Normal approximation for sums of stabilizing functionals

Raphael Lachièze-Rey[*], Matthias Schulte[†], J. E. Yukich[‡]

December 19, 2016

Abstract

We establish presumably optimal rates of normal convergence with respect to the Kolmogorov distance for a large class of geometric functionals of marked Poisson and binomial point processes on general metric spaces. The rates are valid whenever the geometric functional is expressible as a sum of exponentially stabilizing score functions satisfying a moment condition. By incorporating stabilization methods into the Malliavin-Stein theory, we obtain rates of normal approximation for sums of stabilizing score functions which either improve upon existing rates or are the first of their kind.

Our general rates hold for functionals of marked input on spaces more general than full-dimensional subsets of $\mathbb{R}^d$, including m-dimensional Riemannian manifolds, $m \leq d$. We use the general results to deduce improved and new rates of normal convergence for functionals whose variances re-scale as either the volume or the surface area of an underlying set. In particular, we improve upon rates of normal convergence for the k-face and ith intrinsic volume functionals of the convex hull of Poisson and binomial random samples in a smooth convex body in dimension $d \geq 2$. We also provide improved rates of normal convergence for (i) statistics of nearest neighbor graphs and high-dimensional data sets, (ii) estimators of surface area and volume arising in set approximation via Voronoi tessellations, (iii) the number of maximal points in a random sample, and (iv) clique counts in generalized random geometric graphs.

[*]Université Paris Descartes, raphael.lachieze-rey@parisdescartes.fr

[†]Universität Bern, matthias.schulte@stat.unibe.ch

[‡]Lehigh University, jey0@lehigh.edu

Research supported in part by NSF grant DMS-1406410 (JY). This research was initiated during the workshop 'New Directions in Stein's Method' at the Institute for Mathematical Sciences in Singapore. The authors are very grateful to the IMS for their hospitality and support.

Key words and phrases. Stein's method, Malliavin calculus, stabilization, random Euclidean graphs, statistics of data sets, statistics of convex hulls, Voronoi set approximation, maximal points

AMS 2010 Subject Classification: Primary 60F05: Central limit and other weak theorems; Secondary 60D05: Geometric probability and stochastic geometry

1 Introduction

Let $(\mathbb{X}, \mathcal{F})$ be a measurable space equipped with a σ-finite measure $\mathbb{Q}$ and a measurable semi-metric $\mathrm{d} : \mathbb{X} \times \mathbb{X} \to [0, \infty)$. For all $s \geq 1$ let $\mathcal{P}_s$ be a Poisson point process with intensity measure $s\mathbb{Q}$. When $\mathbb{Q}$ is a probability measure we let $\mathcal{X}_n$ be a binomial point process of n points which are i.i.d. according to $\mathbb{Q}$. Consider the statistics

$$H_s := h_s(\mathcal{P}_s) := \sum_{x \in \mathcal{P}_s} \xi_s(x, \mathcal{P}_s), \quad s \geq 1, \tag{1.1}$$

and

$$H'_n := h_n(\mathcal{X}_n) := \sum_{x \in \mathcal{X}_n} \xi_n(x, \mathcal{X}_n), \quad n \in \mathbb{N}, \tag{1.2}$$

where, roughly speaking, the scores $\xi_s(x, \mathcal{P}_s)$ and $\xi_n(x, \mathcal{X}_n)$ represent the local contributions to the global statistics H_s and H'_n, respectively. Functionals such as H_s and H'_n, which are in some sense locally defined, are called stabilizing functionals. The concept of stabilization and the systematic investigation of stabilizing functionals go back to the papers [29, 30]. In the following we are interested in quantitative central limit theorems for stabilizing functionals, whereas laws of large numbers are shown in [27, 30] and moderate deviations are considered in [14]. For a survey on limit theorems in stochastic geometry with a particular focus on stabilization we refer to [41]. Statistics H_s and H'_n typically describe a global property of a random geometric structure on $\mathbb{X}$ in terms of local contributions exhibiting spatial interaction and dependence. Functionals in stochastic geometry which may be cast in the form of (1.1) and (1.2) include total edge length and clique counts in random graphs, statistics of Voronoi set approximation, the k-face and volume functional of convex hulls of random point samples, as well as statistics of RSA packing models and spatial birth growth models.

Functionals H_s and H'_n sometimes involve marked point processes. To handle this generality we consider the product of $(\mathbb{X}, \mathcal{F}, \mathbb{Q})$ with an external probability 'marks' space $(\mathbb{M}, \mathcal{F}_{\mathbb{M}}, \mathbb{Q}_{\mathbb{M}})$, and we attach to each element of the point process an independent random mark.

Throughout this paper we denote by N a standard Gaussian random variable and by

$$d_K(Y, Z) := \sup_{t \in \mathbb{R}} |\mathbb{P}(Y \leq t) - \mathbb{P}(Z \leq t)| \tag{1.3}$$

the Kolmogorov distance of two random variables Y and Z. For a sum $S_n = \sum_{i=1}^n Y_i$ of n i.i.d. random variables $Y_1, \dots, Y_n$ such that $\mathbb{E}\,|Y_1|^3 < \infty$ it is known from the classical Berry-Esseen theorem that

$$d_K\left(\frac{S_n - \mathbb{E}\,S_n}{\sqrt{\operatorname{Var} S_n}}, N\right) \leq \frac{C\mathbb{E}\,|Y_1 - \mathbb{E}\,Y_1|^3}{\operatorname{Var} Y_1} \frac{1}{\sqrt{\operatorname{Var} S_n}}, \quad n \in \mathbb{N}, \tag{1.4}$$

with $C \in (0, \infty)$ a universal constant. By considering special choices for $Y_1, \dots, Y_n$, one can show that the rate $1/\sqrt{\operatorname{Var} S_n}$ in (1.4) is optimal. The main contribution of this paper is to show for exponentially stabilizing functionals H_s and H'_n bounds similar to those at (1.4), with rates $1/\sqrt{\operatorname{Var} H_s}$ and $1/\sqrt{\operatorname{Var} H'_n}$, respectively. Here the scores $(\xi_s)_{s\geq 1}$ and $(\xi_n)_{n\geq 1}$ have uniformly bounded $(4+p)$th moment, the analog of the assumption $\mathbb{E}\,|Y_1|^3 < \infty$ at (1.4). In contrast to the summands of S_n, the summands of H_s and H'_n are dependent in general, but nevertheless by comparison with the classical Berry-Esseen theorem, one can expect the rates $1/\sqrt{\operatorname{Var} H_s}$ and $1/\sqrt{\operatorname{Var} H'_n}$ to be optimal.

In stochastic geometry, it is frequently the case that $(H_s - \mathbb{E}\,H_s)/\sqrt{\operatorname{Var} H_s}$ converges to the standard normal, and likewise for $(H'_n - \mathbb{E}\,H'_n)/\sqrt{\operatorname{Var} H'_n}$. However up to now there has been no systematic treatment which establishes presumably optimal rates of convergence to the normal. For example, in [9] a central limit theorem for functionals of nearest neighbor graphs is derived, but no rate of convergence is given. Dependency graph methods are used in [1] to show asymptotic normality of the total edge length of the nearest neighbor graph as well as of the Voronoi and Delaunay tessellations, but lead to suboptimal rates of convergence. Anticipating stabilization methods, the authors of [20] proved asymptotic normality for the total edge length of the Euclidean minimal spanning tree, though they did not obtain a rate of convergence. In the papers [7, 26, 29] abstract central limit theorems for stabilizing functionals are derived and applied to several problems from stochastic geometry. Quantitative bounds for the normal approximation of stabilizing functionals of an underlying Poisson point process are given in [5, 28, 31, 32, 50]. These results yield rates of convergence for the Kolmogorov distance of the order $1/\sqrt{\operatorname{Var} H_s}$ times some extraneous logarithmic factors. For stabilizing functionals of an underlying binomial point process we are unaware of analogous results. The paper [11] uses Stein's method to provide rates of normal convergence for functionals on binomial input satisfying a type of local dependence, though these rates are in the Wasserstein distance.

Recent work [23] shows that the Malliavin calculus, combined with Stein's method of normal approximation, yields rates of normal approximation for general Poisson functionals. The rates are in the Kolmogorov distance, they are presumably optimal, and the authors use their general results to deduce rates of normal convergence (cf. Proposition 1.4 and Theorem 6.1 of [23]) for Poisson functionals satisfying a type of stabilization.

That paper states that ‘the new connection between the Stein-Malliavin approach and the theory of stabilization has a great potential for further generalisations and applications’, though it stops short of linking these two fertile research areas.

The first main goal of this paper is to fully develop this connection, showing that the theory of stabilization neatly dovetails with Malliavin-Stein methods, giving presumably optimal rates of normal convergence. Malliavin-Stein rates of normal convergence, expressed in terms of moments of first and second order difference operators [23], seemingly consist of unwieldy terms. However, if ξ_s is exponentially stabilizing and satisfies a moment condition, then our first main goal is to show that the Malliavin-Stein bounds remarkably simplify, showing that

$$d_K\left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{\operatorname{Var} H_s}}, N\right) \leq \frac{\tilde{C}}{\sqrt{\operatorname{Var} H_s}}, \quad s \geq 1, \tag{1.5}$$

as explained in Corollary 2.2. These rates, presumed optimal, remove extraneous logarithmic factors appearing in [5, 28, 31, 32, 50].

