(x_1,\ldots,x_{p+q})-\tilde{m}^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p)\tilde{m}^{(k_{p+1},\ldots,k_q)}(x_{p+1},\ldots,x_{p+q})|\\
&\leq\phi(\frac{s}{2})\sum_{l=0}^{\infty}\sum_{j=0}^{l}\frac{C_{l+p+q}}{j!(l-j)!}\int_{(\cup_{i=1}^{p}B_t(x_i))^{j}\times(\cup_{i=1}^{q}B_t(x_{p+i}))^{l-j}}2^l(\hat{c}t)^{jK_p+(l-j)K_q+K}\,\mathrm{d}y_1\ldots\mathrm{d}y_l
\end{aligned}\tag{3.26}$$

Consequently,

$$
\begin{aligned}
&|\tilde{m}^{(k_1,\dots,k_{p+q})}(x_1,\dots,x_{p+q}) - \tilde{m}^{(k_1,\dots,k_p)}(x_1,\dots,x_p)\tilde{m}^{(k_{p+1},\dots,k_q)}(x_{p+1},\dots,x_{p+q})| \\
&\le \phi(\frac{s}{2}) \sum_{l=0}^{\infty} C_{l+p+q} 2^l (\hat{c}t)^{(l+1)K} ((p+q)\theta_d t^d)^l \sum_{j=0}^{l} \frac{1}{j!(l-j)!} \\
&\le \phi(\frac{s}{2}) \sum_{l=0}^{\infty} \frac{C_{l+p+q}}{l!} 4^l (\hat{c}t)^{(l+1)K} ((p+q)\theta_d t^d)^l \\
&\le \phi(\frac{s}{2}) \sum_{l=0}^{\infty} \frac{C_{l+K}}{l!} 4^l (\hat{c}t)^{(l+1)K} (K\theta_d t^d)^l, \qquad (3.27)
\end{aligned}
$$

where $\theta_d := \pi^{d/2}/\Gamma(d/2+1)$ is the volume of the unit ball in $\mathbb{R}^d$ and where the second inequality follows from $jK_p + (l-j)K_q < jK + (l-j)K = lK$ and where the last inequality uses $p+q \le K$. We observe, using the bound (1.13), that there are constants c_1, c_2 and c_3 depending only on a, d and $K = \sum_{i=1}^{p+q} k_i$ such that

$$
\sum_{l=0}^{\infty} \frac{C_{l+K}}{l!} 4^l (\hat{c}t)^{(l+1)K} (K\theta_d t^d)^l \le t^K \sum_{l=0}^{\infty} \frac{c_1 c_2^l l^{c_3} (t^{K+d})^l \cdot l^{al}}{l!}
$$

Using Stirling's formula we find that there are constants c_4, c_5 and c_6 depending only on a, d and K, such that

$$
t^K \sum_{l=0}^{\infty} \frac{C_{l+K}}{l!} 4^l (\hat{c}t)^{lK} (K\theta_d t^d)^l \le t^K \sum_{l=0}^{\infty} \frac{c_4 c_5^l l^{c_6} (t^{K+d})^l}{(\lfloor l(1-a) \rfloor)!},
$$

where $\lfloor r \rfloor$ is the greatest integer less than the real r. We compute

$$
\begin{aligned}
t^K \sum_{l=0}^{\infty} \frac{C_{l+K}}{l!} 4^l (\hat{c}t)^{lK} (K\theta_d t^d)^l \quad &\le \quad t^K \sum_{n=0}^{\infty} \sum_{\{l:\ \lfloor l(1-a) \rfloor = n\}} \frac{c_4 c_5^l l^{c_6} (t^{K+d})^l}{n!} \\
&\le \quad t^K \sum_{n=0}^{\infty} \frac{c_4 c_5^n n^{c_6} (t^{K+d})^{(n+1)/(1-a)}}{(1-a)n!} \\
&\le \quad c_7 \exp(c_8 t^{(K+d)/(1-a)}) \qquad (3.28)
\end{aligned}
$$

where c_7 and c_8 depend only on a, d and K.

Recalling from (3.16) that $t := (s/4)^{b(1-a)/(2(K+d))}$ we obtain

$$
\sum_{l=0}^{\infty} \frac{C_{l+K}}{l!} 4^l (\hat{c}t)^{(l+1)K} (K\theta_d t^d)^l \le c_7 \exp\left(c_8 (\frac{s}{4})^{\frac{b}{2}}\right).
$$

By (1.14), there is a constant c_9 depending only on a such that for all s we have

$\phi(s) \le c_9 \exp(-s^b/c_9)$. Combining this with (3.27) and (3.28) gives

$$|\tilde{m}^{(k_1,\dots,k_{p+q})} - \tilde{m}^{(k_1,\dots,k_p)}\tilde{m}^{(k_{p+1},\dots,k_{p+q})}| \le c_7 c_9 \exp\left(\frac{-(s/2)^b}{c_9} + c_8(\frac{s}{4})^{\frac{b}{2}}\right).$$

This along with (3.19) establishes (1.17) when $(\xi, \mathcal{P})$ is an admissible pair of class (A2).

Now we turn to the case where $(\xi, \mathcal{P})$ is of class (A1). Follow the arguments for case (A2) word for word using that $\sup_{x\in\mathcal{P}} R^\xi(x, \mathcal{P}) \le r$. Notice that when $l > (k-1)K$, $l_1 > (k-1)K_p$, and $l_2 > (k-1)K_q$ the respective summands in (3.20), (3.23), and (3.27) all vanish. Hence, the finiteness of $\tilde{C}_K$ trivially follows (without the need for exponential decay of ϕ or a growth-rate bound on C_k), establishing (1.17) when $(\xi, \mathcal{P})$ is of class (A1). $\square$

4 Proof of main results

We provide the proofs of Theorems 1.11, 1.14, and 1.12 in this order.

4.1 Proof of Theorem 1.11

4.1.1 Proof of expectation asymptotics (1.19).

We have by the definition of the Palm probabilities.

$$n^{-1}\mathbb{E}\mu_n^\xi(f) = n^{-1}\int_{W_n} f(n^{-1/d}u)\mathbb{E}_u\xi(u, \mathcal{P}_n)\rho^{(1)}(u)\,\mathrm{d}u.$$

By the stationarity of $\mathcal{P}$ and translation invariance of ξ we have $\mathbb{E}_{\mathbf{0}}\xi(\mathbf{0}, \mathcal{P})\,\mathrm{d}u = \mathbb{E}_u\xi(u, \mathcal{P})\,\mathrm{d}u$. Using this we have

$$\begin{aligned}
&\left|n^{-1}\mathbb{E}\mu_n^\xi(f) - \mathbb{E}_{\mathbf{0}}\xi(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0})\int_{W_1} f(x)\,\mathrm{d}x\right| \\
&= \left|n^{-1}\int_{W_n} f(n^{-1/d}u)\{\mathbb{E}_u\xi(u, \mathcal{P}_n)\rho^{(1)}(u) - \mathbb{E}_{\mathbf{0}}\xi(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0})\}\,\mathrm{d}u\right| \\
&= \left|n^{-1}\int_{W_n} f(n^{-1/d}u)\mathbb{E}_u[(\xi(u, \mathcal{P}_n) - \xi(u, \mathcal{P}))\rho^{(1)}(u)]\mathrm{d}u\right| \\
&\le \|f\|_\infty n^{-1}\int_{W_n} \mathbb{E}_u[|\xi(u, \mathcal{P}_n) - \xi(u, \mathcal{P})|\mathbf{1}[R^\xi(u, \mathcal{P}) \ge d(u, \partial W_n)]]\rho^{(1)}(u)\mathrm{d}u \\
&\le 2\kappa_1\|f\|_\infty n^{-1}\tilde{M}_p\int_{W_n} (\mathbb{P}_u(R^\xi(u, \mathcal{P}) \ge d(u, \partial W_n))^{1/q}\mathrm{d}u,
\end{aligned}$$

where the last inequality follows from the Hölder inequality, the bound (1.8), the p-moment condition (1.16) (recall $p \in (1, \infty)$ and $\tilde{M}_p \in [1, \infty)$), and where $1/p + 1/q = 1$.

By (1.12) we have

$$\int_{W_n} (\mathbb{P}_u(R^\xi(u,\mathcal{P}) \geq d(u,\partial W_n))^{1/q} \, \mathrm{d}u = O(n^{(d-1)/d}),$$

which gives (1.19) as desired. If ξ satisfies (1.11), but not (1.12), then we note that

$$\begin{aligned}
&\limsup_{n\to\infty} n^{-1} \int_{W_n} (\mathbb{P}_u(R^\xi(u,\mathcal{P}) \geq d(u,\partial W_n))^{1/q} \mathrm{d}u \\
&\leq \limsup_{n\to\infty} n^{-1} \int_{W_n} (\varphi(a_1 d(u,\partial W_n)))^{1/q} \mathrm{d}u \\
&= \limsup_{n\to\infty} \int_{W_1} (\varphi(a_1 n^{1/d} d(z,\partial W_1)))^{1/q} \mathrm{d}z = 0
\end{aligned}$$

where the last equality follows from (1.11) and the bounded convergence theorem. Thus

$$\left| n^{-1}\mathbb{E}\mu_n^\xi(f) - \mathbb{E}_{\mathbf{0}}\xi(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0}) \int_{W_1} f(x)\,\mathrm{d}x \right| = o(1)$$

which gives expectation asymptotics under (1.11).

4.1.2 Proof of variance asymptotics (1.20).

Recall the definiton of mixed moments from (1.5).

$$\begin{aligned}
\mathrm{Var}\mu_n^\xi(f) &= \mathbb{E}\sum_{x\in\mathcal{P}_n} f(n^{-1/d}x)^2 \xi^2(x,\mathcal{P}_n) \\
&\quad + \mathbb{E}\sum_{x,y\in\mathcal{P}_n, x\neq y} f(n^{-1/d}x) f(n^{-1/d}y)\xi(x,\mathcal{P}_n)\xi(y,\mathcal{P}_n) - \Big(\mathbb{E}\sum_{x\in\mathcal{P}_n} f(n^{-1/d}x)\xi(x,\mathcal{P}_n)\Big)^2 \\
&= \int_{W_n} f(n^{-1/d}u)^2 \mathbb{E}_u(\xi^2(u,\mathcal{P}_n))\rho^{(1)}(u)\,\mathrm{d}u \qquad (4.1) \\
&\quad + \int_{W_n\times W_n} f(n^{-1/d}u) f(n^{-1/d}v)\big(m_{(2)}(u,v;n) - m_{(1)}(u;n)m_{(1)}(v;n)\big)\mathrm{d}u\mathrm{d}v. \qquad (4.2)
\end{aligned}$$

Since ξ satisfies the p-moment condition (1.16) for $p > 2$, we have that ξ^2 satisfies the p-moment condition for $p > 1$. Also, ξ and ξ^2 have the same radius of stabilization. Thus, the proof of expectation asymptotics, with ξ replaced by ξ^2, shows that the first term in (4.1), multiplied by n^{-1}, converges to

$$\mathbb{E}_{\mathbf{0}}\xi^2(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0}) \int_{W_1} f(x)^2\,\mathrm{d}x\,;$$

cf. expectation asymptotics (1.19). The second term in (4.2) multiplied by n^{-1} can be rewritten as follows by setting $x = n^{-1/d}u$ and $z = v - u = v - n^{1/d}x$

$$\int_{W_1} \int_{W_n - n^{1/d}x} f(x + n^{-1/d}z) f(x) \times [m_{(2)}(n^{1/d}x, n^{1/d}x + z; n) - m_{(1)}(n^{1/d}x; n) m_{(1)}(n^{1/d}x + z; n)] \, \mathrm{d}z\mathrm{d}x. \tag{4.3}$$

Setting $\mathcal{P}_n^x := \mathcal{P} \cap (W_n - n^{1/d}x)$, translation invariance of ξ and stationarity of $\mathcal{P}$ yields

$$\begin{aligned} m_{(2)}(n^{1/d}x, n^{1/d}x + z; n) &= m_{(2)}(\mathbf{0}, z; \mathcal{P}_n^x) \\ m_{(1)}(n^{1/d}x; n) &= m_{(1)}(\mathbf{0}; \mathcal{P}_n^x) \\ m_{(1)}(n^{1/d}x + z; n) &= m_{(1)}(z; \mathcal{P}_n^x) \,. \end{aligned}$$

Putting aside for the moment technical details one expects that the above moments converge to $m_{(2)}(\mathbf{0}, z)$, $m_{(1)}(\mathbf{0})$ and $m_{(1)}(z) = m_{(1)}(\mathbf{0})$, respectively, when $n \to \infty$. Moreover, splitting the inner integral in (4.3) into two terms

$$\int_{W_n - n^{1/d}x} (\dots) \mathrm{d}z = \int_{W_n - n^{1/d}x} \mathbf{1}[|z| \le M](\dots) \mathrm{d}z + \int_{W_n - n^{1/d}x} \mathbf{1}[|z| > M](\dots) \mathrm{d}z \tag{4.4}$$

for any $M > 0$, we see (at least when f is continuous) that the first term in the right-hand side of (4.4) converges to the desired value

$$\int_{\mathbb{R}^d} f(x)^2 [m_{(2)}(\mathbf{0}, z) - m_{(1)}(\mathbf{0})^2] \, \mathrm{d}z$$

when first $n \to \infty$ and then $M \to \infty$. The absolute value of the second term in (4.4), by the strong clustering of the second mixed moment, cf (1.17), can be bounded uniformly in n by

$$\|f\|_\infty^2 \tilde{C}_2 \int_{|z| > M} \tilde{\phi}(\tilde{c}_2 z) \, \mathrm{d}z \,,$$

which goes to 0 when $M \to \infty$ since $\tilde{\phi}(\cdot)$ is fast decreasing (and thus integrable).