Our second main goal is to show that (1.5) holds when H_s is replaced by H'_n, thus giving analogous rates of normal convergence when Poisson input is replaced by binomial input. Recall that the paper [22] (see Theorem 5.1 there) uses Stein's method and difference operators to establish rates of normal convergence in the Kolmogorov distance for general functionals of binomial point processes. Though [22] deduces rates of normal convergence for some statistics of binomial input in geometric probability, it too stops short of systematically developing the connection between stabilization, Stein's method, and difference operators. Our second goal is to explicitly and fully develop this connection. As a by-product, we show that the ostensibly unmanageable bounds in the Kolmogorov distance may be re-cast into bounds which collapse into a single term $1/\sqrt{\operatorname{Var} H'_n}$. In other words, when ξ_n has a a radius of stabilization (with respect to binomial input $\mathcal{X}_n$) which decays exponentially fast, then subject to a moment condition on ξ_n, Corollary 2.2 shows

$$d_K\left(\frac{H'_n - \mathbb{E}\, H'_n}{\sqrt{\operatorname{Var} H'_n}}, N\right) \leq \frac{\tilde{C}'}{\sqrt{\operatorname{Var} H_n}}, \quad n \geq 9. \tag{1.6}$$

The main finding of this paper, culminating much research related to stabilizing score functionals and captured by the rate results (1.5) and (1.6), is this: Statistics (1.1) and (1.2) enjoy presumably optimal rates of normal convergence once the scores ξ_s and ξ_n satisfy exponential stabilization and a moment condition. In problems of interest, the verification of these conditions is sometimes a straightforward exercise, as seen in Section 5, the applications section. On the other hand, for statistics involving convex hulls of random point samples, the verification of these conditions involves a judicious choice of the underlying metric space, one which allows us to express complicated spatial dependencies in relatively simply fashion. This is all illustrated in Section 5, where it is shown for both the intrinsic volumes of the convex hull and for the count of its lower dimensional faces, that the convergence rates (1.5) and (1.6) are either the first of their kind or that they significantly improve upon existing rates of convergence in the literature, for both Poisson and binomial input in all dimensions $d \geq 2$.

Our third and final goal is to broaden the scope of existing central limit theory in such a way that:

(i) The presumably optimal rates (1.5) and (1.6) are applicable both in the context of volume order and of surface area order scaling of the variance of the functional. By this we mean that the variance of H_s (resp. H_n') is of order s (resp. n) or $s^{1-1/d}$ (resp. $n^{1-1/d}$), after renormalising so that the score of an arbitrary point is of constant order. The notions volume order scaling and surface area order scaling come from a different (but for many problems equivalent) formulation where the intensity of the underlying point process is kept fixed and a set carrying the input is dilated instead. In this set-up the variance may be asymptotically proportional to the volume or surface area of the carrying set. Surface order scaling of the variance typically arises when the scores are non-vanishing only for points close to a $(d-1)$-dimensional subset of $\mathbb{R}^d$. As shown in Theorem 5.4, this generality yields improved rates of normal convergence for statistics arising in Voronoi set approximation. It also brings improved rates of normal convergence for the number of maximal points in a random sample.

(ii) The methods are sufficiently general so that they bring within their purview score functions of data on spaces $(\mathbb{X}, \mathrm{d})$, with d an *arbitrary semi-metric*. We illustrate the power of our general approach by establishing a self-contained, relatively short proof of the asymptotic normality of statistics of convex hulls of random point samples in a smooth compact convex set as discussed earlier in this introduction. Our methods also deliver rates of convergence for statistics of k-nearest neighbors graphs and clique counts on both Poisson and binomial input on general metric spaces $(\mathbb{X}, \mathrm{d})$, as seen in Theorems 5.1 and 5.15.

We anticipate that the generality of the methods here will find further non-trivial applications in the central limit theory for functionals in stochastic geometry.

This paper is organized as follows. In Section 2 we give abstract bounds for the normal approximation of stabilizing functionals with respect to Poisson or binomial input, which are our main findings. These are proven in Section 4, which we prepare by recalling and rewriting some existing Malliavin-Stein bounds in Section 3. In Section 5 we demonstrate the power of our general bounds by applying them to several problems from stochastic geometry.

2 Main results

In this section we present our main results in detail. We first spell out assumptions on the measurable space $(\mathbb{X}, \mathcal{F})$, the σ-finite measure $\mathbb{Q}$ and the measurable semi-metric $\mathrm{d} : \mathbb{X} \times \mathbb{X} \to [0, \infty)$. By $B(x, r)$ we denote the ball of radius $r > 0$ around $x \in \mathbb{X}$, i.e. $B(x, r) := \{y \in \mathbb{X} : \mathrm{d}(x, y) \leq r\}$. In the standard set-up for stabilizing functionals, $\mathbb{X}$ is a subset of $\mathbb{R}^d$ and $\mathbb{Q}$ has a bounded density with respect to the Lebesgue measure (see, for example, [26, 31, 50]). To handle more general $\mathbb{X}$ and $\mathbb{Q}$, including $\mathbb{Q}$ having an unbounded density, we replace this standard assumption by the following growth condition on the $\mathbb{Q}$- surface area of spheres: There are constants $\gamma, \kappa > 0$ such that

$$\limsup_{\varepsilon \to 0} \frac{\mathbb{Q}(B(x, r + \varepsilon) \setminus B(x, r))}{\varepsilon} \leq \kappa \gamma r^{\gamma - 1}, \quad r \geq 0, x \in \mathbb{X}. \tag{2.1}$$

Two examples for measure spaces $(\mathbb{X}, \mathcal{F}, \mathbb{Q})$ and semi-metrics d satisfying the assumption (2.1) are the following:

- Example 1. Let $\mathbb{X}$ be a full-dimensional subset of $\mathbb{R}^d$ equipped with the induced Borel-σ-field $\mathcal{F}$ and the usual Euclidean distance d, assume that $\mathbb{Q}$ is a measure on $\mathbb{X}$ with a density g with respect to the Lebesgue measure, and put $\gamma := d$. Then condition (2.1) reduces to the standard assumption that g is bounded. Indeed, if $\|g\|_\infty := \sup_{x \in \mathbb{X}} |g(x)| < \infty$, then (2.1) is obviously satisfied with $\kappa := \|g\|_\infty \kappa_d$, where $\kappa_d := \pi^{d/2}/\Gamma(d/2 + 1)$ is the volume of the d-dimensional unit ball in $\mathbb{R}^d$. On the other hand, if (2.1) holds, then $\mathbb{Q}(B(x, r)) \leq \kappa r^d$ as seen by Lemma 4.1(a) below. This gives an upper bound of κ/κ_d for g since, by Lebesgue's differentiation theorem, Lebesgue almost all points x in $\mathbb{R}^d$ are Lebesgue points, that is to say

  $$g(x) = \lim_{r \to 0} (\kappa_d r^d)^{-1} \int_{y \in B(x,r)} g(y) \, \mathrm{d}y = \lim_{r \to 0} (\kappa_d r^d)^{-1} \mathbb{Q}(B(x, r)) \leq \kappa/\kappa_d.$$

  In case of an unbounded density (2.1) could be satisfied with some $\gamma < d$.

- Example 2. Let $\mathbb{X} \subset \mathbb{R}^d$ be a smooth m-dimensional subset of $\mathbb{R}^d$, $m \leq d$, equipped with a semi-metric d, and a measure $\mathbb{Q}$ on $\mathbb{X}$ with a bounded density g with respect to the uniform surface measure Vol_m on $\mathbb{X}$. We assume that the Vol_{m-1} measure of the sphere $\partial(B(x, r))$ is bounded by the surface area of the Euclidean sphere $\mathbb{S}^{m-1}(0, r)$ of the same radius, that is to say

  $$\mathrm{Vol}_{m-1}(\partial B(x, r)) \leq m \kappa_m r^{m-1}, \ x \in \mathbb{X}, \ r > 0. \tag{2.2}$$

  When $\mathbb{X}$ is an m-dimensional affine space and d is the usual Euclidean metric on $\mathbb{R}^d$, (2.2) holds with equality, naturally. However (2.2) holds in more general situations. For example, by Bishop's comparison theorem (Theorem 1.2 of [40], along with (1.15) there), (2.2) holds for Riemannian manifolds $\mathbb{X}$ with non-negative Ricci curvature, with d the geodesic distance. Given the bound (2.2), one obtains (2.1) with $\kappa = \|g\|_\infty \kappa_m$ and $\gamma = m$. This example includes the case $\mathbb{X} = \mathbb{S}^m$, the unit sphere in $\mathbb{R}^{m+1}$ equipped with the geodesic distance.

In order to deal with marked point processes, let $\widehat{\mathbb{X}} := \mathbb{X} \times \mathbb{M}$, put $\widehat{\mathcal{F}}$ to be the product σ-field of $\mathcal{F}$ and $\mathcal{F}_\mathbb{M}$, and let $\widehat{\mathbb{Q}}$ be the product measure of $\mathbb{Q}$ and $\mathbb{Q}_\mathbb{M}$. When $(\mathbb{M}, \mathcal{F}_\mathbb{M}, \mathbb{Q}_\mathbb{M})$ is a singleton endowed with a Dirac point mass, $\widehat{\mathbb{X}}$ reduces to $\mathbb{X}$ and the 'hat' superscript can be removed in all occurrences.

Let $\mathbf{N}$ be the set of σ-finite counting measures on $\widehat{\mathbb{X}}$, which can be interpreted as point configurations in $\widehat{\mathbb{X}}$. Thus, we treat the elements from $\mathbf{N}$ as sets in our notation. The set $\mathbf{N}$ is equipped with the smallest σ-field $\mathcal{N}$ such that the maps $m_A : \mathbf{N} \to \mathbb{N} \cup \{0, \infty\}, \mathcal{M} \mapsto \mathcal{M}(A)$ are measurable for all $A \in \widehat{\mathcal{F}}$. A point process is now a random element in $\mathbf{N}$. In this paper we consider two different classes of point processes, namely Poisson and binomial point processes. For $s \geq 1$, update the notation $\mathcal{P}_s$ to represent a Poisson point process with intensity measure $s\widehat{\mathbb{Q}}$. This means that the numbers of points of $\mathcal{P}_s$ in disjoint sets $A_1, \ldots, A_m \in \widehat{\mathcal{F}}$, $m \in \mathbb{N}$, are independent and that the number of points of $\mathcal{P}_s$ in a set $A \in \widehat{\mathcal{F}}$ follows a Poisson distribution with mean $s\widehat{\mathbb{Q}}(A)$. In case $\mathbb{Q}$ is a probability measure, we denote similarly by $\mathcal{X}_n$ a binomial point process of $n \in \mathbb{N}$ points that are independently distributed according to $\widehat{\mathbb{Q}}$. Whenever we state a result involving the binomial point process $\mathcal{X}_n$, we implicitly assume that $\mathbb{Q}$, and hence $\widehat{\mathbb{Q}}$, are probability measures.

As mentioned in the first section, we seek central limit theorems for H_s and H'_n defined at (1.1) and (1.2), respectively. We assume that the scores $(\xi_s)_{s\geq 1}$ are measurable functions from $\widehat{\mathbb{X}} \times \mathbf{N}$ to $\mathbb{R}$. To derive central limit theorems for H_s and H'_n, we impose several conditions on the scores. For $s \geq 1$ a measurable map $R_s : \widehat{\mathbb{X}} \times \mathbf{N} \to \mathbb{R}$ is called a *radius of stabilization* for ξ_s if for all $\hat{x} := (x, m_x) \in \widehat{\mathbb{X}}$, $\mathcal{M} \in \mathbf{N}$ and finite $\widehat{\mathcal{A}} \subset \widehat{\mathbb{X}}$ with $|\widehat{\mathcal{A}}| \leq 7$ we have

$$\xi_s(\hat{x}, (\mathcal{M} \cup \{\hat{x}\} \cup \widehat{\mathcal{A}}) \cap \widehat{B}(x, R_s(\hat{x}, \mathcal{M} \cup \{\hat{x}\}))) = \xi_s(\hat{x}, \mathcal{M} \cup \{\hat{x}\} \cup \widehat{\mathcal{A}}), \tag{2.3}$$

where $\widehat{B}(y, r) := B(y, r) \times \mathbb{M}$ for $y \in \mathbb{X}$ and $r > 0$.

For a given point $x \in \mathbb{X}$ we denote by M_x the corresponding random mark, which is distributed according to $\mathbb{Q}_\mathbb{M}$ and is independent of everything else. Say that $(\xi_s)_{s\geq 1}$ (resp. $(\xi_n)_{n\in\mathbb{N}}$) are *exponentially stabilizing* if there are radii of stabilization $(R_s)_{s\geq 1}$ (resp. $(R_n)_{n\in\mathbb{N}}$) and constants $C_{stab}, c_{stab}, \alpha_{stab} \in (0, \infty)$ such that, for $x \in \mathbb{X}$, $r \geq 0$ and $s \geq 1$,

$$\mathbb{P}(R_s((x, M_x), \mathcal{P}_s \cup \{(x, M_x)\}) \geq r) \leq C_{stab} \exp(-c_{stab}(s^{1/\gamma} r)^{\alpha_{stab}}), \tag{2.4}$$

resp. for $x \in \mathbb{X}$, $r \geq 0$ and $n \geq 9$,

$$\mathbb{P}(R_n((x, M_x), \mathcal{X}_{n-8} \cup \{(x, M_x)\}) \geq r) \leq C_{stab} \exp(-c_{stab}(n^{1/\gamma} r)^{\alpha_{stab}}), \qquad (2.5)$$

where γ is the constant from (2.1).

For a finite set $\mathcal{A} \subset \mathbb{X}$ we denote by $(\mathcal{A}, M_{\mathcal{A}})$ the random set obtained by equipping each point of $\mathcal{A}$ with a random mark distributed according to $\mathbb{Q}_{\mathbb{M}}$ and independent of everything else. Given $p \in [0, \infty)$, say that $(\xi_s)_{s \geq 1}$ or $(\xi_n)_{n \in \mathbb{N}}$ satisfy a $(4+p)$-moment condition if there is a constant $C_p \in (0, \infty)$ such that for all $\mathcal{A} \subset \mathbb{X}$ with $|\mathcal{A}| \leq 7$,

$$\sup_{s \in [1,\infty)} \sup_{x \in \mathbb{X}} \mathbb{E}\, |\xi_s((x, M_x), \mathcal{P}_s \cup \{(x, M_x)\} \cup (\mathcal{A}, M_{\mathcal{A}}))|^{4+p} \leq C_p \qquad (2.6)$$

or

$$\sup_{n \in \mathbb{N}, n \geq 9} \sup_{x \in \mathbb{X}} \mathbb{E}\, |\xi_n((x, M_x), \mathcal{X}_{n-8} \cup \{(x, M_x)\} \cup (\mathcal{A}, M_{\mathcal{A}}))|^{4+p} \leq C_p. \qquad (2.7)$$

Let K be a measurable subset of $\mathbb{X}$. By $\mathrm{d}(z, K) := \inf_{y \in K} \mathrm{d}(z, y)$ we denote the distance between a point $z \in \mathbb{X}$ and K. Moreover, we use the abbreviation $\mathrm{d}_s(\cdot, \cdot) := s^{1/\gamma}\, \mathrm{d}(\cdot, \cdot)$, $s \geq 1$. We introduce another notion relevant for functionals whose variances exhibit surface area order scaling. Say that $(\xi_s)_{s \geq 1}$, resp. $(\xi_n)_{n \in \mathbb{N}}$, *decay exponentially fast with the distance to* K if there are constants $C_K, c_K, \alpha_K \in (0, \infty)$ such that for all $\mathcal{A} \subset \mathbb{X}$ with $|\mathcal{A}| \leq 7$ we have

$$\mathbb{P}(\xi_s((x, M_x), \mathcal{P}_s \cup \{(x, M_x)\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \leq C_K \exp(-c_K\, \mathrm{d}_s(x, K)^{\alpha_K}) \qquad (2.8)$$

for $x \in \mathbb{X}$ and $s \geq 1$ resp.

$$\mathbb{P}(\xi_n((x, M_x), \mathcal{X}_{n-8} \cup \{(x, M_x)\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \leq C_K \exp(-c_K\, \mathrm{d}_n(x, K)^{\alpha_K}) \qquad (2.9)$$

for $x \in \mathbb{X}$ and $n \geq 9$. For functionals whose variances have volume order we will make the choice $K = \mathbb{X}$, in which case (2.8) and (2.9) are obviously satisfied with $C_K = 1$. Later we will have that $\mathbb{X}$ is $\mathbb{R}^d$ or a compact convex subset of $\mathbb{R}^d$ such as the unit cube and that K is a $(d-1)$-dimensional subset of $\mathbb{R}^d$. This situation arises, for example, in statistics of random convex polytopes and Voronoi set approximation. Moreover, problems with surface order scaling of the variance are typically of this form.

The following general theorem gives rates of normal convergence for H_s and H_n' in terms of the Kolmogorov distance defined at (1.3). This theorem is a consequence of general theorems from [23] and [22] giving Malliavin-Stein bounds for functionals of Poisson and binomial point processes (see Theorems 3.1 and 3.2 below). Let $\alpha := \min\{\alpha_{stab}, \alpha_K\}$ and

$$I_{K,s} := s \int_{\mathbb{X}} \exp\left(- \frac{\min\{c_{stab}, c_K\} p\, \mathrm{d}_s(x, K)^{\alpha}}{36 \cdot 4^{\alpha+1}} \right) \mathbb{Q}(\mathrm{d}x), \quad s \geq 1. \qquad (2.10)$$

Throughout this paper N always denotes a standard Gaussian random variable. The proofs of the following results are postponed to Section 4.

Theorem 2.1. *(a) Assume that the score functions* $(\xi_s)_{s\geq 1}$ *are exponentially stabilizing* (2.4)*, satisfy the moment condition* (2.6) *for some* $p \in (0,1]$*, decay exponentially fast with the distance to a measurable set* $K \subset \mathbb{X}$*, as at* (2.8)*. Then there is a constant* $\tilde{C} \in (0,\infty)$ *only depending on the constants in* (2.1)*,* (2.4)*,* (2.6) *and* (2.8) *such that*

$$d_K\left(\frac{H_s - \mathbb{E}\,H_s}{\sqrt{\operatorname{Var} H_s}}, N\right) \leq \tilde{C}\left(\frac{\sqrt{I_{K,s}}}{\operatorname{Var} H_s} + \frac{I_{K,s}}{(\operatorname{Var} H_s)^{3/2}} + \frac{I_{K,s}^{5/4} + I_{K,s}^{3/2}}{(\operatorname{Var} H_s)^2}\right), \quad s \geq 1. \tag{2.11}$$

(b) Assume that the score functions $(\xi_n)_{n\in\mathbb{N}}$ *are exponentially stabilizing* (2.5)*, satisfy the moment condition* (2.7) *for some* $p \in (0,1]$*, decay exponentially fast with the distance to a measurable set* $K \subset \mathbb{X}$*, as at* (2.9)*. Let* $(I_{K,n})_{n\in\mathbb{N}}$ *be as in* (2.10)*. Then there is a constant* $\tilde{C} \in (0,\infty)$ *only depending on the constants in* (2.1)*,* (2.5)*,* (2.7) *and* (2.9) *such that*

$$d_K\left(\frac{H_n' - \mathbb{E}\,H_n'}{\sqrt{\operatorname{Var} H_n'}}, N\right) \leq \tilde{C}\left(\frac{\sqrt{I_{K,n}}}{\operatorname{Var} H_n'} + \frac{I_{K,n}}{(\operatorname{Var} H_n')^{3/2}} + \frac{I_{K,n} + I_{K,n}^{3/2}}{(\operatorname{Var} H_n')^2}\right), \quad n \geq 9. \tag{2.12}$$