To formally justify the above statements we need the following lemma. Denote

$$h_n^\xi(x, z) := m_{(2)}(\mathbf{0}, z; \mathcal{P}_n^x) - m_{(1)}(\mathbf{0}; \mathcal{P}_n^x) \, m_{(1)}(z; \mathcal{P}_n^x) \,.$$

Lemma 4.1. *Assume that translation invariant score function ξ on the input process $\mathcal{P}$ satisfies (1.11) and the p-moment condition (1.16) for $p \in (2, \infty)$. Then $h_n^\xi(x, z)$ is uniformly bounded*

$$\sup_{n \le \infty} \sup_{x \in W_1} \sup_{z \in W_n - n^{1/d}x} |h_n^\xi(x, z)| \le C_h < \infty$$

for some constant C_h and

$$\lim_{n \to \infty} h_n^\xi(x, z) = h_\infty^\xi(x, z) = m_{(2)}(\mathbf{0}, z) - (m_{(1)}(\mathbf{0}))^2 \,.$$

Proof. Denote $X_n := \xi(\mathbf{0}, \mathcal{P}_n^x)$, $Y_n := \xi(z, \mathcal{P}_n^x)$, $X := \xi(\mathbf{0}, \mathcal{P})$, and $Y := \xi(z, \mathcal{P})$. We shall prove first that all expectations $\mathbb{E}_{\mathbf{0},z}(X_n^2)$, $\mathbb{E}_{\mathbf{0},z}(Y_n^2)$, $\mathbb{E}_{\mathbf{0},z}(X^2)$ $\mathbb{E}_{\mathbf{0},z}(Y^2)$, $\mathbb{E}_{\mathbf{0}}|X_n|$, $\mathbb{E}_z|Y_n|$, $\mathbb{E}_{\mathbf{0}}|X|$ and $\mathbb{E}_z|Y|$ are uniformly bounded. Indeed, by the Hölder inequality

$$\mathbb{E}_{\mathbf{0},z}(X_n^2) \le (\mathbb{E}_{\mathbf{0},z}|X_n|^p)^{2/p} = (\mathbb{E}_{n^{1/d}x,z}|\xi(n^{1/d}x, \mathcal{P}_n)|^p)^{2/p} \le \tilde{M}_p^{2/p} \tag{4.5}$$

where in the last inequality we have used p-moment condition (1.16) for $p > 2$. Similarly $\mathbb{E}_{\mathbf{0},z}(Y_n^2)$ and $\mathbb{E}_{\mathbf{0},z}(X^2)$, $\mathbb{E}_{\mathbf{0},z}(Y^2)$ are bounded by $\tilde{M}_p^{2/p}$. Again using p-moment condition (1.16), we obtain

$$\mathbb{E}_{\mathbf{0}}|X_n| \le (\mathbb{E}_{\mathbf{0}}(X_n)^2)^{1/2} \le (\mathbb{E}_{n^{1/d}x}|\xi^2(n^{1/d}x, \mathcal{P}_n)|)^{1/2} \le \tilde{M}_p^{1/2}$$

and similarly for $\mathbb{E}_z|Y_n|$, $\mathbb{E}_{\mathbf{0}}|X|$ and $\mathbb{E}_z|Y|$. This proves the uniform bound of $|h_n^\xi(x, z)|$. To prove the convergence notice that

$$\begin{aligned} |m_{(2)}(\mathbf{0}, z; \mathcal{P}_n^x) - m_{(2)}(\mathbf{0}, z)| &= |\mathbb{E}_{\mathbf{0},z}(X_nY_n) - \mathbb{E}_{\mathbf{0},z}(XY)|\rho^{(2)}(\mathbf{0}, z) &(4.6)\\ &\le \kappa_2\big(\mathbb{E}_{\mathbf{0},z}|X_nY_n - X_nY| + \mathbb{E}_{\mathbf{0},z}|X_nY - XY|\big) \\ &\le \kappa_2(\mathbb{E}_{\mathbf{0},z}(X_n^2)\mathbb{E}_{\mathbf{0},z}(Y_n - Y)^2)^{1/2} + \kappa_2(\mathbb{E}_{\mathbf{0},z}(Y^2)\mathbb{E}_{\mathbf{0},z}(X_n - X)^2)^{1/2}, &(4.7)\end{aligned}$$

where κ_2 bounds the second correlation function as at (1.8). We have already proved that $\mathbb{E}_{\mathbf{0},z}(X_n^2)$, $\mathbb{E}_{\mathbf{0},z}(Y^2)$ are bounded. Moreover

$$\begin{aligned} \mathbb{E}_{\mathbf{0},z}(X_n - X)^2 &= \mathbb{E}_{\mathbf{0},z}((X_n - X)^2\mathbf{1}[X_n \ne X]) \\ &\le \mathbb{E}_{\mathbf{0},z}(X_n^2\mathbf{1}[X_n \ne X]) + 2\mathbb{E}_{\mathbf{0},z}(|X_nX|\mathbf{1}[X_n \ne X]) + \mathbb{E}_{\mathbf{0},z}(X^2\mathbf{1}[X_n \ne X]). \end{aligned}$$

The Hölder inequality gives for $p > 2$ and $2/p + 1/q = 1$,

$$\mathbb{E}_{\mathbf{0},z}(X_n^2\mathbf{1}[X_n \ne X]) \le (\mathbb{E}_{\mathbf{0},z}(X_n^p))^{2/p}(\mathbb{P}_{\mathbf{0},z}(X_n \ne X))^{1/q}$$

$$\begin{aligned} \mathbb{E}_{\mathbf{0},z}(|X_nX|\mathbf{1}[X_n \ne X]) &\le (\mathbb{E}_{\mathbf{0},z}(X_n^p)\mathbb{E}_{\mathbf{0},z}(X^p))^{1/p}(\mathbb{P}_{\mathbf{0},z}(X_n \ne X))^{1/q} \\ \mathbb{E}_{\mathbf{0},z}(X^2\mathbf{1}[X_n \ne X]) &\le (\mathbb{E}_{\mathbf{0},z}(X^p)^{2/p}(\mathbb{P}_{\mathbf{0},z}(X_n \ne X))^{1/q}. \end{aligned}$$

The pth moment of X_n and X under $\mathbb{E}_{\mathbf{0},z}$ can be bounded by $\tilde{M}_p$ using the p-moment condition (1.16) with $p > 2$ as in (4.5). By stabilization (1.11) with $l = 2$

$$\mathbb{P}_{\mathbf{0},z}(X_n \ne X) \le \mathbb{P}_{\mathbf{0},z}(R^\xi(\mathbf{0}, \mathcal{P}) > n^{1/d}d(x, \partial W_1)) \le \varphi(a_2 n^{1/d} d(x, \partial W_1)) \tag{4.8}$$

with the right-hand side converging to 0 for all $x \notin \partial W_1$. This proves that $\mathbb{E}_{\mathbf{0},z}(X_n - X)^2$ and (by the very same arguments) $\mathbb{E}_{\mathbf{0},z}(Y_n - Y)^2$ converge to 0 as $n \to \infty$ for all $x \notin \partial W_1$. Concluding this part of the proof, we have shown that the expression in (4.7) converges to 0 and thus $m_{(2)}(\mathbf{0}, z; \mathcal{P}_n^x)$ converges to $m_{(2)}(\mathbf{0}, z)$.

Using similar arguments with

$$\begin{aligned}|m_{(1)}(\mathbf{0},\mathcal{P}_n^x)-m_{(1)}(\mathbf{0})| &= |\mathbb{E}_{\mathbf{0}}(X_n)-\mathbb{E}_{\mathbf{0}}(X)|\rho^{(1)}(\mathbf{0})\\ &\le \kappa_1((\mathbb{E}_{\mathbf{0}}(X_n)^2)^{1/2}+(\mathbb{E}_{\mathbf{0}}(X^2))^{1/2})(\mathbb{P}_{\mathbf{0}}(X_n\neq X)))^{1/2},\end{aligned}$$

by the p-moment condition (1.16) and the stabilization property (1.11) for $p=1$ one can show that $m_{(1)}(\mathbf{0},\mathcal{P}_n^x)$ converges to $m_{(1)}(\mathbf{0})$ uniformly in x for all $x\in W_1\setminus\partial W_1$. Exactly the same arguments assure convergence of $m_{(1)}(z,\mathcal{P}_n^x)$ to $m_{(1)}(z)=m_{(1)}(\mathbf{0})$. This concludes the proof of Lemma 4.1. □

In order to complete the proof of variance asymptotics for general $f\in\mathcal{B}(W_1)$ (not necessarily continuous) we use arguments borrowed from the proof of [56, Theorem 2.1]. Recall that $x\in W_1$ is a Lebesgue point for f if $(\mathrm{Vol}B_\epsilon(x))^{-1}\int_{B_\epsilon(x)}|f(z)-f(x)|\mathrm{d}z\to 0$ as $\epsilon\to 0$. Denote by $\mathcal{C}_f$ all Lebesgue points of f in W_1. By the Lebesgue density theorem almost every $x\in W_1$ is a Lebesgue point of f and thus for any $M>0$ and n large enough the double integral in (4.3) is equal to

$$\begin{aligned}&\int_{W_1}\mathbf{1}[x\in\mathcal{C}_f]f(x)\int_{W_n-n^{1/d}x}f(x+n^{-1/d}z)h_n^\xi(x,z)\,\mathrm{d}z\mathrm{d}x\\ &=\int_{W_1}\mathbf{1}[x\in\mathcal{C}_f]f(x)\int_{|z|\le M}f(x+n^{-1/d}z)h_n^\xi(x,z)\,\mathrm{d}z\mathrm{d}x\\ &+\int_{W_1}\mathbf{1}[x\in\mathcal{C}_f]f(x)\int_{W_n-n^{1/d}x}1(|z|>M)f(x+n^{-1/d}z)h_n^\xi(x,z)\,\mathrm{d}z\mathrm{d}x\,.\end{aligned}$$

As already explained, by the strong clustering property of the second mixed moment, the second term converges to 0 as first $n\to\infty$ and then $M\to\infty$. Considering the first term, by the uniform boundedness of $h_n^\xi(x,z)$, using the dominated convergence theorem, it is enough to prove for any Lebesgue point x of f and fixed M that

$$\lim_{n\to\infty}\int_{|z|<M}h_n^\xi(x,z)f(x+n^{-1/d}z)\,\mathrm{d}z=f(x)\int_{|z|<M}h_\infty^\xi(x,z)\,\mathrm{d}z\,.$$

It this regard notice that

$$\begin{aligned}&\int_{|z|<M}|h_n^\xi(x,z)f(x+n^{-1/d}z)-h_\infty^\xi(x,z)f(x)|\,\mathrm{d}z\\ &\le\int_{|z|<M}C_h\times|f(x+n^{-1/d}z)-f(x)|+|h_n^\xi(x,z)-h_\infty^\xi(x,z)|\times\|f\|_\infty\,\mathrm{d}z\\ &\le C_hn\int_{|z|<n^{-1/d}M}|f(x+z)-f(x)|\,\mathrm{d}z+\|f\|_\infty\times\int_{|z|<M}|h_n^\xi(x,z)-h_\infty^\xi(x,z)|\,\mathrm{d}z\,.\end{aligned}$$

Both terms converge to 0 as $n\to\infty$: the first since x is a Lebesgue point of x, the second by the dominated convergence of $h_n^\xi(x,z)$; cf. Lemma 4.1. Note that

$\int_{W_1}\int_{\mathbb{R}^2}|h^\xi_\infty(x,z)|\,\mathrm{d}z\mathrm{d}x<\infty$, which follows again from the clustering of the second mixed moment (1.17). Thus letting M go to infinity in $\int_{W_1} f^2(x)\int_{|z|<M} h^\xi_\infty(x,z)\,\mathrm{d}z\mathrm{d}x$ one completes the proof of variance asymptotics. $\square$

4.2 Proof of Theorem 1.14.

The proof is inspired by the proofs of [43, Propositions 1 and 2]. By the refined Campbell theorem and stationarity of $\mathcal{P}$, we have

$$\begin{aligned} n^{-1}\mathrm{Var}\widehat{H}^\xi_n(\mathcal{P}) &= \int_{W_n}\mathbb{E}_x\xi^2(x;\mathcal{P})\rho^{(1)}(x)\mathrm{d}x+\int_{W_n}\int_{W_n}[m_{(2)}(x,y)-m_{(1)}(x)m_{(1)}(y)]\mathrm{d}y\mathrm{d}x\\ &= \mathbb{E}_{\mathbf{0}}\xi^2(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})+n^{-1}\int_{W_n}\int_{W_n}(m_{(2)}(x,y)-m_{(1)}(x)m_{(1)}(y))\mathrm{d}y\mathrm{d}x. \end{aligned}\tag{4.9}$$

Now we write $c(x,y):=m_{(2)}(x,y)-m_{(1)}(x)m_{(1)}(y)$. The double integral in (4.9) becomes ($z=y-x$)

$$\begin{aligned} n^{-1}\int_{W_n}\int_{W_n}(m_{(2)}(x,y)-m_{(1)}(x)m_{(1)}(y))\mathrm{d}y\mathrm{d}x &= n^{-1}\int_{W_n}\int_{\mathbb{R}^d}c(\mathbf{0},z)\mathbf{1}[x+z\in W_n]\mathrm{d}z\mathrm{d}x\\ &= n^{-1}\int_{W_n}\int_{\mathbb{R}^d}c(\mathbf{0},z)\mathbf{1}[x\in W_n-z]\mathrm{d}z\mathrm{d}x. \end{aligned}$$