Assuming growth bounds on $I_{K,s}/\operatorname{Var} H_s$ and $I_{K,n}/\operatorname{Var} H_n'$, the rates (2.11) and (2.12) nicely simplify into presumably optimal rates, ready for off-the-shelf use in applications. Notice that if $K = \mathbb{X}$, then (2.8) and (2.9) hold with $c_K = 0$. Thus, we have

$$I_{\mathbb{X},s} = s\mathbb{Q}(\mathbb{X}), \quad s \geq 1, \qquad \text{and} \qquad I_{\mathbb{X},n} = n\mathbb{Q}(\mathbb{X}), \quad n \in \mathbb{N}. \tag{2.13}$$

Corollary 2.2. *Let the conditions of Theorem 2.1 prevail. Assume further that there is a* $C \in (0,\infty)$ *such that* $\sup_{s\geq 1} I_{K,s}/\operatorname{Var} H_s \leq C$*. Then there is a* $\tilde{C}' \in (0,\infty)$ *only depending on* C *and the constants in* (2.1)*,* (2.4)*,* (2.6) *and* (2.8) *such that*

$$d_K\left(\frac{H_s - \mathbb{E}\,H_s}{\sqrt{\operatorname{Var} H_s}}, N\right) \leq \frac{\tilde{C}'}{\sqrt{\operatorname{Var} H_s}}, \quad s \geq 1. \tag{2.14}$$

If there is a $C \in (0,\infty)$ *such that* $\sup_{n\geq 1} I_{K,n}/\operatorname{Var} H_n' \leq C$*, then there is a* $\tilde{C}' \in (0,\infty)$ *only depending on* C *and the constants in* (2.1)*,* (2.5)*,* (2.7) *and* (2.9) *such that*

$$d_K\left(\frac{H_n' - \mathbb{E}\,H_n'}{\sqrt{\operatorname{Var} H_n'}}, N\right) \leq \frac{\tilde{C}'}{\sqrt{\operatorname{Var} H_n'}}, \quad n \geq 9. \tag{2.15}$$

This corollary is applied in the context of the convex hull of a random sample of points in a smooth convex set in Section 5.4. In this case, the variance is of order $s^{\frac{d-1}{d+1}}$ ($n^{\frac{d-1}{d+1}}$ in the binomial setting), and we obtain rates of normal convergence of order $(\operatorname{Var} H_s)^{-1/2} = \Omega(s^{-(d-1)/(2(d+1))})$ (resp. $(\operatorname{Var} H_n')^{-1/2} = \Omega(n^{-(d-1)/(2(d+1))})$), which improves upon rates obtained via other methods.

In the setting $\mathbb{X} \subset \mathbb{R}^d$, our results admit further simplification, which goes as follows. For $K \subset \mathbb{X} \subset \mathbb{R}^d$ and $r \in (0, \infty)$, let $K_r := \{y \in \mathbb{R}^d : \ \mathrm{d}(y, K) \leq r\}$ denote the r-parallel set of K. Recall that the $(d-1)$-dimensional upper Minkowski content of K is given by

$$\overline{\mathcal{M}}^{d-1}(K) := \limsup_{r \to 0} \frac{\mathrm{Vol}_d(K_r)}{2r}. \tag{2.16}$$

If K is a closed $(d-1)$-rectifiable set in $\mathbb{R}^d$ (i.e., the Lipschitz image of a bounded set in $\mathbb{R}^{d-1}$), then $\overline{\mathcal{M}}^{d-1}(K)$ exists and coincides with a scalar multiple of $\mathcal{H}^{d-1}(K)$, the $(d-1)$-dimensional Hausdorff measure of K. Given an unbounded set $I \subset (0, \infty)$ and two families of real numbers $(a_i)_{i \in I}, (b_i)_{i \in I}$, we use the Landau notation $a_i = O(b_i)$ to indicate that $\limsup_{i \in I, i \to \infty} |a_i|/|b_i| < \infty$. If $b_i = O(a_i)$ we write $a_i = \Omega(b_i)$, whereas if $a_i = O(b_i)$ and $b_i = O(a_i)$ we write $a_i = \Theta(b_i)$.

Theorem 2.3. *Let $\mathbb{X} \subset \mathbb{R}^d$ be full-dimensional, let $\mathbb{Q}$ have a bounded density with respect to Lebesgue measure and let the conditions of Theorem 2.1 prevail with $\gamma = d$.*
(a) Let K be a full-dimensional compact subset of $\mathbb{X}$ with $\overline{\mathcal{M}}^{d-1}(\partial K) < \infty$. If $\mathrm{Var}\, H_s = \Omega(s)$, resp. $\mathrm{Var}\, H'_n = \Omega(n)$, then there is a constant $c \in (0, \infty)$ such that

$$d_K\left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{\mathrm{Var}\, H_s}}, N\right) \leq \frac{c}{\sqrt{s}}, \quad s \geq 1, \quad \textit{resp.} \quad d_K\left(\frac{H'_n - \mathbb{E}\, H'_n}{\sqrt{\mathrm{Var}\, H'_n}}, N\right) \leq \frac{c}{\sqrt{n}}, \quad n \geq 9. \tag{2.17}$$

(b) Let K be a $(d-1)$-dimensional compact subset of $\mathbb{X}$ with $\overline{\mathcal{M}}^{d-1}(K) < \infty$. If $\mathrm{Var}\, H_s = \Omega(s^{(d-1)/d})$, resp. $\mathrm{Var}\, H'_n = \Omega(n^{(d-1)/d})$, then there is a constant $c \in (0, \infty)$ such that

$$d_K\left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{\mathrm{Var}\, H_s}}, N\right) \leq \frac{c}{s^{\frac{1}{2} - \frac{1}{2d}}}, \quad s \geq 1, \quad \textit{resp.} \quad d_K\left(\frac{H'_n - \mathbb{E}\, H'_n}{\sqrt{\mathrm{Var}\, H'_n}}, N\right) \leq \frac{c}{n^{\frac{1}{2} - \frac{1}{2d}}}, \quad n \geq 9. \tag{2.18}$$

Remarks. (i) *Comparing* (2.17) *with existing results.* The results at (2.17) are applicable in the setting of volume order scaling of the variances, i.e., when the variances of H_s and H'_n exhibit scaling proportional to s and n. The rate for Poisson input in (2.17) significantly improves upon the rate given by Theorem 2.1 of [31] (see also Lemma 4.4 of [26]), Corollary 3.1 of [5], and Theorem 2.3 in [28], which all contain extraneous logarithmic factors and which rely on dependency graph methods. The rate in (2.17) for binomial input is new, as up to now there are no explicit general rates of normal convergence for sums of stabilizing score functions ξ_n of binomial input.

(ii) *Comparing* (2.18) *with existing results.* The rates at (2.18) are relevant for statistics with surface area rescaling of the variances, i.e., when the variance of H_s (resp. H'_n) exhibits scaling proportional to $s^{1-1/d}$ (resp. $n^{1-1/d}$). These rates both improve and extend upon the rates given in the main result (Theorem 1.3) in [50]. First, in the case of Poisson input, the rates remove the logarithmic factors present in Theorem 1.3 of [50]. Second, we obtain rates of normal convergence for binomial input, whereas [50] does not treat this situation.

(iii) *Extensions to random measures.* Up to a constant factor, the rates of normal convergence in Theorem 2.1, Corollary 2.2, and Theorem 2.3 hold for the non-linear statistics $H_s(f) = \sum_{x\in\mathcal{P}_s} f(x)\xi_s(x,\mathcal{P}_s)$ and $H'_n(f) = \sum_{x\in\mathcal{X}_n} f(x)\xi_n(x,\mathcal{X}_n)$, obtained by integrating the random measures $\sum_{x\in\mathcal{P}_s}\xi_s(x,\mathcal{P}_s)\delta_x$ and $\sum_{x\in\mathcal{X}_n}\xi_n(x,\mathcal{X}_n)\delta_x$ with a bounded measurable test function f on $\mathbb{X}$. For example, if $K=\mathbb{X}$, $\mathrm{Var}(H_s(f)) = \Omega(s)$, and $\mathrm{Var}(H'_n(f)) = \Omega(n)$, then there is a constant $c\in(0,\infty)$ such that

$$d_K\left(\frac{H_s(f)-\mathbb{E}\,H_s(f)}{\sqrt{\mathrm{Var}\,H_s(f)}}, N\right) \le \frac{c}{\sqrt{s}}, \quad s\ge 1, \tag{2.19}$$

and

$$d_K\left(\frac{H'_n(f)-\mathbb{E}\,H'_n(f)}{\sqrt{\mathrm{Var}\,H'_n(f)}}, N\right) \le \frac{c}{\sqrt{n}}, \quad n\ge 9. \tag{2.20}$$

The rate (2.19) improves upon the main result (Theorem 2.1) of [31] whereas the rate (2.20) is new.

(iv) *Extensions to the Wasserstein distance.* All quantitative bounds presented in this section also hold for the Wasserstein distance (see also the discussion at the end of Section 3). The Wasserstein distance between random variables Y and Z with $\mathbb{E}\,|Y|, \mathbb{E}\,|Z| < \infty$ is given by

$$d_W(Y,Z) := \sup_{h\in\mathrm{Lip}(1)} |\mathbb{E}\,h(Y) - \mathbb{E}\,h(Z)|, \tag{2.21}$$

where Lip(1) stands for the set of all functions $h:\mathbb{R}\to\mathbb{R}$ whose Lipschitz constant is at most one. Since we believe that the Kolmogorov distance d_K is more prominent than the Wasserstein distance, the applications in Section 5 are formulated only for d_K.

(v) *Subsets without influence.* Assume that there is a measurable set $\tilde{\mathbb{X}}\subset\mathbb{X}$ such that the scores satisfy

$$\xi_s(x,\mathcal{M}) = \mathbf{1}_{\{x\in\tilde{\mathbb{X}}\}}\xi_s(x,\mathcal{M}\cap\tilde{\mathbb{X}}), \quad \mathcal{M}\in\mathbf{N}, x\in\mathcal{M}, s\ge 1,$$

where $\mathcal{M}\cap\tilde{\mathbb{X}}$ stands for the restriction of the point configuration $\mathcal{M}$ to $\tilde{X}$. In other words, the sum of scores $\sum_{x\in\mathcal{M}}\xi_s(x,\mathcal{M})$ only depends on the points of $\mathcal{M}$ which belong to $\tilde{\mathbb{X}}$. In this case our previous results are still valid if the assumptions (2.1)-(2.9) hold for all $x\in\tilde{\mathbb{X}}$.

(vi) *Null sets.* In our assumptions (2.1)-(2.9) we require, for simplicity, that some inequalities are satisfied for all $x\in\mathbb{X}$. In case that these only hold for $\mathbb{Q}$-a.e. $x\in\mathbb{X}$, our results are still true. This also applies to comment (v).

3 Malliavin-Stein bounds

For any measurable $f : \mathbf{N} \to \mathbb{R}$ and $\mathcal{M} \in \mathbf{N}$ we define

$$D_{\hat{x}} f(\mathcal{M}) = f(\mathcal{M} \cup \{\hat{x}\}) - f(\mathcal{M}), \quad \hat{x} \in \widehat{\mathbb{X}},$$

and

$$D^2_{\hat{x}_1,\hat{x}_2} f(\mathcal{M}) = f(\mathcal{M} \cup \{\hat{x}_1, \hat{x}_2\}) - f(\mathcal{M} \cup \{\hat{x}_1\}) - f(\mathcal{M} \cup \{\hat{x}_2\}) + f(\mathcal{M}), \quad \hat{x}_1, \hat{x}_2 \in \widehat{\mathbb{X}}.$$

Our key tool for the proof of the bound (2.11) is the following marked version of a result from [23] (see Proposition 1.4 and Theorem 6.1 in [23]) for square integrable Poisson functionals.

Theorem 3.1. *Let $s > 0$ and let $f : \mathbf{N} \to \mathbb{R}$ be measurable with $\mathbb{E} f(\mathcal{P}_s)^2 < \infty$. Assume there are constants $c, p \in (0, \infty)$ such that*

$$\mathbb{E} |D_{(x,M_x)} f(\mathcal{P}_s \cup \{(\mathcal{A}, M_{\mathcal{A}})\})|^{4+p} \le c, \quad \mathbb{Q}\text{-a.e. } x \in \mathbb{X}, \mathcal{A} \subset \mathbb{X}, |\mathcal{A}| \le 1. \tag{3.1}$$

Let $F := f(\mathcal{P}_s)$. Then there is a constant $C := C(c, p) \in (0, \infty)$ such that

$$d_K\left(\frac{F - \mathbb{E} F}{\sqrt{\operatorname{Var} F}}, N\right) \le C(S_1 + S_2 + S_3), \tag{3.2}$$

with

$$\begin{aligned}
\Gamma_s :=& s \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} f(\mathcal{P}_s) \neq 0)^{\frac{p}{8+2p}} \, \mathbb{Q}(\mathrm{d}x), \\
\psi_s(x_1, x_2) :=& \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} f(\mathcal{P}_s) \neq 0)^{\frac{p}{16+4p}}, \\
S_1 :=& \frac{s}{\operatorname{Var} F} \sqrt{\int_{\mathbb{X}^2} \psi_s(x_1, x_2)^2 \, \mathbb{Q}^2(\mathrm{d}(x_1, x_2))}, \\
S_2 :=& \frac{s^{3/2}}{\operatorname{Var} F} \sqrt{\int_{\mathbb{X}} \left(\int_{\mathbb{X}} \psi_s(x_1, x_2) \, \mathbb{Q}(\mathrm{d}x_2) \right)^2 \mathbb{Q}(\mathrm{d}x_1)}, \\
S_3 :=& \frac{\sqrt{\Gamma_s}}{\operatorname{Var} F} + \frac{2\Gamma_s}{(\operatorname{Var} F)^{3/2}} + \frac{\Gamma_s^{5/4} + 2\Gamma_s^{3/2}}{(\operatorname{Var} F)^2}.
\end{aligned}$$

Proof. In case that there are no marks, this is Theorem 6.1 in [23]. The marked version can be obtained in the following way: In Theorem 1.2 in [23] one can use the product form of $\widehat{\mathbb{Q}}$ and Hölder's inequality to bring the marks under the expectations. Evaluating this new bound along the lines of the proof of Theorem 6.1 in [23] yields (3.2). □

For the case of binomial input, we do not have the same ready-made bounds at our disposal. We fill this lacuna with the following analogous bound, bringing [22] and [23] into a satisfying alignment.

Theorem 3.2. *Let $n \geq 3$ and let $f : \mathbf{N} \to \mathbb{R}$ be measurable with $\mathbb{E}\, f(\mathcal{X}_n)^2 < \infty$. Assume that there are constants $c, p \in (0, \infty)$ such that*

$$\mathbb{E}\, |D_{(x,M_x)} f(\mathcal{X}_{n-1-|\mathcal{A}|} \cup \{(\mathcal{A}, M_{\mathcal{A}})\})|^{4+p} \leq c, \quad \mathbb{Q}\text{-a.e. } x \in \mathbb{X}, \mathcal{A} \subset \mathbb{X}, |\mathcal{A}| \leq 2. \tag{3.3}$$

Let $F := f(\mathcal{X}_n)$. Then there is a constant $C := C(c, p) \in (0, \infty)$ such that

$$d_K\left(\frac{F - \mathbb{E}\, F}{\sqrt{\operatorname{Var} F}}, N\right) \leq C(S_1' + S_2' + S_3'), \tag{3.4}$$

with

$$\begin{aligned}
\Gamma_n' :=& n \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} f(\mathcal{X}_{n-1}) \neq 0)^{\frac{p}{8+2p}}\, \mathbb{Q}(\mathrm{d}x),\\
\psi_n'(x, x') :=& \sup_{\mathcal{A} \subset \mathbb{X}: |\mathcal{A}| \leq 1} \mathbb{P}(D^2_{(x,M_x),(x',M_{x'})} f(\mathcal{X}_{n-2-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0)^{\frac{p}{8+2p}},\\
S_1' :=& \frac{n}{\operatorname{Var} F} \sqrt{\int_{\mathbb{X}^2} \psi_n'(x, x')\, \mathbb{Q}^2(\mathrm{d}(x, x'))},\\
S_2' :=& \frac{n^{3/2}}{\operatorname{Var} F} \sqrt{\int_{\mathbb{X}} \left(\int_{\mathbb{X}} \psi_n'(x, x')\, \mathbb{Q}(\mathrm{d}x')\right)^2 \mathbb{Q}(\mathrm{d}x)},\\
S_3' :=& \frac{\Gamma_n'}{\sqrt{\operatorname{Var} F}^3} + \frac{\sqrt{\Gamma_n'}^3 + \Gamma_n'}{(\operatorname{Var} F)^2} + \frac{\sqrt{\Gamma_n'}}{\operatorname{Var} F}.
\end{aligned}$$

Before proving Theorem 3.2 we require two auxiliary results, the first of which involves some additional notation. For a measurable $f : \mathbf{N} \to \mathbb{R}$ extend the notation $f(x_1, \ldots, x_q) := f(\{x_1, \ldots, x_q\})$ for $x_1, \ldots, x_q \in \mathbb{X}$.

For a fixed $n \geq 1$ let $X := (X_1, \ldots, X_n)$, where $X_1, \ldots, X_n$ are independent random elements in $\widehat{\mathbb{X}}$ distributed according to $\widehat{\mathbb{Q}}$. Let $X', \widetilde{X}$ be independent copies of X. We write $U \overset{a.s.}{=} V$ if two variables U and V satisfy $\mathbb{P}(U = V) = 1$. In the vocabulary of [22], a random vector $Y := (Y_1, \ldots, Y_n)$ is a recombination of $\{X, X', \widetilde{X}\}$ if for each $1 \leq i \leq n$, either $Y_i \overset{a.s.}{=} X_i, Y_i \overset{a.s.}{=} X_i'$ or $Y_i \overset{a.s.}{=} \widetilde{X}_i$. For a vector $x = (x_1, \ldots, x_p) \in \widehat{\mathbb{X}}^p$, and indices $I := \{i_1, \ldots, i_q\} \subset [p] := \{1, 2, ..., p\}$, define $x^{i_1,\ldots,i_q} := (x_j, j \notin I)$, the vector x with the components indexed by I removed. For $i, j \in [n]$, introduce the index derivatives

$$\begin{aligned}
\mathbf{D}_i f(X) &:= f(X) - f(X^i)\\
\mathbf{D}^2_{i,j} f(X) &:= f(X) - f(X^i) - f(X^j) + f(X^{i,j}) = \mathbf{D}^2_{j,i} f(X).
\end{aligned}$$

We note that the derivatives D and $\mathbf{D}$ obey the relation $\mathbf{D}_i f(X) = D_{X_i} f(\mathcal{X}_n^i)$.