Write $\mathbf{1}[x\in W_n-z]$ as $1-\mathbf{1}[x\in(W_n-z)^c]$ to obtain

$$n^{-1}\int_{W_n}\int_{W_n}(m_{(2)}(x,y)-m_{(1)}(x)m_{(1)}(y))\mathrm{d}y\mathrm{d}x$$

$$=\int_{\mathbb{R}^d}c(\mathbf{0},z)\mathrm{d}z-n^{-1}\int_{\mathbb{R}^d}\int_{W_n}c(\mathbf{0},z)\mathbf{1}[x\in\mathbb{R}^d\setminus(W_n-z)]\mathrm{d}x\mathrm{d}z.$$

From (1.25), we have that $\gamma_{W_n}(z):=\mathrm{Vol}_d(W_n\cap(\mathbb{R}^d\setminus(W_n-z)))$ and thus rewrite (4.9) as

$$n^{-1}\mathrm{Var}\widehat{H}^\xi_n(\mathcal{P})=\mathbb{E}_{\mathbf{0}}\xi^2(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})+\int_{\mathbb{R}^d}c(\mathbf{0},z)\mathrm{d}z-n^{-1}\int_{\mathbb{R}^d}c(\mathbf{0},z)\gamma_{W_n}(z)\mathrm{d}z. \tag{4.10}$$

Now we claim that

$$\lim_{n\to\infty}n^{-1}\int_{\mathbb{R}^d}c(\mathbf{0},z)\gamma_{W_n}(z)\mathrm{d}z=0.$$

Indeed, as noted in Lemma 1 of [43], for all $z\in\mathbb{R}^d$ we have $\lim_{n\to\infty}n^{-1}\gamma_{W_n}(z)=0$. Since $n^{-1}c(\mathbf{0},z)\gamma_{W_n}(z)$ is dominated by the fast decreasing function $c(\mathbf{0},z)$, the dominated convergence theorem gives the claimed limit. Letting $n\to\infty$ in (4.10) gives

$$\lim_{n\to\infty}n^{-1}\mathrm{Var}\widehat{H}^\xi_n(\mathcal{P})=\mathbb{E}_{\mathbf{0}}\xi^2(\mathbf{0},\mathcal{P})\rho^{(1)}(\mathbf{0})+\int_{\mathbb{R}^d}c(\mathbf{0},z)\mathrm{d}z=\sigma^2(\xi), \tag{4.11}$$

where the last equality follows by the definition of $\sigma^2(\xi)$ in (1.18) and the finiteness follows by the fast decreasing property of $c(\mathbf{0}, z, \mathcal{P})$ (which follows from the assumption of strong clustering of mixed moments).

Now if $\sigma^2(\xi) = 0$ then the right hand side of (4.11) vanishes, i.e.,

$$\mathbb{E}_{\mathbf{0}}\xi^2(\mathbf{0}, \mathcal{P})\rho^{(1)}(\mathbf{0}) + \int_{\mathbb{R}^d} c(\mathbf{0}, z)\mathrm{d}z = 0.$$

Applying this identity to the right hand side of (4.10), then multiplying (4.10) by $n^{1/d}$ and taking limits we obtain

$$\lim_{n\to\infty} n^{-(d-1)/d}\mathrm{Var}\widehat{H}_n^{\xi}(\mathcal{P}) = -\lim_{n\to\infty} n^{-(d-1)/d}\int_{\mathbb{R}^d} c(\mathbf{0}, z)\gamma_{W_n}(z)\mathrm{d}z. \tag{4.12}$$

As in [43], we have $n^{-(d-1)/d}\gamma_{W_n}(z) \le C|z|$, and therefore again, by the fast decreasing property of $c(\mathbf{0}, z)$ we conclude that $n^{-(d-1)/d}c(\mathbf{0}, z)\gamma_{W_n}(z)$ is dominated by an integrable function of z. Also, as in [43, Lemma 1], for all $z \in \mathbb{R}^d$ we have $\lim_{n\to\infty} n^{-(d-1)/d}\gamma_{W_n}(z) = \gamma(z)$. The dominated convergence theorem yields (1.27) as desired,

$$\lim_{n\to\infty} n^{-(d-1)/d}\mathrm{Var}\widehat{H}_n^{\xi}(\mathcal{P}) = -\int_{\mathbb{R}^d} c(\mathbf{0}, z)\gamma(z)\mathrm{d}z. \quad \square$$

4.3 First proof of the central limit theorem

4.3.1 The method of cumulants

We use the method of cumulants to prove Theorem 1.12. Recall that we write $\langle f, \mu\rangle$ for $\int f d\mu$. The guiding principle is that as soon as the kth order cumulants C_n^k for $(\mathrm{Var}\langle f, \mu_n^{\xi}\rangle)^{-1/2}\langle f, \overline{\mu}_n\rangle$ vanish as $n \to \infty$ for k large, then

$$(\mathrm{Var}\langle f, \mu_n^{\xi}\rangle)^{-1/2}\langle f, \overline{\mu}_n\rangle \xrightarrow{\mathcal{D}} N(0, 1). \tag{4.13}$$

We establish the vanishing of C_n^k for k large by showing that the kth order cumulant for $\langle f, \overline{\mu}_n\rangle$ is of order $O(n)$, $k \ge 2$, and then use the assumption $\mathrm{Var}\langle f, \mu_n^{\xi}\rangle = \Omega(n^{\nu})$.

Our approach. The $O(n)$ growth of the kth order cumulant for $\langle f, \overline{\mu}_n\rangle$ is established by controlling the growth of cumulant measures for μ_n, which are defined analogously to moment measures. We first prove a general result (see (4.18) and (4.19) below) showing that integrals of cumulant measures for μ_n^{ξ} can be controlled by a finite sum of integrals of so-called (S, T) semi-cluster measures, where (S, T) is a generic partition of $\{1, ..., k\}$. This result holds for any μ_n^{ξ} of the form (1.3) and depends neither on choice of input $\mathcal{P}$ nor on the localization properties of ξ. Semi-cluster measures for μ_n^{ξ} have the appealing property that they involve differences of measures on product spaces with product measures, and thus their Radon-Nikodym derivatives involve differences of mixed moment functions.

In general, bounds on cumulant measures in terms of semi-cluster measures are not terribly informative. However, when ξ satisfies moment bounds and strong clustering (1.17), then the situation changes. First, integrals of (S,T) semi-cluster measures on properly chosen subsets $W(S,T)$ of W_n^k, with (S,T) ranging over partitions of $\{1,...,k\}$, exhibit $O(n)$ growth. This is because the subsets $W(S,T)$ are chosen so that the Radon-Nikodym derivative of the (S,T) semi-cluster measure, being a difference of mixed moment functions, may be controlled by the strong clustering bound (1.17) for points $(v_1,...,v_k) \in W(S,T)$. Second, it conveniently happens that W_n^k is precisely the union of $W(S,T)$, as (S,T) ranges over partitions of $\{1,...,k\}$. Therefore, combining these observations, we see that every cumulant measure on W_n^k is a sum ranging over partitions (S,T) of $\{1,...,k\}$ of linear combinations of (S,T) semi-cluster measures on $W(S,T)$, each of which exhibits $O(n)$ growth.

Thus cumulant measures exhibit growth *proportional to the volume of the window* W_n carrying $\mathcal{P}_n$, namely

$$\langle f^k, c_n^k \rangle = O(n), \quad f \in \mathcal{B}(W_1), \tag{4.14}$$

The remainder of Section 4.3 provides the details justifying (4.14).

Remarks on related work. (a) The estimate (4.14) first appeared in [7, Lemma 5.3], but the work of [21] (and to some extent [72]) was the first to rigorously control the growth of c_n^k on the diagonal subspaces, where two or more coordinates coincide. In fact Section 3 of [21] shows the estimate $\langle f^k, c_n^k \rangle \leq L^k (k!)^\beta n$, where L and β are constants independent of n and k. We assert that the clustering and cumulant arguments behind (4.14) are not restricted to Poisson input, but depend only on clustering (1.17) and moment bounds (1.16). Since these arguments are not well known we present them in a way which is hopefully accessible, reasonably self-contained, and rigorous. Since we do not care about the constants in (4.14), we shall suitably adopt the arguments of [7, Lemma 5.3] and [72], taking the the opportunity to make those arguments more rigorous. Indeed those arguments did not adequately explain clustering on diagonal subspaces.

(b) The breakthrough paper [51] shows that the kth order cumulant for the *linear statistic* $(\mathrm{Var}\langle f, \sum_x \delta_{n^{-1/d}x} \rangle)^{-1/2} \langle f, \sum_x \delta_{n^{-1/d}x} \rangle$ vanishes as $n \to \infty$ and k large. This approach is extended to linear statistics of random measures μ_n^ξ in Section 4.4 thereby giving a second proof of the central limit theorem.

4.3.2 Properties of cumulant and semi-cluster measures

Moments and cumulants. For a random variable Y with all finite moments, expanding the logarithm of the Laplace transform (in the negative domain) in a formal power series gives

$$\log \mathbb{E}(e^{tY}) = \log\big(1 + \sum_{k=1}^{\infty} \frac{M_k t^k}{k!}\big) = \sum_{k=1}^{\infty} \frac{S_k t^k}{k!}\,, \tag{4.15}$$

where $M_k = \mathbb{E}(Y^k)$ is the k th moment of Y and $S_k = S_k(Y)$ denotes the k th cumulant of Y. Both series in (4.15) can be considered as formal ones and no additional condition (on exponential moments of Y) are required for the cumulants to exist. Explicit relations between cumulants and moments can established by formal manipulations of these series, see e.g. [17, Lemma 5.2.VI]. In particular

$$S_k = \sum_{\gamma \in \Pi[k]} (-1)^{|\gamma|-1} (|\gamma| - 1)! \prod_{i=1}^{|\gamma|} M^{|\gamma(i)|}\,, \tag{4.16}$$

where $\Pi[k]$ is the set of all unordered partitions of the set $\{1, ..., k\}$, and for a partition $\gamma = \{\gamma(1), \ldots, \gamma(l)\} \in \Pi[k]$, $|\gamma| = l$ denotes the number of its elements, while $|\gamma(i)|$ the number of elements of subset $\gamma(i)$. (Although elements of $\Pi[k]$ are unordered partitions, we need to adopt some convention for the labeling of their elements: let $\gamma(1), \ldots, \gamma(l)$ correspond to the ordering of the smallest elements in the partition sets.) In view of (4.16) the existence of the kth cumulant S_k follows from the finiteness of the moment M_k.

Moment measures. Given a random measure μ on $\mathbb{R}^d$, the k-th moment measure $M^k = M^k(\mu)$ is the one (Sect 5.4 and Sect 9.5 of [17]) satisfying

$$\langle f_1 \otimes ... \otimes f_k, M^k(\mu)\rangle = \mathbb{E}[\langle f_1, \mu\rangle ... \langle f_k, \mu\rangle] = \mathbb{E}[\sum_{x \in \mathcal{P}_n} f_1(\frac{x}{n^{1/d}})\xi(x, \mathcal{P}_n) \cdots \sum_{x \in \mathcal{P}_n} f_k(\frac{x}{n^{1/d}})\xi(x, \mathcal{P}_n)]$$

for all $f_1, ..., f_k \in \mathbb{B}(\mathbb{R}^d)$, where $f_1 \otimes ... \otimes f_k : (\mathbb{R}^d)^k \to \mathbb{R}$ is given by $f_1 \otimes ... \otimes f_k(x_1, ..., x_k) = f_1(x_1)...f_k(x_k)$.

As on p. 143 of [17], when μ is a counting measure, M^k may be expressed as a sum of factorial moment measures $M_{[j]}, 1 \le j \le k$, (as defined on p. 133 of [17]):

$$M^k(\mathrm{d}(x_1 \times \cdots \times x_k)) = \sum_{j=1}^{k} \sum_{\mathcal{V}} M_{[j]}(\Pi_{i=1}^{j} \mathrm{d}y_i(\mathcal{V}))\delta(\mathcal{V}),$$

where, to quote from [17], the inner sum is taken over all partitions $\mathcal{V}$ of the k coordinates into j non empty disjoint subsets, the $y_i(\mathcal{V}), 1 \le i \le j$, constitute an arbitrary selection of one coordinate from each subset, and $\delta(\mathcal{V})$ is a δ function which equals zero unless equality holds among the coordinates in each non-empty subset of $\mathcal{V}$.

When μ is the atomic measure μ_n^ξ, we write M_n^k for $M^k(\mu_n^\xi)$. By the Campbell formula, considering repetitions in the k-fold product of $\mathbb{R}^d$, and putting $\tilde{y}_i := y_i(\mathcal{V})$ and $\mathcal{V} := (\mathcal{V}_1, ..., \mathcal{V}_j)$ we have that

$$\langle f \otimes ... \otimes f, M_n^k \rangle = \mathbb{E}[\langle f, \mu_n^\xi \rangle ... \langle f, \mu_n^\xi \rangle]$$

$$= \sum_{j=1}^k \sum_{\mathcal{V}} \int_{(W_n)^j} \Pi_{i=1}^k f(\frac{y_i}{n^{1/d}}) \mathbb{E}_{\tilde{y}_1 ... \tilde{y}_j}[\Pi_{i=1}^j \xi^{|\mathcal{V}_i|}(\tilde{y}_i, \mathcal{P}_n)] \rho^{(j)}(\tilde{y}_1, ..., \tilde{y}_j) \Pi_{i=1}^j \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V}).$$

In other words, recalling Lemma 9.5.IV of [17] we get

$$dM_n^k(y_1, ..., y_k) = \sum_{j=1}^k \sum_{\mathcal{V}} m^{(|\mathcal{V}_1|, ..., |\mathcal{V}_j|)}(\tilde{y}_1, ..., \tilde{y}_j; n) \Pi_{i=1}^j \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V}). \tag{4.17}$$

Cumulant measures. The kth cumulant measure $c_n^k := c^k(\mu_n)$ is defined analogously to the kth moment measure via

$$\langle f_1 \otimes ... \otimes f_k, c^k(\mu_n) \rangle = c(\langle f_1, \mu_n \rangle ... \langle f_k, \mu_n \rangle)$$

where $c(X_1, ..., X_k)$ denotes the mixed cumulant of the random variables $X_1, ..., X_k$.

The existence of the cumulant measures c_n^l, $l = 1, 2, ...$ follows from the existence of moment measures in view of the representation (4.16). Thus, we have the following representation for cumulant measures :

$$c_n^l = \sum_{T_1, ..., T_p} (-1)^{p-1}(p-1)! M_n^{T_1} \cdots M_n^{T_p},$$

where $T_1, ..., T_p$ ranges over all unordered partitions of the set $1, ..., l$ (see p. 30 of [42]). Henceforth for $T_i \subset \{1, ..., l\}$, let $M_n^{T_i}$ denote a copy of the moment measure $M^{|T_i|}$ on the product space W^{T_i}. Multiplication denotes the usual product of measures: For T_1, T_2 disjoint sets of integers and for measurable $B_1 \subset (\mathbb{R}^d)^{T_1}, B_2 \subset (\mathbb{R}^d)^{T_2}$ we have $M_n^{T_1} M_n^{T_2}(B_1 \times B_2) = M_n^{T_1}(B_1) M_n^{T_2}(B_2)$. The first cumulant measure coincides with the expectation measure and the second cumulant measure coincides with the covariance measure.

Cluster and semi-cluster measures. We show that every cumulant measure c_n^k is a linear combination of products of moment and cluster measures. We first recall the definition of cluster and semi-cluster measures. A cluster measure $U_n^{S,T}$ on $W_n^S \times W_n^T$

for non-empty $S, T \subset \{1, 2, ...\}$ is defined by

$$U_n^{S,T}(B \times D) = M_n^{S \cup T}(B \times D) - M_n^S(B) M_n^T(D)$$

for Borel sets B and D in W_n^S and W_n^T, respectively, and where multiplication means product measure.

Let S_1, S_2 be a partition of S and let T_1, T_2 be a partition of T. A product of a cluster measure $U_n^{S_1,T_1}$ on $W_n^{S_1} \times W_n^{T_1}$ with products of moment measures $M_n^{|S_2|}$ and $M_n^{|T_2|}$ on $W_n^{S_2} \times W_n^{T_2}$ is an (S, T) *semi-cluster measure*.

For each non-trivial partition (S, T) of $\{1, ..., k\}$ the k-th cumulant c_n^k measure is represented as

$$c_n^k = \sum_{(S_1,T_1),(S_2,T_2)} \alpha((S_1, T_1), (S_2, T_2)) U_n^{S_1,T_1} M_n^{|S_2|} M_n^{|T_2|}, \tag{4.18}$$

where the sum ranges over partitions of $\{1, ..., k\}$ consisting of pairings (S_1, T_1), (S_2, T_2), where $S_1, S_2 \subset S$ and $T_1, T_2 \subset T$, where S_1 *and* T_1 *are non-empty*, and where $\alpha((S_1, T_1), (S_2, T_2))$ are integer valued pre-factors. In other words, for any non-trivial partition (S, T) of $\{1, ..., k\}$, c_n^k is a linear combination of (S, T) semi-cluster measures. We prove this exactly as in the proof of Lemma 5.1 of [7], as that proof involves only combinatorics and does not depend on the nature of the input. For an alternate proof, with good growth bounds on the integer pre-factors $\alpha((S_1, T_1), (S_2, T_2))$, we refer to Lemma 3.2 of [21].

Let $\Xi(k)$ be the collection of partitions of $\{1, ..., k\}$ into two subsets S and T. If W_n^k may be expressed as the union of sets $W(S, T)$, $(S, T) \in \Xi(k)$, then

$$|\langle f^k, c_n^k \rangle| \le \sum_{(S,T) \in \Xi(k)} \int_{W(S,T)} |f(v_1)...f(v_k)||dc_n^k(v_1, ..., v_k)| \tag{4.19}$$

$$\le ||f||_\infty^k \sum_{(S,T) \in \Xi(k)} \sum_{(S_1,T_1),(S_2,T_2)} \alpha((S_1, T_1), (S_2, T_2)) \int_{W(S,T)} d(U_n^{S_1,T_1} M_n^{|S_2|} M_n^{|T_2|})(v_1, ..., v_k),$$

where the last inequality follows by (4.18). As noted at the outset, this bound is valid for any $f \in \mathbb{B}(\mathbb{R}^d)$ and any measure μ_n^ξ of the form (1.3).

We now specify the collection of sets $W(S, T)$, $(S, T) \in \Xi(k)$, to be used in all that follows. Given $v := (v_1, ..., v_k) \in W_n^k$, let

$$D_k(v) := D_k(v_1, ..., v_k) := \max_{i \le k} (||v_1 - v_i|| + ... + ||v_k - v_i||)$$

be the l^1 diameter for v. For all such partitions consider the subset $W(S, T)$ of $W_n^S \times W_n^T$ having the property that $v \in W(S, T)$ implies $d(v^S, v^T) \ge D_k(v)/k^2$, where v^S and v^T are the projections of v onto W_n^S and W_n^T, respectively, and where $d(v^S, v^T)$ is the minimal Euclidean distance between pairs of points from v^S and v^T.

It is easy to see that for every $v := (v_1, ..., v_k) \in W_n^k$, there is a partition (S, T) of $\{1, ..., k\}$ such that $d(v^S, v^T) \geq D_k(v)/k^2$. If this were not the case then given $v := (v_1, ..., v_k)$, the distance between any two components of v must be strictly less than $D_k(v)/k^2$ and we would get $\max_{i \leq k} \sum_{j=1}^k ||v_i - v_j|| \leq (k-1)kD_k/k^2$.

This would contradict the definition of $D_k(v)$. Thus W_n^k is the union of sets $W(S,T)$, $(S,T) \in \Xi(k)$, as desired. We next describe the behavior of the differential $d(U_n^{S_1,T_1} M_n^{|S_2|} M_n^{|T_2|})$ on $W(S,T)$.

Semi-cluster measures on $W(S,T)$. Next, given $S_1 \subset S$ and $T_1 \subset T$, notice that $d(v^{S_1}, v^{T_1}) \geq d(v^S, v^T)$ where v^{S_1} denotes the projection of v^S onto $W_n^{S_1}$ and v^{T_1} denotes the projection of v^T onto $W_n^{T_1}$. Let $\Pi(S_1, T_1)$ be the partitions of S_1 into j_1 sets $\mathcal{V}_1, ..., \mathcal{V}_{j_1}$, with $1 \leq j_1 \leq |S_1|$, and the partitions of T_1 into j_2 sets $\mathcal{V}_{j_1+1}, ..., \mathcal{V}_{j_1+j_2}$, with $1 \leq j_2 \leq |T_1|$. Thus an element of $\Pi(S_1, T_1)$ is a partition of $S_1 \cup T_1$.

If a partition $\mathcal{V}$ of $S_1 \cup T_1$ does not belong to $\Pi(S_1, T_1)$, then there is a partition element of $\mathcal{V}$ containing points in S_1 and T_1 and so $\delta(\mathcal{V}) = 0$ on the set $W(S,T)$. Thus we make the crucial observation that, *on the subset $W(S,T)$ of W_n^k the differential $d(M_n^{S_1 \cup T_1})$ collapses into a sum over partitions in $\Pi(S_1, T_1)$.* Thus $d(M_n^{S_1 \cup T_1})$ and $d(M_n^{S_1} M_n^{T_1})$ both involve sums of measures on common diagonal subspaces, made precise as follows.

Lemma 4.2. *On the set $W(S,T)$ we have*

$$d(U_n^{S_1,T_1}) = \sum_{j_1=1}^{|S_1|} \sum_{j_2=1}^{|T_1|} \sum_{\mathcal{V} \in \Pi(S_1,T_1)} [....] \Pi_{i=1}^{j_1+j_2} \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V}) \tag{4.20}$$

where

$$[....] := m^{(|\mathcal{V}_1|, ..., |\mathcal{V}_{j_1}|, |\mathcal{V}_{j_1+1}|, ..., |\mathcal{V}_{j_1+j_2}|)}(\tilde{y}_1, ..., \tilde{y}_{j_1} \tilde{y}_{j_1+1}, ...\tilde{y}_{j_1+j_2}; n)$$
$$- m^{(|\mathcal{V}_1|, ..., |\mathcal{V}_{j_1}|)}(\tilde{y}_1, ..., \tilde{y}_{j_1}; n) m^{(|\mathcal{V}_{j_1+1}|, ..., |\mathcal{V}_{j_1+j_2}|)}(\tilde{y}_{j_1+1}, ...\tilde{y}_{j_1+j_2}; n).$$

The representation of $M_n^{|S_2|}$ and $M_n^{|T_2|}$ follows from (4.17), that is to say

$$dM_n^{|S_2|} = \sum_{j_3=1}^{|S_2|} \sum_{\mathcal{V} \in \Pi(S_2)} m^{(|\mathcal{V}_1|, ..., |\mathcal{V}_{j_3}|)}(\tilde{y}_1, ..., \tilde{y}_{j_3}; n) \Pi_{i=1}^{j_3} \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V}), \tag{4.21}$$

where $\Pi(S_2)$ runs over partitions of S_2 into j_3 sets, $1 \leq j_3 \leq |S_2|$. Similarly

$$dM_n^{|T_2|} = \sum_{j_4=1}^{|T_2|} \sum_{\mathcal{V} \in \Pi(T_2)} m^{(|\mathcal{V}_1|, ..., |\mathcal{V}_{j_4}|)}(\tilde{y}_1, ..., \tilde{y}_{j_4}; n) \Pi_{i=1}^{j_4} \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V}), \tag{4.22}$$

where $\Pi(T_2)$ runs over partitions of T_2 into j_4 sets, $1 \leq j_4 \leq |T_2|$.

4.3.3 Strong clustering and semi-cluster measures

The previous section established properties of semi-cluster and cumulant measures valid for any μ_n^ξ of the form (1.3). When ξ satisfies strong clustering (1.17) and moment bounds, we now assert that each integral in (4.19) is $O(n)$.

Lemma 4.3. *Assume ξ satisfies strong clustering (1.17) and moment bounds for all $p \geq 1$. For each partition element (S, T) of $\Xi(k)$ we have*

$$\int_{W(S,T) \subset W_n^S \times W_n^T} |d(U_n^{S_1,T_1} M_n^{|S_2|} M_n^{|T_2|})| = O(n). \tag{4.23}$$

Proof. The differential $d(U_n^{S_1,T_1} M_n^{|S_2|} M_n^{|T_2|})$ is a sum

$$\sum_{j_1=1}^{|S_1|} \sum_{j_2=1}^{|T_1|} \sum_{j_3=1}^{|S_2|} \sum_{j_4=1}^{|T_2|} [...][...][...]$$

of products of three factors, one factor coming from each of the summands in (4.20)-(4.22). By Theorem 1.10, on the set $W(S,T)$ the factor arising from (4.20) is bounded in absolute value by

$$\tilde{C}_k \tilde{\phi}(\tilde{c}_k D_k(y)/k^2).$$

By the moment bound (1.16) the two remaining factors arising from summands in (4.21)- (4.22) are bounded by a constant depending only on k, say $M'(k)$.

Thus we have

$$\begin{aligned} \int_{W(S,T)} |d(U_n^{S_1,T_1} M_n^{|S_2|} M_n^{|T_2|})| &\leq \tilde{C}_k (M'(k))^2 \sum_{j=1}^{k} \sum_{\mathcal{V}} \int_{W(S,T)} \tilde{\phi}(\tilde{c}_k D_k(y)/k^2) \Pi_{i=1}^j \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V})|| \\ &\leq \tilde{C}_k (M'(k))^2 \sum_{j=1}^{k} \sum_{\mathcal{V}} \int_{(W_n)^j} \tilde{\phi}(\tilde{c}_k D_k(y)/k^2) \Pi_{i=1}^j \mathrm{d}y_i(\mathcal{V}) \delta(\mathcal{V}). \end{aligned}$$

Here $\mathcal{V}$ runs over all partitions of the k coordinates into j non-empty disjoint subsets. We assert that all summands are $O(n)$. We show this when $j = k$, as the proof for the summands $j \in \{1, ..., k-1\}$ is the same. We bound

$$\begin{aligned} &\int_{y_1 \in W_n} ... \int_{y_k \in W_n} \tilde{\phi}(\tilde{c}_k D_k(y)/k^2) \mathrm{d}y_1 ... \mathrm{d}y_k \\ &= \int_{y_1 \in W_n} \int_{w_2 \in W_n - y_1} ... \int_{w_k \in W_n - y_1} \tilde{\phi}(\tilde{c}_k D_k(0, w_2, ..., w_k)/k^2) \mathrm{d}y_1 \mathrm{d}w_2 ... \mathrm{d}w_k. \end{aligned}$$

Now $D_k(0, w_2, ..., w_k) \geq \sum_{i=2}^k ||w_i||$ and so letting $\tilde{c}'_k := \tilde{c}_k/k^2$ gives

$$\begin{aligned} &\int_{y_1 \in W_n} \cdots \int_{y_k \in W_n} \tilde{\phi}(\tilde{c}_k D_k(y)/k^2) \mathrm{d}y_1 \ldots \mathrm{d}y_k \\ &\leq n \int_{w_2 \in \mathbb{R}^d} \cdots \int_{w_k \in \mathbb{R}^d} \tilde{\phi}(\tilde{c}'_k \sum_{i=2}^k ||w_i||) \mathrm{d}w_2 \ldots \mathrm{d}w_k \\ &\leq n \left(\int_{\mathbb{R}^d} (\tilde{\phi}(\tilde{c}'_k ||w||))^{1/k} \mathrm{d}w \right)^{k-1} = O(n), \end{aligned}$$

where the last inequality follows since $\tilde{\phi}$ decreasing implies

$$\tilde{\phi}(\tilde{c}'_k \sum_{i=2}^k ||w_i||) \leq (\Pi_{i=2}^k \tilde{\phi}(\tilde{c}'_k ||w_i||))^{1/k}.$$

We bound the other summands for $j \in \{1, ..., k-1\}$ in a similar manner, completing the proof of Lemma 4.3. □

4.3.4 Proof of Theorem 1.12

By the bound (4.19) and Lemma 4.3 we obtain (4.14). Letting C_n^k be the kth cumulant for $(\mathrm{Var}\langle f, \mu_n^\xi \rangle)^{-1/2} \langle f, \mu_n^\xi \rangle$, we obtain $C_n^1 = 0, C_n^2 = 1$, and for all $k = 3, 4, ...$

$$C_n^k = O(n(\mathrm{Var}\langle f, \mu_n^\xi \rangle)^{-k/2}).$$

Since $\mathrm{Var}\langle f, \mu_n^\xi \rangle = \Omega(n^\nu)$ by assumption, it follows that if $k \in (2/\nu, \infty)$, then the kth cumulant tends to zero as $n \to \infty$. By a classical result of Marcinkiewicz (see e.g. [68, Lemma 3]), we get that all cumulants $C_n^k,\ k \geq 3$, converge to zero as $n \to \infty$. This gives (4.13) as desired and completes the proof of Theorem 1.12. □

4.4 Second proof of the central limit theorem

We shall now give a second proof of the central limit theorem which we believe is of independent interest. Even though this proof is also based on the cumulant method as outlined in Section 4.3.1, we shall bound the cumulants using different ideas as indicated in Remark (b) in Section 4.3.1. Though much of this proof can be read independently of the proof in Section 4.3, we repeatedly use the definition of moments and cumulants from Section 4.3.2.