We introduce, for n-dimensional random vectors Y, Y' and Z,

$$\begin{aligned}
\gamma_{Y,Z}(f) &:= \mathbb{E}\left[\mathbf{1}_{\{\mathbf{D}^2_{1,2}f(Y)\neq 0\}}\mathbf{D}_2 f(Z)^4\right]\\
\gamma'_{Y,Y',Z}(f) &:= \mathbb{E}\left[\mathbf{1}_{\{\mathbf{D}^2_{1,2}f(Y)\neq 0,\, \mathbf{D}^2_{1,3}f(Y')\neq 0\}}\mathbf{D}_2 f(Z)^4\right]\\
B_n(f) &:= \sup\{\gamma_{Y,Z}(f);\ Y, Z \text{ recombinations of } \{X, X', \widetilde{X}\}\}\\
B'_n(f) &:= \sup\{\gamma'_{Y,Y',Z}(f);\ Y, Y', Z \text{ recombinations of } \{X, X', \widetilde{X}\}\}.
\end{aligned}$$

Theorem 5.1 of [22], simplified by [22, Remark 5.2] and [22, Proposition 5.3], gives the following:

Theorem 3.3. *Let $n \geq 2$, $f : \mathbf{N} \to \mathbb{R}$ measurable with $\mathbb{E}\, f(\mathcal{X}_n)^2 < \infty$, and $F := f(\mathcal{X}_n)$. Then there is a constant $c_0 \in (0, \infty)$, depending neither on n nor f, such that*

$$d_K\left(\frac{F - \mathbb{E}\, F}{\sqrt{\operatorname{Var} F}}, N\right) \leq c_0 \left[\frac{\sqrt{n}}{\operatorname{Var} F}\left(\sqrt{nB_n(f)} + \sqrt{n^2 B'_n(f)} + \sqrt{\mathbb{E}\, \mathbf{D}_1 f(X)^4}\right)\right. \tag{3.5}$$
$$\left. + \sup_Y \frac{n}{(\operatorname{Var} F)^2}\mathbb{E}\, |(f(X) - \mathbb{E}\, F)(\mathbf{D}_1 f(Y))^3| + \frac{n}{(\operatorname{Var} F)^{\frac{3}{2}}}\mathbb{E}\, |\mathbf{D}_1 f(X)|^3\right],$$

where the $\sup_Y$ *runs over recombinations Y of $\{X, X', \tilde{X}\}$.*

To control the fourth centered moment of $F := f(\mathcal{X}_n)$, we use the following bound. For a similar bound for Poisson functionals we refer to [23, Lemma 4.2].

Lemma 3.4. *For a measurable $f : \mathbf{N} \to \mathbb{R}$, $n \in \mathbb{N}$ and $F := f(\mathcal{X}_n)$ assume that* $\operatorname{Var} F = 1$. *Then*

$$\mathbb{E}\,(F - \mathbb{E}\, F)^4 \leq 9 \max\left\{\left(32n \int_{\mathbb{X}} \sqrt{\mathbb{E}\,(D_{(x,M_x)} f(\mathcal{X}_{n-1}))^4}\, \mathbb{Q}(\mathrm{d}x)\right)^2, 4n\mathbb{E}\,(\mathbf{D}_1 f(\mathcal{X}_n))^4 + 1\right\}.$$

Proof. The Efron-Stein inequality implies that for measurable $g : \mathbf{N} \to \mathbb{R}$ and $n \in \mathbb{N}$ such that $\mathbb{E}\, g(\mathcal{X}_n)^2 < \infty$,

$$\operatorname{Var} g(\mathcal{X}_n) \leq 2n\mathbb{E}\,(\mathbf{D}_1 g(\mathcal{X}_n))^2.$$

Using $\operatorname{Var} F = 1$ and the Efron-Stein bound in this order gives

$$\mathbb{E}\,(F - \mathbb{E}\, F)^4 = \operatorname{Var}\left((f(\mathcal{X}_n) - \mathbb{E}\, F)^2\right) + 1 \leq 2n\mathbb{E}\left(\mathbf{D}_1((f(\mathcal{X}_n) - \mathbb{E}\, F)^2)\right)^2 + 1.$$

Combining the identity

$$\begin{aligned}
\mathbf{D}_1(g(\mathcal{X}_n)^2) = g(\mathcal{X}_n)^2 - g(\mathcal{X}_n^1)^2 &= (g(\mathcal{X}_n^1) + \mathbf{D}_1 g(\mathcal{X}_n))^2 - g(\mathcal{X}_n^1)^2\\
&= 2g(\mathcal{X}_n^1)\mathbf{D}_1 g(\mathcal{X}_n) + (\mathbf{D}_1 g(\mathcal{X}_n))^2
\end{aligned}$$

with Jensen's inequality, we obtain

$$\begin{aligned}\mathbb{E}\,(F-\mathbb{E}\,F)^4 &\le 2n\mathbb{E}\,\Big[\big(2\mathbf{D}_1 f(\mathcal{X}_n)(f(\mathcal{X}_n^1)-\mathbb{E}\,F)+(\mathbf{D}_1 f(\mathcal{X}_n))^2\big)^2\Big]+1\\ &\le 4n\mathbb{E}\,\big[4(\mathbf{D}_1 f(\mathcal{X}_n))^2(f(\mathcal{X}_n^1)-\mathbb{E}\,F)^2+(\mathbf{D}_1 f(\mathcal{X}_n))^4\big]+1.\end{aligned}$$

Hölder's inequality and a combination of the triangle inequality and Jensen's inequality imply that

$$\begin{aligned}&\mathbb{E}\,(\mathbf{D}_1 f(\mathcal{X}_n))^2(f(\mathcal{X}_n^1)-\mathbb{E}\,F)^2\\ &\le \int_{\widehat{\mathbb{X}}}\sqrt{\mathbb{E}\,(f(\mathcal{X}_n^1\cup\{y\})-f(\mathcal{X}_n^1))^4}\,\widehat{\mathbb{Q}}(\mathrm{d}y)\sqrt{\mathbb{E}\,(f(\mathcal{X}_n^1)-\mathbb{E}\,F)^4}\\ &\le \int_{\mathbb{X}}\sqrt{\mathbb{E}\,(D_{(x,M_x)}f(\mathcal{X}_{n-1}))^4}\,\mathbb{Q}(\mathrm{d}x)\ 2\big(\sqrt{\mathbb{E}\,(f(\mathcal{X}_n)-\mathbb{E}\,F)^4}+\sqrt{\mathbb{E}\,(\mathbf{D}_1 f(\mathcal{X}_n))^4}\big).\end{aligned}$$

Combining the above estimates we arrive at

$$\begin{aligned}\mathbb{E}\,(F-\mathbb{E}\,F)^4 \le 32n\int_{\mathbb{X}}\sqrt{\mathbb{E}\,(D_{(x,M_x)}f(\mathcal{X}_{n-1}))^4}\,\mathbb{Q}(\mathrm{d}x)\big(\sqrt{\mathbb{E}\,(F-\mathbb{E}\,F)^4}+\sqrt{\mathbb{E}\,(\mathbf{D}_1 f(\mathcal{X}_n))^4}\big)\\ +4n\mathbb{E}(\mathbf{D}_1 f(\mathcal{X}_n))^4+1,\end{aligned}$$

which implies the asserted inequality. □

Given Lemma 3.4, we deduce Theorem 3.2 from Theorem 3.3 as follows.

Proof of Theorem 3.2. It suffices to show that each of the five terms in (3.5) is bounded by a scalar multiple of S_1', S_2', or S_3'. We first show that the terms in (3.5) involving $B_n(f)$ and $B_n'(f)$ are bounded resp. by scalar multiples of S_1' and S_2'. Let us estimate first $B_n(f)$. By $\widehat{\mathbb{Q}}^{Y_1,Y_2,Z_1,Z_2}$ we denote the joint probability measure of Y_1,Y_2,Z_1,Z_2 and by $\mathbb{Q}^{Y_1,Y_2,Z_1,Z_2}$ the joint probability measure of Y_1,Y_2,Z_1,Z_2 without marks. By Hölder's inequality, the fact that $\widehat{\mathbb{Q}}^{Y_1,Y_2,Z_1,Z_2}$ factorizes into $\mathbb{Q}^{Y_1,Y_2,Z_1,Z_2}$ and a part controlling the marks, the independence of Y_1,Y_2, and (3.3), we obtain that

$$\begin{aligned}\gamma_{Y,Z}(f) &= \mathbb{E}\,[\mathbf{1}_{\{\mathbf{D}^2_{1,2}f(Y)\neq 0\}}(\mathbf{D}_2 f(Z))^4]\\ &= \int_{\widehat{\mathbb{X}}^4}\mathbb{E}\,[\mathbf{1}_{\{D^2_{\hat{y}_1,\hat{y}_2}f(Y^{1,2})\neq 0\}}(D_{\hat{z}_2}f(Z^{1,2}\cup\{\hat{z}_1\}))^4]\,\widehat{\mathbb{Q}}^{Y_1,Y_2,Z_1,Z_2}(\mathrm{d}(\hat{y}_1,\hat{y}_2,\hat{z}_1,\hat{z}_2))\\ &\le \int_{\widehat{\mathbb{X}}^4}\mathbb{P}(D^2_{\hat{y}_1,\hat{y}_2}f(Y^{1,2})\neq 0)^{\frac{p}{4+p}}\,\mathbb{E}\,[|D_{\hat{z}_2}f(Z^{1,2}\cup\{\hat{z}_1\}))|^{4+p}]^{\frac{4}{4+p}}\\ &\qquad\qquad \widehat{\mathbb{Q}}^{Y_1,Y_2,Z_1,Z_2}(\mathrm{d}(\hat{y}_1,\hat{y}_2,\hat{z}_1,\hat{z}_2))\\ &\le \int_{\mathbb{X}^4}\mathbb{P}(D^2_{(y_1,M_{y_1}),(y_2,M_{y_2})}f(Y^{1,2})\neq 0)^{\frac{p}{4+p}}\,\mathbb{E}\,[|D_{(z_2,M_{z_2})}f(Z^{1,2}\cup\{(z_1,M_{z_1})\}))|^{4+p}]^{\frac{4}{4+p}}\\ &\qquad\qquad \mathbb{Q}^{Y_1,Y_2,Z_1,Z_2}(\mathrm{d}(y_1,y_2,z_1,z_2))\\ &\le c^{\frac{4}{4+p}}\int_{\mathbb{X}^2}\mathbb{P}(D^2_{(y_1,M_{y_1}),(y_2,M_{y_2})}f(\mathcal{X}_{n-2})\neq 0)^{\frac{p}{4+p}}\,\mathbb{Q}^2(\mathrm{d}(y_1,y_2)).\end{aligned}$$