Our approach. We shall adapt the approach in [51, Sec. 4] replacing $\mathcal{P}_{GEF}$ by μ_n^ξ, which is a purely atomic measure, and considering its linear statistic $\mu_n^\xi(f)$. Our mixed moment functions play the same role as the k-point correlation functions of the point process, in the sense that they are densities of the moment measures of μ_n^ξ when arguments are all distinct. Some care is required to take properly into account repeated arguments, when these functions no longer simply 'collapse' to appropriate lower dimensional ones, but change their structure due to corresponding Palm conditioning of the mass attached to the repeated points. This is captured by our generalized mixed moments at (1.5).

4.4.1 Ursell functions of a purely atomic measure

Recall the definition of the generalized mixed moment functions at (1.5):

$$m^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n) := \mathbb{E}_{x_1,\dots,x_p}\big((\xi(x_1,\mathcal{P}_n))^{k_1}\dots(\xi(x_p,\mathcal{P}_n))^{k_p}\big)\rho^{(p)}(x_1,\dots,x_p).$$

We will drop dependence on n, i.e., $m^{(k_1,\dots,k_p)}(x_1,\dots,x_p;n) = m^{(k_1,\dots,k_p)}(x_1,\dots,x_p)$ unless asymptotics in n is considered.

Inspired by the approach in [8, Section 2] we now introduce truncated mixed moment functions. Define *truncated mixed moment (Ursell) functions* $m_\top^{(k_1,\dots,k_p)}$ by taking $m_\top^{(k)}(x) := m^{(k)}(x)$ for all $k \in \mathbb{N}$ and inductively

$$m_\top^{(k_1,\dots,k_p)}(x_1,\dots,x_p) := m^{(k_1,\dots,k_p)}(x_1,\dots,x_p) - \sum_{\substack{\gamma\in\Pi[p]\\|\gamma|>1}}\prod_{i=1}^{|\gamma|} m_\top^{(k_j:j\in\gamma(i))}(x_j : j\in\gamma(i))\,. \tag{4.24}$$

for distinct $x_1,\dots,x_p \in W_n$ and all integers $k_1,\dots,k_p$, $p \geq 1$, and (implicitly) $n \leq \infty$. It is straightforward to prove that these functions satisfy the following relations. They extend the known relations for point processes, where $m^{(k_1,\dots,k_p)}(x_1,\dots,x_p) = \rho^{(p)}(x_1,\dots,x_p)$ depend only on p, but we were unable to find them in the literature for purely atomic random measure. Assuming $1 \in \gamma(1)$ in (4.24) and summing over partitions of $\{1,\dots,p\}\setminus\gamma(1)$, we get the following relation :

$$\begin{aligned} m^{(k_1,\dots,k_p)}(x_1,\dots,x_p) = m_\top^{(k_1,\dots,k_p)}&(x_1,\dots,x_p) + \\ &\sum_{\substack{I\subsetneq\{1,\dots,p\}\\1\in I}} m_\top^{(k_j:j\in I)}(x_j : j\in I)\, m^{(k_j:j\in I^c)}(x_j : j\in I^c)\,, \end{aligned} \tag{4.25}$$

where $I^c := \{1,\dots,p\}\setminus I$. Using (4.25), by induction with respect to p, one obtains the direct relation to the mixed moment functions

$$m_\top^{(k_1,\dots,k_p)}(x_1,\dots,x_p) = \sum_{\gamma\in\Pi[p]}(-1)^{|\gamma|-1}(|\gamma|-1)!\prod_{i=1}^{|\gamma|} m^{(k_j:j\in\gamma(i))}(x_j : j\in\gamma(i))\,. \tag{4.26}$$

This extends the relation [51, (27)], valid for point processes. We say that a partition $\gamma = \{\gamma(1),\dots,\gamma(l)\} \in \Pi(p)$ *refines* partition $\sigma = \{\sigma(1),\dots,\sigma(l_1)\} \in \Pi(p)$ if for all

$i \in \{1,\ldots,l\}$, $\gamma(i) \subset \sigma(j)$ for some $j \in \{1,\ldots,l_1\}$. Otherwise, the partition γ is said to *mix* partition σ. Now using (4.24), we get for any $I \subsetneqq \{1,\ldots,p\}$

$$m^{(k_j:j\in I)}(x_j : j \in I)\, m^{(k_j:j\in I^c)}(x_j : j \in I^c) = \sum_{\substack{\gamma\in\Pi[p] \\ \gamma \text{ refines } \{I,I^c\}}} \prod_{i=1}^{|\gamma|} m_\top^{(k_j:j\in\gamma(i))}(x_j : j \in \gamma(i)) \,, \tag{4.27}$$

and therefore, again in view of (4.24)

$$m_\top^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p) = m^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p) - m^{(k_j:j\in I)}(x_j : j \in I)\, m^{(k_j:j\in I^c)}(x_j : j \in I^c) + \sum_{\substack{\gamma\in\Pi[p], |\gamma|>1 \\ \gamma \text{ mixes } \{I,I^c\}}} \prod_{i=1}^{|\gamma|} m_\top^{(k_j:j\in\gamma(i))}(x_j : j \in \gamma(i)) \,. \tag{4.28}$$

This extends the relation [51, last displayed formula in the proof of Claim 4.1] valid for point processes.

4.4.2 Clustering and bounds for Ursell functions

We show now that clustering of the generalized mixed moments (1.17) implies some bounds on the Ursell functions. Since $m^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p;n)$ is invariant with respect to any joint permutation of its arguments $(k_1,\ldots,k_p)$ and $(x_1,\ldots,x_p)$, clustering (1.17) may be rephrased as follows : There exist a fast decreasing function $\tilde{\phi}$ and constants $\tilde{C}_k$, $\tilde{c}_k$, such that for any collection of positive integers $k_1,\ldots,k_p$, $p \geq 2$, satisfying $k_1+\ldots+k_p = k$, for any nonempty, proper subset $I \subsetneqq \{1,\ldots,p\}$, for all $n \leq \infty$ and all configurations $x_1,\ldots,x_p \in W_n$ of distinct points we have

$$\Big| m^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p;n) - m^{(k_j:j\in I)}(x_j : j \in I;n)\, m^{(k_j:j\in I^c)}(x_j : j \in I^c;n) \Big| \leq \tilde{C}_k \tilde{\phi}(\tilde{c}_k s) \,, \tag{4.29}$$

where $s := d\big(\{x_j : j \in I\}, \{x_j : j \in I^c\}\big)$.

Now we consider the bounds of Ursell functions of clustering measures. Following the idea of [51, Claim 4.1] one proves that clustering (1.17) and the p-moment condition (1.16) imply that there exists a fast decreasing function $\tilde{\phi}_\top$ and constants $\tilde{C}_k^\top$, $\tilde{c}_k^\top$, such that for any collection of positive integers $k_1,\ldots,k_p$, $p \geq 2$, satisfying $k_1+\ldots+k_p = k$, for all $n \leq \infty$ and all configurations $x_1,\ldots,x_p \in W_n$ of distinct points we have

$$|m_\top^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p;n)| \leq \tilde{C}_k^\top \tilde{\phi}_\top\big(c_k^\top \operatorname{diam}(x_1,\ldots,x_p)\big) \,, \tag{4.30}$$

where $\operatorname{diam}(x_1,\ldots,x_p) := \max_{i,j=1\ldots p}(|x_i - x_j|)$. The proof uses the semi-cluster representation (4.28), clustering (1.17), together with the fact that there exist constants $c_p^\top$ (depending on the dimension d) such that for each configuration $x_1,\ldots,x_p \in W_n$, there exists a partition $\{I, I^c\}$ of $\{1,\ldots,p\}$ such that $d(\{x_j : j \in I\}, \{x_j : j \in I^c\}) \geq$

$\tilde{c}_p^{\top}\operatorname{diam}(x_1,\dots,x_p)$.

Next, inequality (4.30) allows one to bound integrals

$$\sup_{n\le\infty}\sup_{x_1\in W_n}\sup_{\substack{k_1+\ldots+k_p=k\\ k_i>0}}\int_{(W_n)^{p-1}}|m_{\top}^{(k_1,\ldots,k_p)}(x_1,\ldots,x_p;n)|\,\mathrm{d}x_2\cdots\mathrm{d}x_p<\infty. \tag{4.31}$$

Indeed, for a fix point $x_1\in W_n$, we split $(W_n)^{p-1}$ into disjoint sets:

$$\begin{aligned} G_0 &:= \{(x_2,\ldots,x_p)\in (W_n)^{p-1} : \operatorname{diam}(x_1,\ldots,x_p)\le 1\}\\ G_l &:= \{(x_2,\ldots,x_p)\in (W_n)^{p-1} : 2^{l-1}<\operatorname{diam}(x_1,\ldots,x_p)\le 2^l\},\quad l\ge 1\end{aligned}$$

and use estimate (4.30) to bound the integral on the left-hand side of (4.31) by

$$\tilde{C}_k^{\top}+\tilde{C}_k^{\top}\sum_{l=1}^{\infty}2^{dl(k-1)}\tilde{\phi}_{\top}(\tilde{c}_k^{\top}2^{l-1})<\infty$$

since $\tilde{\phi}_{\top}$ is fast decreasing; cf. [51, Claim 4.2].

4.4.3 Proof of Theorem 1.12

The cumulant of order one is equal to the expectation and hence disappears for the considered (centered) random variable $\overline{\mu}_n^{\xi}(f)$. The cumulant of order 2 is equal to the variance and hence equal to 1 in our case. For $k\ge 2$, note the following relation between the normalized and the unnormalized cumulants :

$$S_k((\mathrm{Var}\,\overline{\mu}_n^{\xi}(f))^{-1/2}\mu_n^{\xi}(f)) = (\mathrm{Var}\,\mu_n^{\xi}(f))^{-k/2}\times S_k(\mu_n^{\xi}(f)). \tag{4.32}$$

We establish the vanishing of (4.32) for k large by showing that the kth order cumulant $S_k(\mu_n^{\xi}(f)$ is of order $O(n)$, $k\ge 2$, and then use assumption (1.22), i.e., $\mathrm{Var}\langle f,\mu_n^{\xi}\rangle=\Omega(n^{\nu})$. We have

$$M_n^k := \mathbb{E}(\langle f,\mu_n^{\xi}\rangle)^k = \mathbb{E}\Big(\sum_{x_i\in\mathcal{P}_n} f_n(x_i)\xi(x_i,\mathcal{P}_n)\Big)^k,$$

where $f_n(\cdot)=f(\cdot/n^{1/d})$. Considering appropriately the repetitions of points x_i in the k th product of the sum and using the Campbell's formula one obtains

$$M_n^k=\sum_{\sigma\in\Pi[k]}\langle\bigotimes_{i=1}^{|\sigma|}f_n^{|\sigma(i)|}m^{(\sigma)},\lambda_n^{|\sigma|}\rangle\,, \tag{4.33}$$

where $\bigotimes$ denotes the tensor product of functions

$$(\bigotimes_{i=1}^{p}f_n^{k_j})(x_1,\ldots,x_p)=\prod_{i=1}^{p}(f_n)^{k_j}(x_j)\,,$$

$$m^{(\sigma)}(x_1, \ldots, x_{|\sigma|}; n) := m^{(|\sigma(1)|, \ldots, |\sigma(|\sigma|)|)}(x_1, \ldots, x_{|\sigma|}; n)$$

and λ_n^l denotes the Lebesgue measure on $(W_n)^l$. Using the above representation and (4.16) the kth cumulant $S_k(\mu_n^\xi(f))$ can be expressed as follows

$$\begin{aligned} S_k(\mu_n^\xi(f)) &= \sum_{\gamma \in \Pi[k]} (-1)^{|\gamma|-1}(|\gamma|-1)! \sum_{\substack{\sigma \in \Pi[k] \\ \sigma \text{ refines } \gamma}} \prod_{i=1}^{|\gamma|} \langle \bigotimes_{j=1}^{|\gamma(i)/\sigma|} f_n^{(\gamma(i)/\sigma)(j)} m^{(\gamma(i)/\sigma)}, \lambda_n^{|\gamma(i)/\sigma|} \rangle \\ &= \sum_{\sigma \in \Pi[k]} \sum_{\substack{\gamma \in \Pi[k] \\ \sigma \text{ refines } \gamma}} (-1)^{|\gamma|-1}(|\gamma|-1)! \prod_{i=1}^{|\gamma|} \langle \bigotimes_{j=1}^{|\gamma(i)/\sigma|} f_n^{(\gamma(i)/\sigma)(j)} m^{(\gamma(i)/\sigma)}, \lambda_n^{|\gamma(i)/\sigma|} \rangle \,, \end{aligned} \tag{4.34}$$

where $\gamma(i)/\sigma$ is the partition of $\gamma(i)$ induced by σ. Note that for any partition $\sigma \in \Pi[k]$, with $|\sigma(j)| = k_j$, $j = 1, \ldots, |\sigma| = p$, the inner sum in (4.34) can be rewritten as follows:

$$\sum_{\gamma \in \Pi[p]} (-1)^{|\gamma|-1}(|\gamma|-1)! \prod_{i=1}^{|\gamma|} \langle \bigotimes_{j \in \gamma(i)} f_n^{k_j} m^{(k_j : j \in \gamma(i))}, \lambda_n^{|\gamma(i)|} \rangle = \langle \bigotimes_{j=1}^{p} f_n^{k_j} m_\top^{(k_1, \ldots, k_p)}, \lambda_n^p \rangle \,, \tag{4.35}$$

where the equality is due to (4.26). Consequently

$$S_k(\mu_n^\xi(f)) = \sum_{\sigma \in \Pi[k]} \langle \bigotimes_{j=1}^{|\sigma|} f_n^{|\sigma(j)|} m_\top^{(|\sigma(1)|, \ldots, |\sigma(|\sigma|)|)}, \lambda_n^{|\sigma|} \rangle \,, \tag{4.36}$$

which extends the relation [51, Claim 4.3] valid for point processes. This formula, which expresses the kth cumulant in terms of truncated mixed moment functions, is the counterpart to the standard formula (4.33) expressing kth moments in terms of correlation functions. Now, using (4.31) and denoting the supremum therein by $\hat{C}_k$, we have that

$$\begin{aligned} |\langle \bigotimes_{j=1}^{p} f_n^{k_j} m_\top^{(k_1, \ldots, k_p)}, \lambda_n^p \rangle \,| &\le \int_{W_n^p} |\bigotimes_{j=1}^{p} f_n^{k_j}| |m_\top^{(k_1, \ldots, k_p)}(x_1, \ldots, x_p)| \mathrm{d}x_1 \ldots \mathrm{d}x_p \\ &\le \|f\|_\infty^k \int_{W_n} \mathrm{d}x_1 \int_{W_n^{p-1}} |m_\top^{(k_1, \ldots, k_p)}(x_1, \ldots, x_p)| \mathrm{d}x_2 \ldots \mathrm{d}x_p \\ &\le \|f\|_\infty^k \hat{C}_k \mathrm{Vol}(W_n). \end{aligned}$$

So, the above bound along with (4.34) and(4.35) gives us that $S_k(\mu_n^\xi(f)) = O(n)$ for all $k \ge 2$. Thus, using the variance lower bound condition (1.22) and the relation (4.32), we get for large enough k, that $S_k((\mathrm{Var}\mu_n^\xi(f))^{-1/2}\mu_n^\xi(f)) \to 0$ as $n \to \infty$. Now, as discussed in (4.13), this suffices to guarantee normal convergence. $\square$

5 Appendix

5.1 Facts needed in the proof of clustering of mixed moments

The following facts regarding U-statistics are used in the proof of Lemma 3.2.

Lemma 5.1. *Let f, g be two real valued, symmetric functions defined on $(\mathbb{R}^d)^k$ and $(\mathbb{R}^d)^l$ respectively. Let $F := \frac{1}{k!}\sum_{\mathbf{x}\in\mathcal{X}^{(k)}} f(\mathbf{x})$ and $G := \frac{1}{l!}\sum_{\mathbf{x}'\in\mathcal{X}^{(l)}} g(\mathbf{x}')$ be the corresponding U-statistics of order k and l respectively, on the input $\mathcal{X} \subset \mathbb{R}^d$. Then we have:*
(i) The product $F\,G$ is a sum of U-statistics of order not greater than $k + l$.
(ii) Let $\mathcal{A}$ be a fixed, finite subset of $\mathbb{R}^d$. The statistic $F_{\mathcal{A}} := 1/k! \sum_{\mathbf{x}\in(\mathcal{X}\cup\mathcal{A})^{(k)}} f(\mathbf{x})$ is a sum of U-statistics of $\mathcal{X}$ of order not greater than k.

Proof. The first statement follows from the representation

$$F\,G = \sum_{m=\max(k,l)}^{k+l} \frac{1}{m!} \sum_{\mathbf{z}\in\mathcal{X}^{(m)}} h_m(\mathbf{z})\,,$$

where

$$h_m(z_1,\ldots,z_m) := \frac{1}{k!l!} \sum_{\pi\in S_m} f\big(z_{\pi(1)}, z_{\pi(2)},\ldots,z_{\pi(k)}\big) g\big(z_{\pi(m-l+1)}, z_{\pi(m-l+2)}\ldots,z_{\pi(m)}\big)\,,$$

with S_m denoting the permutation group of the first m integers. For the second statement observe that

$$F_{\mathcal{A}} = \sum_{m=0}^{\min(|\mathcal{A}|,k)} \sum_{\mathbf{a}\in\mathcal{A}^{(m)}} \frac{1}{m!} \sum_{\mathbf{z}\in\mathcal{X}^{(k-m)}} h_{k-m,\mathbf{a}}(\mathbf{z})\,,$$

where

$$h_{k-m,\mathbf{a}}(z_1,\ldots,z_{k-m}) := \frac{1}{m!(m-k)!} \sum_{\pi\in S_{k-m}} f\big(a_1,\ldots,a_m, z_{\pi(1)},\ldots,z_{\pi(k-m)}\big)\,. \quad \square$$

The following fact regarding the radius of stabilization is used in the proof of (1.17) in Section 3.2.

Lemma 5.2. *Let ξ be a score function on a locally finite input $\mathcal{X}$ and R^ξ its radius of stabilization. For a given $t > 0$ consider score function $\tilde{\xi}(x,\mathcal{X}) := \xi(x,\mathcal{X})\mathbf{1}[R^\xi(x,\mathcal{X}) \le t]$. Then the radius of stabilization $R^{\tilde{\xi}}$ of $\tilde{\xi}$ is bounded by t: $R^{\tilde{\xi}}(x,\mathcal{X}) \le t$ for any locally finite input $\mathcal{X}$ and $x \in \mathcal{X}$.*

Proof. Let $\mathcal{X}, \mathcal{A}$ be locally finite subsets of $\mathbb{R}^d$ with $x \in \mathcal{X}$. We have

$$\begin{aligned}
&\tilde{\xi}(x, (\mathcal{X} \cap B_t(x)) \cup (\mathcal{A} \cap B_t^c(x))) \\
&= \xi(x, (\mathcal{X} \cap B_t(x)) \cup (\mathcal{A} \cap B_t^c(x)))\, \mathbf{1}\big[R^\xi(x, (\mathcal{X} \cap B_t(x)) \cup (\mathcal{A} \cap B_t^c(x))) \le t\big] \\
&= \xi(x, \mathcal{X} \cap B_t(x))\, \mathbf{1}\big[R^\xi(x, (\mathcal{X} \cap B_t(x)) \cup (\mathcal{A} \cap B_t^c(x))) \le t\big],
\end{aligned}$$

where the last equality follows from the definition of R^ξ. Notice

$$\mathbf{1}\big[R^\xi(x, (\mathcal{X} \cap B_t(x)) \cup (\mathcal{A} \cap B_t^c(x))) \le t\big] = \mathbf{1}\big[R^\xi(x, \mathcal{X} \cap B_t(x)) \le t\big]$$

and so $\tilde{\xi}(x, (\mathcal{X} \cap B_t(x)) \cup (\mathcal{A} \cap B_t^c(x))) = \tilde{\xi}(x, \mathcal{X} \cap B_t(x))$, which was to be shown. □

5.2 Determinantal and permanental point process lemmas

We collect various facts about determinantal and permanental point processes needed in our approach. These facts, of independent interest, illustrate the tractability of these point processes. First,we show that if determinantal and permanental point processes have a kernel K decreasing fast enough, then they generate admissible clustering point processes satisfying clustering conditions (1.13) and (1.7) respectively. We are indebted to Manjunath Krishnapur, who sketched to us the proof of the next result.

Lemma 5.3. *Let $\mathcal{P}$ be a stationary determinantal point process on $\mathbb{R}^d$ with a kernel satisfying $K(x,y) \le \omega(|x-y|)$, where ω is at (2.7). Then*

$$|\rho^{(n)}(x_1, \ldots, x_{p+q}) - \rho^{(p)}(x_1, \ldots, x_p)\rho^{(q)}(x_{p+1}, \ldots, x_{p+q})| \le n^{1+\frac{n}{2}} \omega(s) \|K\|^{n-1}, \qquad (5.1)$$

where $\|K\| := \sup_{x,y \in \mathbb{R}^d} |K(x,y)|$, s is at (1.2), and $n = p+q$.

Proof. Define the matrices $K_0 := ((K(x_i, x_j))_{1 \le i,j \le n}, K_1 := ((K(x_i, x_j))_{1 \le i,j \le p}$, and $K_2 := ((K(x_i, x_j))_{p+1 \le i,j \le n}$. Let L be the block diagonal matrix with blocks K_1, K_2. We define $\|K_0\| := \sup_{1 \le i,j \le n} |K_0(x_i, x_j)|$ and similarly for the other matrices. Then

$$\begin{aligned}
|\rho^{(n)}(x_1, \ldots, x_{p+q}) - \rho^{(p)}(x_1, \ldots, x_p)\rho^{(q)}(x_{p+1}, \ldots, x_{p+q})| &= |\det(K_0) - \det(K_1)\det(K_2)| \\
&= |\det(K_0) - \det(L)| \\
&\le n^{1+\frac{n}{2}} \|K_0 - L\| \|K_0\|^{n-1} \qquad (5.2) \\
&\le n^{1+\frac{n}{2}} \omega(s) \|K\|^{n-1},
\end{aligned}$$

where the inequality follows by [1, (3.4.5)]. This gives (5.1). □

As a first step to prove the analogue of Lemma 5.3 for permanental point processes, we prove an analogue of (5.2). We follow verbatim the proof of (5.2) as given in [1, (3.4.5)]. Instead of using Hadamard's inequality for determinants as in [1], we use the following version of Hadamard's inequality for permanents ([16, Theorem 1.1]): For any column vectors $v_1, \ldots, v_n$ of length n with complex entries, it holds that

$$|\mathrm{per}([v_1, \ldots, v_n])| \le \frac{n!}{n^{\frac{n}{2}}} \prod_{i=1}^{n} \sqrt{\bar{v_i}^T v_i} \le n! \prod_{i=1}^{n} \|v_i\|,$$

where $\|v_i\|$ is the l_∞-norm of v_i viewed as an n-dimensional complex vector.

Lemma 5.4. *Let $n \in \mathbb{N}$. For any two kernels K and L, we have*

$$|\mathrm{per}(K) - \mathrm{per}(L)| \le n n! \|K - L\| \max\{\|K\|, \|L\|\}^{n-1}.$$

Now, in the proof of Lemma 5.3, using the above estimate instead of (5.2), we establish weak clustering (1.7) of permanental point processes with fast decreasing kernels K.

Lemma 5.5. *Let $\mathcal{P}$ be a stationary permanental point process on $\mathbb{R}^d$ with a fast-decreasing kernel satisfying $K(x, y) \le \omega(|x - y|)$ where ω is at (2.7). Then*

$$|\rho^{(n)}(x_1, \ldots, x_{p+q}) - \rho^{(p)}(x_1, \ldots, x_p)\rho^{(q)}(x_{p+1}, \ldots, x_{p+q})| \le n n! \omega(s) \|K\|^{n-1},$$

where s is at (1.2) and $n = p + q$.

Recall that clustering of α-determinantal point processes, $|\alpha| = 1/m, m \in \mathbb{N}$, relies heavily on Proposition 2.3, whose proof we now give.