This implies that

$$\gamma_{Y,Z}(f) \leq c^{\frac{4}{4+p}} \int_{\mathbb{X}^2} \psi'_n(y_1, y_2)\, \mathbb{Q}^2(\mathrm{d}(y_1, y_2)),$$

which gives the desired bound

$$\frac{\sqrt{n}}{\operatorname{Var} F} \sqrt{n B_n(f)} \leq c^{\frac{2}{4+p}} \frac{n}{\operatorname{Var} F} \sqrt{\int_{\mathbb{X}^2} \psi'_n(x, x')\, \mathbb{Q}^2(\mathrm{d}(x, x'))} \leq C(c, p) S'_1.$$

To estimate $B'_n(f)$, let $\widehat{\mathbb{Q}}^{(Y_1,\dots,Z_3)}$ be the joint probability measure of

$$(Y_1, \dots, Y_3, Y'_1, \dots, Y'_3, Z_1, \dots, Z_3)$$

and let $\mathbb{Q}^{(Y_1,\dots,Z_3)}$ be the corresponding probability measure without marks. By similar arguments as above, we obtain that

$$\begin{aligned}
\gamma'_{Y,Y',Z}(f) &= \mathbb{E}\,[\mathbf{1}_{\{\mathbf{D}^2_{1,2} f(Y) \neq 0, \mathbf{D}^2_{1,3} f(Y') \neq 0\}} (\mathbf{D}_2 f(Z))^4] \\
&= \int_{\widehat{\mathbb{X}}^9} \mathbb{E}\,[\mathbf{1}_{\{D^2_{\hat{y}_1,\hat{y}_2} f(Y^{1,2,3} \cup \{\hat{y}_3\}) \neq 0\}} \mathbf{1}_{\{D^2_{\hat{y}'_1,\hat{y}'_3} f(Y'^{1,2,3} \cup \{\hat{y}'_2\}) \neq 0\}} (D_{\hat{z}_2} f(Z^{1,2,3} \cup \{\hat{z}_1, \hat{z}_3\}))^4] \\
&\qquad\qquad \widehat{\mathbb{Q}}^{Y_1,\dots,Z_3}(\mathrm{d}(\hat{y}_1, \dots, \hat{z}_3)) \\
&\leq \int_{\widehat{\mathbb{X}}^9} \mathbb{P}(D^2_{\hat{y}_1,\hat{y}_2} f(Y^{1,2,3} \cup \{\hat{y}_3\}) \neq 0)^{\frac{p}{8+2p}} \mathbb{P}(D^2_{\hat{y}'_1,\hat{y}'_3} f(Y'^{1,2,3} \cup \{\hat{y}'_2\}) \neq 0)^{\frac{p}{8+2p}} \\
&\qquad\qquad \mathbb{E}\,[|D_{\hat{z}_2} f(Z^{1,2,3} \cup \{\hat{z}_1, \hat{z}_3\})|^{4+p}]^{\frac{4}{4+p}}\, \widehat{\mathbb{Q}}^{Y_1,\dots,Z_3}(\mathrm{d}(\hat{y}_1, \dots, \hat{z}_3)) \\
&\leq \int_{\mathbb{X}^9} \mathbb{P}(D^2_{(y_1,M_{y_1}),(y_2,M_{y_2})} f(\mathcal{X}_{n-3} \cup \{(y_3, M_{y_3})\}) \neq 0)^{\frac{p}{8+2p}} \\
&\qquad \mathbb{P}(D^2_{(y'_1,M_{y'_1}),(y'_3,M_{y'_3})} f(\mathcal{X}_{n-3} \cup \{(y'_2, M_{y'_2})\}) \neq 0)^{\frac{p}{8+2p}} \\
&\qquad \mathbb{E}\,[|D_{(z_2,M_{z_2})} f(\mathcal{X}_{n-3} \cup \{(z_1, M_{z_1}), (z_3, M_{z_3})\})|^{4+p}]^{\frac{4}{4+p}}\, \mathbb{Q}^{Y_1,\dots,Z_3}(\mathrm{d}(y_1, \dots, z_3)) \\
&\leq c^{\frac{4}{4+p}} \int_{\mathbb{X}^9} \psi'_n(y_1, y_2)\, \psi'_n(y'_1, y'_3)\, \mathbb{Q}^{Y_1,\dots,Z_3}(\mathrm{d}(y_1, \dots, z_3)).
\end{aligned}$$

If $Y_1 \overset{a.s.}{=} Y'_1$, this simplifies to

$$\gamma'_{Y,Y',Z}(f) \leq c^{\frac{4}{4+p}} \int_{\mathbb{X}} \left(\int_{\mathbb{X}} \psi'_n(x, x')\, \mathbb{Q}(\mathrm{d}x') \right)^2 \mathbb{Q}(\mathrm{d}x).$$

If Y_1 and Y'_1 are independent, the Cauchy-Schwarz inequality leads to

$$\gamma'_{Y,Y',Z}(f) \leq c^{\frac{4}{4+p}} \left(\int_{\mathbb{X}^2} \psi'_n(x, x')\, \mathbb{Q}^2(\mathrm{d}(x, x')) \right)^2 \leq c^{\frac{4}{4+p}} \int_{\mathbb{X}} \left(\int_{\mathbb{X}} \psi'_n(x, x')\, \mathbb{Q}(\mathrm{d}x') \right)^2 \mathbb{Q}(\mathrm{d}x).$$

Thus, we obtain the desired bound

$$\frac{\sqrt{n}}{\operatorname{Var} F} \sqrt{n^2 \gamma'_{Y,Y',Z}(f)} \leq c^{\frac{2}{4+p}} \frac{n^{\frac{3}{2}}}{\operatorname{Var} F} \sqrt{\int_{\mathbb{X}} \left(\int_{\mathbb{X}} \psi'_n(x, x')\, \mathbb{Q}(\mathrm{d}x') \right)^2 \mathbb{Q}(\mathrm{d}x)} \leq C(c, p) S'_2.$$

We now show that the remaining terms in (3.5) are bounded by a scalar multiple of S_3'. For $1 \le m \le 4$ and $\mathbb{Q}$-a.e. $x \in \mathbb{X}$, Hölder's inequality and (3.3) lead to

$$\begin{aligned}\mathbb{E}\,|D_{(x,M_x)}f(\mathcal{X}_{n-1})|^m &\le \mathbb{E}\,[|D_{(x,M_x)}f(\mathcal{X}_{n-1})|^{4+p}]^{\frac{m}{4+p}}\,\mathbb{P}(D_{(x,M_x)}f(\mathcal{X}_{n-1}) \neq 0)^{\frac{4+p-m}{4+p}} \\ &\le c^{\frac{m}{4+p}}\mathbb{P}(D_{(x,M_x)}f(\mathcal{X}_{n-1}) \neq 0)^{\frac{p}{4+p}},\end{aligned} \tag{3.6}$$

where we have also used that $\frac{4+p-m}{4+p} \ge \frac{p}{4+p}$. For $1 \le m \le 4$ and $u \in [1/2, 1]$ we derive from (3.6) that

$$\int_{\mathbb{X}} \mathbb{E}\left[|D_{(x,M_x)}f(\mathcal{X}_{n-1})|^m\right]^u \mathbb{Q}(\mathrm{d}x) \le c^{\frac{mu}{4+p}} \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)}f(\mathcal{X}_{n-1}) \neq 0)^{\frac{up}{4+p}}\,\mathbb{Q}(\mathrm{d}x) \le c^{\frac{mu}{4+p}}\frac{\Gamma_n'}{n}. \tag{3.7}$$

This implies immediately that, for $1 \le m \le 4$,

$$\mathbb{E}\,|\mathbf{D}_1 f(X)|^m \le c^{\frac{m}{4+p}} \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)}f(\mathcal{X}_{n-1}) \neq 0)^{\frac{p}{4+p}}\,\mathbb{Q}(\mathrm{d}x) \le c^{\frac{m}{4+p}}\frac{\Gamma_n'}{n}.$$

This gives for $m = 3$ and $m = 4$ that the third and fifth terms in (3.5) are bounded by

$$\frac{n\mathbb{E}\,|\mathbf{D}_1 f(X)|^3}{(\mathrm{Var}\,F)^{\frac{3}{2}}} + \frac{\sqrt{n}\sqrt{\mathbb{E}\,\mathbf{D}_1 f(X)^4}}{\mathrm{Var}\,F} \le \frac{c^{\frac{3}{4+p}}\Gamma_n'}{(\mathrm{Var}\,F)^{\frac{3}{2}}} + \frac{c^{\frac{2}{4+p}}\sqrt{\Gamma_n'}}{\mathrm{Var}\,F} \le C(c,p)S_3'.$$