Proof of Proposition 2.3. We shall prove the proposition in the case $m = 2$; the general case follows in the same fashion albeit with considerably more notation. Let $x_1, \ldots, x_{p+q}$ be distinct points in $\mathbb{R}^d$ with s at (1.2) as usual. For a subset $S \subset [p+q]$, we abbreviate $\rho^{|S|}(x_j : j \in S)$ by $\rho(S)$. Using (2.17) we have that

$$\begin{aligned}
\rho_0^{(p+q)}([p+q]) &= \sum_{S_1 \sqcup S_2 = [p+q]} \rho(S_1)\rho(S_2) = 2\rho([p+q]) + 2\rho([p])\rho([q]) \\
&+ \sum_{S_1 \sqcup S_2 = [p+q], S_2 \cap [p] = \emptyset, S_i \neq \emptyset} \rho(S_1)\rho(S_2) + \sum_{S_1 \sqcup S_2 = [p+q], S_1 \cap [p] = \emptyset, S_i \neq \emptyset} \rho(S_1)\rho(S_2) \\
&+ \sum_{S_1 \sqcup S_2 = [p+q], S_2 \cap [q] = \emptyset, S_i \neq \emptyset} \rho(S_1)\rho(S_2) + \sum_{S_1 \sqcup S_2 = [p+q], S_1 \cap [q] = \emptyset, S_i \neq \emptyset} \rho(S_1)\rho(S_2) \\
&+ \sum_{S_1 \sqcup S_2 = [p+q], S_i \cap [p] \neq \emptyset, S_i \cap [q] \neq \emptyset} \rho(S_1)\rho(S_2)
\end{aligned}$$

$$
\begin{aligned}
&= 2\rho([p+q]) + 2\rho([p])\rho([q]) + \sum_{S_{21}\sqcup S_{22}=[q], S_{ij}\neq\emptyset} (\rho(S_{21}\cup[p])\rho(S_{22}) + \rho(S_{22}\cup[p])\rho(S_{21})) \\
&\quad + \sum_{S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} (\rho(S_{11}\cup[q])\rho(S_{12}) + \rho(S_{12}\cup[q])\rho(S_{11})) \\
&\quad + \sum_{S_{21}\sqcup S_{22}=[q], S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} \rho(S_{11}\cup S_{21})\rho(S_{12}\cup S_{22}).
\end{aligned}
$$

On the other hand the product of correlation functions is

$$
\begin{aligned}
&\rho_0([p])\rho_0([q]) = (\sum_{S_{11}\sqcup S_{12}=[p]} \rho(S_{11})\rho(S_{12}))(\sum_{S_{21}\sqcup S_{22}=[q]} \rho(S_{21})\rho(S_{22})) \\
&= (2\rho([p]) + \sum_{S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} \rho(S_{11})\rho(S_{12})) \\
&\quad \times(2\rho([q]) + \sum_{S_{21}\sqcup S_{22}=[q], S_{ij}\neq\emptyset} \rho(S_{21})\rho(S_{22})) \\
&= 4\rho([p])\rho([q]) \\
&\quad +2 \sum_{S_{21}\sqcup S_{22}=[q], S_{ij}\neq\emptyset} \rho(S_{21})\rho([p])\rho(S_{22}) + 2 \sum_{S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} \rho(S_{11})\rho([q])\rho(S_{12}) \\
&\quad + \sum_{S_{21}\sqcup S_{22}=[q], S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} \rho(S_{11})\rho(S_{21})\rho(S_{12})\rho(S_{22}).
\end{aligned}
$$

Now, we shall match the two summations term-wise and bound the differences using correlation bound (1.8) and clustering condition (1.7):

$$
\begin{aligned}
&|\rho_0([p+q]) - \rho_0([p])\rho_0([q])| \le 2|\rho([p+q]) - \rho([p])\rho([q])| \\
&+ \sum_{S_{21}\sqcup S_{22}=[q], S_{ij}\neq\emptyset} |\rho(S_{21}\cup[p])\rho(S_{22}) - \rho(S_{21})\rho([p])\rho(S_{22})| \\
&+ \sum_{S_{21}\sqcup S_{22}=[q], S_{ij}\neq\emptyset} |\rho(S_{22}\cup[p])\rho(S_{21}) - \rho(S_{21})\rho([p])\rho(S_{22})| \\
&+ \sum_{S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} |\rho(S_{11}\cup[q])\rho(S_{12}) - \rho(S_{11})\rho([q])\rho(S_{12})| \\
&+ \sum_{S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} |\rho(S_{12}\cup[q])\rho(S_{11}) - \rho(S_{11})\rho([q])\rho(S_{12})| \\
&+ \sum_{S_{21}\sqcup S_{22}=[q], S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} |\rho(S_{11}\cup S_{21})\rho(S_{12}\cup S_{22}) - \rho(S_{11})\rho(S_{21})\rho(S_{12})\rho(S_{22})| \\
&\le 2C_{p+q}\phi(c_{p+q}s)\kappa_{p+q}[1 + \sum_{S_{21}\sqcup S_{22}=[q], S_{ij}\neq\emptyset} 1 + \sum_{S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} 1 + \sum_{S_{21}\sqcup S_{22}=[q], S_{11}\sqcup S_{12}=[p], S_{ij}\neq\emptyset} 1] \\
&\le 2\kappa_{p+q}C_{p+q}\phi(c_{p+q}s) \sum_{S_1\sqcup S_2=[p+q]} 1 = 2\kappa_{p+q}C_{p+q}\phi(c_{p+q}s)2^{p+q}. \quad \square
\end{aligned}
$$

To bound the radius of stabilization of geometric functionals on a determinantal point process, we shall need the following estimate of exponential decay of Palm void probability. Though the proof is inspired by the proof of a similar estimate in [47, Lemma 2], we derive a more general and explicit estimate.

Lemma 5.6. *Let $\mathcal{P}$ be a stationary determinantal point process on $\mathbb{R}^d$. Then for $p, k \in \mathbb{N}$, $r > 0$, $\mathbf{x} \in \mathbb{R}^{dp}$, and $y \in \mathbb{R}^d$, we have*

$$\mathbb{P}^!_{\mathbf{x}}(\mathcal{P}(B_r(y)) \leq k) \leq e^{(2k+p)/8} e^{-K(\mathbf{0},\mathbf{0})\mathrm{Vol}(B_r)/8}. \tag{5.3}$$

Proof. Note that we may assume $y = \mathbf{0}$ without loss of generality. Let $\mathbf{x} = \{x_1, \ldots, x_p\}$. For any determinantal point process $\mathcal{P}$ (even non-stationary), let $\mathcal{P}_x$ be the reduced Palm point process with respect to $x \in \mathbb{R}^d$. From (2.11) (see also [67, Theorem 6.5]), we can explicitly describe the correlation functions of $\mathcal{P}_x$. Thus, we have that $\mathcal{P}_x$ is also a determinantal point process and its kernel L is given by

$$L(y_1, y_2) = K(y_1, y_2) - \frac{K(y_1, x)K(x, y_2)}{K(x, x)}. \tag{5.4}$$

The simple inequality $\int_{\mathbb{R}^d} |K(x, y)|^2 \mathrm{d}y \leq K(x, x)$ shows that for any bounded Borel subset B

$$\begin{aligned} \mathbb{E}^!_x(\mathcal{P}(B)) &= \int_B L(y, y)\mathrm{d}y \\ &= \int_B K(y, y)\mathrm{d}y - \frac{1}{K(x, x)} \int_B |K(x, y)|^2 \mathrm{d}y \\ &\geq \mathbb{E}(\mathcal{P}(B)) - 1. \end{aligned}$$

Now re-iterating the above inequality, we get that for all $\mathbf{x} \in (\mathbb{R}^d)^{(p)}$ and any bounded Borel subset B

$$\mathbb{E}^!_{\mathbf{x}}(\mathcal{P}(B)) \geq \mathbb{E}(\mathcal{P}(B)) - p. \tag{5.5}$$

Since determinantal point processes are a sum of independent Bernoulli random variables [9, Theorem 4.5.3], the Chernoff-Hoeffding bound [46, Theorem 4.5] yields

$$\mathbb{P}^!_{\mathbf{x}}(\mathcal{P}(B) \leq \mathbb{E}^!_{\mathbf{x}}(\mathcal{P}(B))/2) \leq e^{-\mathbb{E}^!_{\mathbf{x}}(\mathcal{P}(B))/8}. \tag{5.6}$$

Now we return to our stationary determinantal point process $\mathcal{P}$ and note that $\mathbb{E}(\mathcal{P}(B)) = K(\mathbf{0}, \mathbf{0})\mathrm{Vol}(B)$. Choose r_0 (depending only on p, k) large enough such that $K(\mathbf{0}, \mathbf{0})\mathrm{Vol}(B_{r_0}) = 2k + p$. Thus combining (5.5) and (5.6), we have that for $r \in (r_0, \infty)$

$$\mathbb{P}^!_{\mathbf{x}}(\mathcal{P}(B_r) \leq k) \leq \mathbb{P}^!_{\mathbf{x}}(\mathcal{P}(B_r) \leq \mathbb{E}^!_{\mathbf{x}}(\mathcal{P}(B))/2) \leq e^{-(K(\mathbf{0},\mathbf{0})\mathrm{Vol}(B_r)-p)/8}.$$

For $r \in (0, r_0]$, the definition of r_0 shows that the right-hand side of (5.3) is larger than 1 and hence it is a trivial bound. $\square$

Inequality (5.5) can also be deduced from the stronger coupling result of [54, Prop. 5.10(iv)] for determinantal point processes with a continuous kernel but we have given an elementary proof. Given Ginibre input, we can improve the exponent in the void probability bound (5.3). We believe this result to be of independent interest.

This we achieve by generalizing [70, Lemma 6.1] (which treats the case $k = 0$). The proof uses again Cauchy's interlacing theorem to bound the Palm probability of $\{\mathcal{P}(B_r) \leq k\}$ by a scalar multiple of its stationary probability and then we use representation results for the Ginibre process to bound the probability more explicitly.

Lemma 5.7. *Let $B_r := B_r(\mathbf{0}) \subset \mathbb{R}^2$ and let $\mathcal{P}$ be the Ginibre point process. Then for $p, k \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{R}^{2p}$,*

$$\mathbb{P}^!_{\mathbf{x}}(\mathcal{P}(B_r) \leq k) \leq \exp\{p(k+1)r^2\}\mathbb{P}(\mathcal{P}(B_r) \leq k) \leq kr^{2k} \exp\{(p(k+1)+k)r^2 - \frac{1}{4}r^4(1+o(1))\}. \tag{5.7}$$

We remark that stationarity shows the above bound holds for any radius r ball.

Proof. Again, we shall prove the result for $p = 1$ and use induction to deduce the general case. So, let $\mathbf{x} = x \in \mathbb{R}^2$.

Let $\mathcal{K}_{B_r}$ be the restriction to B_r of the integral operator $\mathcal{K}$ (generated by kernel K) corresponding to Ginibre point process and $\mathcal{L}_{B_r}$ be the restriction to B_r of the integral operator $\mathcal{L}$ (generated by kernel L) corresponding to the reduced Palm point process (also a determinantal point process). Let $\lambda_i, i = 1, 2, \dots$ and $\mu_i, i = 1, 2, \dots$ be the eigenvalues of $\mathcal{K}_{B_r}$ and $\mathcal{L}_{B_r}$ in decreasing order respectively.

Then from (5.4) we have that the rank of the integral operator $\mathcal{K}_{B_r} - \mathcal{L}_{B_r}$ is one. Secondly, note that

$$\sum_i \mu_i = \mathbb{E}_x(\mathcal{P}(B_r)) = \int_{B_r} L(y,y)\mathrm{d}y \leq \int_{B_r} K(y,y)\mathrm{d}y = \mathbb{E}(\mathcal{P}(B_r)) = \sum_i \lambda_i.$$

Hence, by a generalisation of Cauchy's interlacement theorem [18, Theorem 4] combined with the above inequality, we get that the respective eigenvalues satisfy the interlacing inequality $\lambda_i \geq \mu_i \geq \lambda_{i+1}$ for $i = 1, 2, \dots$.

Again by [9, Theorem 4.5.3], we have that $P(B_r) \overset{d}{=} \sum_i \text{Bernoulli}(\lambda_i)$ and under Palm measure, $P(B_r) \overset{d}{=} \sum_i \text{Bernoulli}(\mu_i)$ where both the sums involve independent Bernoulli random variables. Independence of the Bernoulli random variables gives

$$\begin{aligned} \mathbb{P}_x(\mathcal{P}(B_r) \leq k) &= \sum_{J \subset \mathbb{N}, |J| \leq k} \prod_{j \in J} \mu_j \prod_{j \notin J} (1 - \mu_j) \\ &\leq \sum_{J \subset \mathbb{N}, |J| \leq k} \prod_{j \in J} \lambda_j \prod_{j \notin J} (1 - \lambda_{j+1}) \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{J\subset\mathbb{N},|J|\leq k}\prod_{j\in J}\lambda_j\prod_{j\notin J}(1-\lambda_j)\prod_{j-1\in J\cup\{0\},j\notin J}(1-\lambda_j)^{-1}\\
&\leq (1-\lambda_1)^{-k-1}\sum_{J\subset\mathbb{N},|J|\leq k}\prod_{j\in J}\lambda_j\prod_{j\notin J}(1-\lambda_j)=(1-\lambda_1)^{-k-1}\mathbb{P}[\mathcal{P}(B_r)\leq k].
\end{aligned}$$

The proof of the first inequality in (5.7) for the case $p=1$ is complete by noting that $\lambda_1=\mathbb{P}(EXP(1)\leq r^2)$ (see [9, Theorems 4.7.1 and 4.7.3]), where $EXP(1)$ stands for an exponential random variable with mean 1. As said before, iteratively the first inequality in (5.7) can be proven for an arbitrary p. To complete the proof of the second inequality, we bound $\mathbb{P}(\mathcal{P}(B_r)\leq k)$ in a manner similar to the proof of [9, Proposition 7.2.1].

Let $\mathcal{P}^*:=\{R_1^2,R_2^2,\ldots,\}=\{|X|^2:X\in\mathcal{P}\}$ be the point process of squared modulii of the Ginibre point process. Then, from [9, Theorem 4.7.3], it is known that $R_i^2\stackrel{d}{=}\Gamma(i,1)$ ($\Gamma(i,1)$ denotes a gamma random variable with parameters $i,1$) and are independently distributed. We shall need the bound that for all $i\geq 1$,

$$\mathbb{P}(R_i^2\geq r^2)\leq e^{-\beta r^2}\mathbb{E}(e^{\beta R_i^2})\leq e^{-\beta r^2}(1-\beta)^{-i},$$

for some constant $\beta\in(0,1)$. For $i<r^2$, the bound is optimal for $\beta=1-\frac{i}{r^2}$. For r, set $r_*:=\ulcorner r^2\urcorner$, the ceiling of r^2. Then,

$$\begin{aligned}
\mathbb{P}(\mathcal{P}(B_r)\leq k) &= \mathbb{P}(\sharp\{i:R_i^2\leq r^2\}\leq k)\leq\mathbb{P}(\sharp\{i\leq r_*:R_i^2\leq r^2\}\leq k)\\
&\leq \sum_{J\subset[r_*],|J|\leq k}\prod_{i\in J}\mathbb{P}(R_j^2\leq r^2)\prod_{i\notin J}\mathbb{P}(R_j^2>r^2)\\
&\leq \sum_{J\subset[r_*],|J|\leq k}\prod_{i\in J}e^{r^2}e^{-\beta r^2}(1-\beta)^{-i}\prod_{i\notin J}e^{-\beta r^2}(1-\beta)^{-i}\\
&\leq kr^{2k}e^{kr^2}\prod_{i=1}^{r_*}e^{-ar^2}(1-\beta)^{-i}\\
&= kr^{2k}e^{kr^2}e^{-\frac{1}{4}r^4(1+o(1))},
\end{aligned}$$

where equality follows by substituting the optimal β for each i, as in [9, Section 7.2]. □

Acknowledgements

The work has benefitted from DY's visits to Lehigh University and IMA, Minneapolis supported in part by the respective institutions. Part of this work was done when DY was a post-doc at Technion, Israel. He is thankful to the institute for its support and to his host Robert Adler for many discussions. JY's research was supported in part by an NSF grant. The authors thank Manjunath Krishnapur for numerous inputs, especially on determinantal point processes and Gaussian analytic functions. The authors are also thankful to Jesper Møller and Günter Last for useful comments on the first draft of this article.

References

[1] G. W. Anderson, A. Guionnet and O. Zeitouni (2010), *An Introduction to Random Matrices,* Cambridge University Press, Cambridge, U.K.

[2] F. Baccelli and P. Bermolen (2012), Extremal versus additive Matérn point processes, *Queueing Syst.* **71**, No. 1-2, 179–197.

[3] A. Baddeley (1980), A limit theorem for statistics of spatial data, *Adv. in Appl. Probab.,* **12**, 2, 447–461.

[4] A. D. Barbour and A. Xia (2006), Normal approximation for random sums, *Adv. Appl. Probab.*, **38**, 693–728.

[5] Yu. Baryshnikov, P. Eichelsbacher, T. Schreiber, and J. E. Yukich (2008), Moderate deviations for some point measures in geometric probability, *Ann. Inst. Henri Poincaré Probab. Stat.*, **44**, 3, 422-446.

[6] Yu. Baryshnikov and J. E. Yukich (2003), Gaussian fields and random packing, *J. of Stat. Phys.*, **111**, 443–463.

[7] Yu. Baryshnikov and J. E. Yukich (2005), Gaussian limits for random measures in geometric probability. *Ann. Appl. Prob.*, **15**, no. 1A, 213–253.

[8] K. Baumann and G. C. Hegerfeldt (1985), A noncommutative Marcinkiewicz theorem, *Publications of the Research Institute for Mathematical Sciences*, **21**, no. 1, 191–204.

[9] J. Ben Hough, M. Krishnapur, Y. Peres, and B. Virág (2009), *Zeros of Gaussian Analytic Functions and Determinantal Point Processes,* Vol. 51, American Mathematical Society, Providence, RI.

[10] A. Björner (1995), Topological methods, In *Handbook of Combinatorics*, Vol. 2, 1819–1872.

[11] B. Błaszczyszyn (1995), Factorial moment expansion for stochastic systems, *Stoch. Proc. Appl.*, **56**, 2, 321–335.

[12] B. Błaszczyszyn, E. Merzbach and V. Schmidt (1997), A note on expansion for functionals of spatial marked point processes, *Stat. Probab. Letters* , **36**, 3, 299–306.

[13] B. Błaszczyszyn and D. Yogeshwaran (2014), On comparison of clustering properties of point processes, *Adv. Appl. Probab.*, **46**, 1, 1–20.

[14] O. Bobrowski and R.J. Adler (2014), Distance functions, critical points, and the topology of random Cech complexes, *Homology Homotopy Appl.*, **16**, 2, 311–344.

[15] A. Bulinski and E. Spodarev (2012), Central limit theorems for weakly dependent random fields. In *Stochastic Geometry, Spatial Statistics and Random Fields : Asymptotic methods. Ed. Spodarev, E.*, Lect. Notes Math., Springer, vol. 2068.

[16] E. Carlen, E. Lieb, M. Loss (2006), An inequality of Hadamard type for permanents, *Meth. Appl. Anal.*, **13**, 1, 1-18.

[17] D. J. Daley and D. Vere-Jones (2003), *An Introduction to the Theory of Point Processes,* Springer, NY.

[18] J. Dancis and C. Davis (1987), An interlacing theorem for eigenvalues of self-adjoint operators, *Lin. Alg. Appl.*, **88–89**, 117–122.

[19] L. Decreusefond, E. Ferraz, H. Randriam, and A. Vergne (2014), Simplicial Homology of Random Configurations, *Adv. in Appl. Probab.* **46**, 2, 1-20.

[20] H. Edelsbrunner and J. L. Harer (2010), *Computational Topology : An Introduction,* American Mathematical Society, Providence, RI.

[21] P. Eichelsbacher, M. Raic, and T. Schreiber (2015), Moderate deviations for stabilizing functionals in geometric probability, *Ann. Inst. Henri Poincaré Probab. Stat.*, **51**, No. 1, 89-128.

[22] P. Eichelsbacher and C. Thäle (2014), New Berry-Esseen bounds for non-linear functionals of Poisson random measures, *Electron. J. Probab.* **102**(19).

[23] R. Forman (2002), A user's guide to discrete Morse theory, *Sém. Lothar. Combin.*, **48**.

[24] P. J. Forrester and G. Honner (1999), Exact statistical properties of the zeros of complex random polynomials, *J. of Physics A: Math. and General*, **32**, 16, 2961.

[25] M. Franceschetti and R. Meester (2007), *Random Networks for Communication,* Cambridge Series in Statistical and Probabilistic Mathematics, Camb. Univ. Press.

[26] V. Gershkovich and H. Rubinstein (1997), Morse theory for Min-type functions, *The Asian J. Math.* **1**, 696–715.

[27] J. Grote and C. Thäle (2016), Gaussian polytopes: a cumulant-based approach, *pre-print* arXiv:1602.06148.

[28] A. Goldman (2010), The Palm measure and the Voronoi tessellation for the Ginibre point process, *Ann. Appl. Probab.*, **20**, No. 1, 90–128.

[29] P. Hall (1988), *Introduction to the Theory of Coverage Processes,* John Wiley and Sons.

[30] K-H. Hanisch (1982), On inversion formulae for n-fold Palm distributions of point processes in LCS-spaces, *Math. Nachr.*, **106**, 1, 171–179.

[31] L. Heinrich (1994), Normal approximation for some mean-value estimates of absolutely regular tessellations, *Math. Meth. Stat.*, **3**, 1, 1–24.

[32] L. Heinrich (2013), Asymptotic methods in statistics of random point processes, In *Stochastic Geometry, Spatial Statistics and Random Fields : Asymptotic methods. Ed. Spodarev, E.*, Lect. Notes Math., Springer, vol. 2068.

[33] L. Heinrich and I. Molchanov (1999), Central limit theorem for a class of random measures associated with germ-grain models, *Adv. Appl. Probab.*, **31**, 2, 283–314.

[34] D. Hug, G. Last and M. Schulte (2016) Second order properties and central limit theorems for geometric functionals of Boolean models, *Ann. Appl. Probab.*, **26**, 1, 73–135.

[35] M. Kahle (2011), Random geometric complexes, *Discrete Comput. Geom.*, **45**, 3, 553–573.

[36] M. Krishnapur (2006), Overcrowding estimates for zeroes of planar and hyperbolic Gaussian analytic functions, *J. Stat. Phys.*, **124**, No. 6, 1399–1423.

[37] R. Lachièze-Rey and G. Peccati (2015), New Kolmogorov bounds for functionals of binomial point processes, *pre-print* arXiv:1505.04640.

[38] G. Last and M. Penrose (2010), Poisson process Fock space representation, chaos expansion and covariance inequalities, *Probab. Th. Rel. Fields*, **150**, 663—690.

[39] G. Last and M. Penrose (2012), Percolation and limit theory for the Poisson lilypond model, *Rand. Struct. Alg.*, **42**, 226-249.

[40] G. Last, G. Peccati and M. Schulte (2016), Normal approximations on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization, *Probab. Th. Rel. Fields.*, to appear.

[41] V. A. Malyshev (1975), The central limit theorem for Gibbsian random fields, *Soviet Mathematics Doklady* **16**, 1141–1145.

[42] V. A. Malyshev and R. A. Minlos (1991), *Gibbs Random Fields,* Kluwer.

[43] Ph. A. Martin and T. Yalcin (1980), The charge fluctuations in classical Coulomb systems, *J. Stat. Phys.*, **22**, No. 4, 435–463.

[44] P. McCullagh and J. Møller (2006), The permanental process, *Adv. Appl. Probab.*, **38**, No. 4, 873–888.

[45] J. Milnor (1963), *Morse Theory,* Princeton University Press.

[46] M. Mitzenmacher and E. Upfal (2005), *Probability and Computing Randomized Algorithms and Probabilistic Analysis,* Cambridge University Press, Cambridge.

[47] N. Miyoshi and T. Shirai (2015), A sufficient condition for the tail asymptotics of SIR distribution in downlink cellular networks, *pre-print* arXiv:1602.00844.

[48] J. Møller (2003), Shot noise Cox processes, *Adv. Appl. Probab.*, **35**, No. 3 614–640.

[49] J. Møller and G.L. Torrisi (2005), Generalised shot noise Cox processes, *Adv. Appl. Probab.*, **37**, No. 1, 48–74.

[50] J. R. Munkres (1995), *Elements of Algebraic Topology,* Westview Press.

[51] F. Nazarov and M. Sodin (2012), Correlation functions for random complex zeroes: Strong clustering and local universality, *Comm. Math. Phys.*, **310**, No. 1, 75–98.

[52] I. Nourdin and G. Peccati (2012), *Normal Approximations with Malliavin Calculus : From Stein's Method to Universality*, Cambridge Tracts in Mathematics, **192**, Cambridge University Press, Cambridge.

[53] G. Peccati and M. Reitzner (Editors) (2016), *Stochastic Analysis for Poisson Point Processes: Malliavin Calculus, Wiener-Itô Chaos Expansions and Stochastic Geometry,* to appear, Springer-Verlag.

[54] R. Pemantle and Y. Peres (2014), Concentration of Lipschitz functionals of determinantal and other strong Rayleigh measures, *Comb. Probab. Comput.*, **23**, 140–160.

[55] M. D. Penrose (2003), *Random Geometric Graphs,* Oxford Univ. Press, New York.

[56] M. D. Penrose (2007), Gaussian limits for random geometric measures, *Elec. J. Probab.*, **12**, 989–1035.

[57] M. D. Penrose (2007), Laws of large numbers in stochastic geometry with statistical applications, *Bernoulli*, **13**, 1124-1150.

[58] M. D. Penrose and J. E. Yukich (2001), Central limit theorems for some graphs in computational geometry, *Ann. Appl. Probab.*, **11**, 1005-1041.

[59] M. D. Penrose and J. E. Yukich (2002), Limit theory for random sequential packing and deposition, *Ann. Appl. Probab.*, **12**, 272-301.

[60] M. D. Penrose and J. E. Yukich (2003), Weak laws of large numbers in geometric probability, *Ann. Appl. Probab.*, **13**, 277-303.

[61] M. D. Penrose and J. E. Yukich (2005), Normal approximation in geometric probability, *In Stein's Method and Applications*, Lecture Note Series, Inst. for Math. Sci., National Univ. of Singapore, **5**, A. D. Barbour and Louis H. Y. Chen, Eds., 37-58.

[62] M. Reitzner and M. Schulte (2013), Central limit theorems for U-statistics of Poisson point processes, *Ann. Probab.*, **41**, No. 6, 3879–3909.

[63] R. Schneider (1993), *Convex Bodies: the Brunn-Minkowski Theory,* Cambridge University Press, Cambridge.

[64] R. Schneider and W. Weil (2008), *Stochastic and Integral geometry,* Probability and its Applications (New York), Springer-Verlag, Berlin.

[65] T. Schreiber and J. E. Yukich (2013), Limit theorems for geometric functionals of Gibbs point processes, *Ann. Inst. Henri Poincaré Probab. Stat.*, **49**, 4, 158–1182.

[66] M. Schulte and C. Thäle (2016), Central limit theorems for the radial spanning tree. *Rand. Struct. Alg.*, online version.

[67] T. Shirai and Y. Takahashi (2003), Random point fields associated with certain Fredholm determinants I: fermion, Poisson and boson point processes, *J. Funct. Anal.*, **205**, 2, 414–463.

[68] A. Soshnikov (2002), Gaussian limit for determinantal random point fields, *Ann. Probab.*, **30**, 171-187.

[69] A. Xia and J. E. Yukich (2015), Normal approximation for statistics of Gibbsian input in geometric probability, *Adv. Appl. Probab.*, **47**, 934-972.

[70] D. Yogeshwaran and R. J. Adler (2015), On the topology of random complexes built over stationary point processes, *Ann. Appl. Prob.*, **25**, No. 6, 3338–3380.

[71] J. E. Yukich (1998), *Probability Theory of Classical Euclidean Optimization Problems,* Lecture Notes in Mathematics, vol. 1675, Springer-Verlag.

[72] J. E. Yukich (2013), Limit theorems in discrete stochastic geometry. In *Stochastic Geometry, Spatial Statistics and Random Fields : Asymptotic methods. Ed. Spodarev, E.*, Lect. Notes Math., Springer, vol. 2068.

[73] A. Zomorodian (2012), Topological data analysis. In *Advances in Applied and Computational Topology*, Proc. Sympos. Appl. Math., Amer. Math. Soc., **70**, 1–39.