Lastly, we bound the fourth term in (3.5) by a scalar multiple of S_3'. Let Y be a recombination of $\{X, X', \tilde{X}\}$. Noting that $Y \overset{(d)}{=} X$, let us estimate

$$\begin{aligned}&\mathbb{E}\,|(f(X) - \mathbb{E}\,F)(\mathbf{D}_1 f(Y))^3| \\ &= \mathbb{E}\,\left|\left(f(X^1) - \mathbb{E}\,F + \mathbf{D}_1 f(X)\right)(\mathbf{D}_1 f(Y))^3\right| \\ &\le \int_{\mathbb{X}} \mathbb{E}\,\left[|f(X^1) - \mathbb{E}\,F|\,|D_{(y_1,M_{y_1})}f(Y^1)|^3\right]\,\mathbb{Q}(\mathrm{d}y_1) + \mathbb{E}\,\left[|\mathbf{D}_1 f(X)|\,|\mathbf{D}_1 f(Y)|^3\right] \\ &\le \mathbb{E}\,[(f(\mathcal{X}_n^1) - \mathbb{E}\,F)^4]^{\frac{1}{4}} \int_{\mathbb{X}} \mathbb{E}\,\left[(D_{(x,M_x)}f(\mathcal{X}_{n-1}))^4\right]^{\frac{3}{4}}\,\mathbb{Q}(\mathrm{d}x) + \mathbb{E}\,(\mathbf{D}_1 f(X))^4 \\ &\le \left(\mathbb{E}\,[(f(\mathcal{X}_n) - \mathbb{E}\,F)^4]^{\frac{1}{4}} + \mathbb{E}\,[(\mathbf{D}_1 f(X))^4]^{\frac{1}{4}}\right) \int_{\mathbb{X}} \mathbb{E}\,\left[(D_{(x,M_x)}f(\mathcal{X}_{n-1}))^4\right]^{\frac{3}{4}}\,\mathbb{Q}(\mathrm{d}x) + c^{\frac{4}{4+p}}\frac{\Gamma_n'}{n}.\end{aligned}$$

By (3.7) we have

$$\int_{\mathbb{X}} \mathbb{E}\,\left[(D_{(x,M_x)}f(\mathcal{X}_{n-1}))^4\right]^{\frac{3}{4}}\,\mathbb{Q}(\mathrm{d}x) \le c^{\frac{3}{4+p}}\frac{\Gamma_n'}{n}.$$

From Lemma 3.4 and (3.7) it follows that

$$\begin{aligned}\frac{\mathbb{E}\,(F - \mathbb{E}\,F)^4}{(\mathrm{Var}\,F)^2} &\le 9\max\Bigg\{\left(\frac{32n}{\mathrm{Var}\,F}\int_{\mathbb{X}}\sqrt{\mathbb{E}\,(D_{(y,M_y)}f(\mathcal{X}_{n-1}))^4}\,\mathbb{Q}(\mathrm{d}y)\right)^2, \\ &\qquad 4n\frac{\mathbb{E}\,(\mathbf{D}_1 f(\mathcal{X}_n))^4}{(\mathrm{Var}\,F)^2} + 1\Bigg\} \\ &\le 9\max\left\{\frac{1024c^{\frac{4}{4+p}}(\Gamma_n')^2}{(\mathrm{Var}\,F)^2}, \frac{4c^{\frac{4}{4+p}}\,\Gamma_n'}{(\mathrm{Var}\,F)^2} + 1\right\}.\end{aligned}$$

All together, the fourth term in (3.5) satisfies the bound

$$
\begin{aligned}
&\frac{n\mathbb{E}\,|(f(X)-\mathbb{E}\,F)(\mathbf{D}_1 f(Y))^3|}{(\operatorname{Var}F)^2}\\
&\le \frac{c^{\frac{4}{4+p}}\Gamma'_n}{(\operatorname{Var}F)^2} + \left(\sqrt{3}\max\left\{\frac{4\cdot\sqrt{2}c^{\frac{1}{4+p}}\sqrt{\Gamma'_n}}{\sqrt{\operatorname{Var}F}}, \frac{\sqrt{2}c^{\frac{1}{4+p}}\Gamma_n'^{\frac{1}{4}}}{\sqrt{\operatorname{Var}F}}+1\right\} + \frac{c^{\frac{1}{4+p}}\Gamma_n'^{\frac{1}{4}}}{n^{\frac{1}{4}}\sqrt{\operatorname{Var}F}}\right)\frac{c^{\frac{3}{4+p}}\Gamma'_n}{(\operatorname{Var}F)^{\frac{3}{2}}}\\
&\le C(c,p)S'_3. \quad \square
\end{aligned}
$$

Remark. The bounds in Theorem 3.1 and Theorem 3.2 are still valid for the Wasserstein distance given in (2.21). This follows from the fact that the underlying bounds in Theorem 6.1 in [23] and Theorem 3.3 (see also Remark 4.3 in [22]) are true for the Wasserstein distance as well.

4 Proofs of Theorem 2.1 and Theorem 2.3

The bounds in Theorems 3.1 and 3.2 are admittedly unwieldy. However when F is a sum of stabilizing score functions, as in (1.1) and (1.2), then the terms on the right-hand side of (3.2) and (3.4) conveniently collapse into the more manageable bounds (2.11) and (2.12), respectively.

We first provide several lemmas giving moment and probability bounds for the first and second order difference operators. Throughout we assume that the hypotheses of Theorem 2.1 are in force. We can assume without loss of generality that $C_{stab} = C_K =: C$, $c_{stab} = c_K =: c$ and $\alpha_{stab} = \alpha_K =: \alpha$.

Lemma 4.1.

(a) For any $x \in \mathbb{X}$ and $r \ge 0$,

$$\mathbb{Q}(B(x,r)) \le \kappa r^{\gamma}. \tag{4.1}$$

(b) For any $\nu > 0$ there is a constant $C_\nu \in (0,\infty)$ such that

$$\int_{\mathbb{X}\setminus B(x,r)} \exp(-(\beta^{1/\gamma}\,\mathrm{d}(x,y))^{\nu})\,\mathbb{Q}(\mathrm{d}y) \le \frac{C_\nu}{\beta}\exp(-(\beta^{1/\gamma}r)^{\nu}/2) \tag{4.2}$$

for all $\beta \ge 1$, $x \in \mathbb{X}$ and $r \ge 0$.

Proof. We prove only (b) since (a) can be proven similarly. For any monotone sequence $(r_n)_{n\in\mathbb{N}}$ with $r_1 > r =: r_0$ and $\lim_{n\to\infty} r_n = \infty$ we have

$$\int_{\mathbb{X}\setminus B(x,r)} \exp(-(\beta^{1/\gamma}\,\mathrm{d}(x,y))^{\nu})\,\mathbb{Q}(\mathrm{d}y) \le \sum_{n=1}^{\infty}\exp(-(\beta^{1/\gamma}r_{n-1})^{\nu})\,\mathbb{Q}(B(x,r_n)\setminus B(x,r_{n-1})).$$

For $\sup_{n\in\mathbb{N}} |r_n - r_{n-1}| \to 0$ assumption (2.1) as well as compactness arguments and the properties of the Riemann integral imply that

$$\int_{\mathbb{X}\setminus B(x,r)} \exp(-(\beta^{1/\gamma}\,\mathrm{d}(x,y))^\nu)\,\mathbb{Q}(\mathrm{d}y) \le \int_r^\infty \exp(-(\beta^{1/\gamma}u)^\nu)\,\kappa\gamma u^{\gamma-1}\,\mathrm{d}u$$
$$= \frac{1}{\beta}\int_{\beta^{1/\gamma}r}^\infty \exp(-w^\nu)\,\kappa\gamma w^{\gamma-1}\,\mathrm{d}w.$$

Now a straightforward computation completes the proof of (b). □

Throughout our proofs we only make use of (4.1) and (4.2) and not of (2.1) so that one could replace the assumption (2.1) by (4.1) and (4.2).

Lemma 4.2. *Let $\mathcal{M} \in \mathbf{N}$ and $\hat{y}, \hat{y}_1, \hat{y}_2 \in \widehat{\mathbb{X}}$. Then, for $s \ge 1$,*

$$D_{\hat{y}} h_s(\mathcal{M}) = \xi_s(\hat{y}, \mathcal{M}\cup\{\hat{y}\}) + \sum_{x\in\mathcal{M}} D_{\hat{y}}\xi_s(x,\mathcal{M})$$
$$D^2_{\hat{y}_1,\hat{y}_2} h_s(\mathcal{M}) = D_{\hat{y}_1}\xi_s(\hat{y}_2, \mathcal{M}\cup\{\hat{y}_2\}) + D_{\hat{y}_2}\xi_s(\hat{y}_1, \mathcal{M}\cup\{\hat{y}_1\}) + \sum_{x\in\mathcal{M}} D^2_{\hat{y}_1,\hat{y}_2}\xi_s(x,\mathcal{M}).$$

Proof. In the following let $h := h_s$ and $\xi := \xi_s$. By the definition of the difference operator we have that

$$\begin{aligned} D_{\hat{y}} h(\mathcal{M}) &= \sum_{x\in\mathcal{M}\cup\{\hat{y}\}} \xi(x, \mathcal{M}\cup\{\hat{y}\}) - \sum_{x\in\mathcal{M}} \xi(x,\mathcal{M}) \\ &= \xi(\hat{y}, \mathcal{M}\cup\{\hat{y}\}) + \sum_{x\in\mathcal{M}} \big(\xi(x, \mathcal{M}\cup\{\hat{y}\}) - \xi(x,\mathcal{M})\big) \\ &= \xi(\hat{y}, \mathcal{M}\cup\{\hat{y}\}) + \sum_{x\in\mathcal{M}} D_{\hat{y}}\xi(x,\mathcal{M}). \end{aligned}$$

For the second-order difference operator this implies that

$$\begin{aligned} &D^2_{\hat{y}_1,\hat{y}_2} h(\mathcal{M}) \\ &= \xi(\hat{y}_2, \mathcal{M}\cup\{\hat{y}_1,\hat{y}_2\}) + \sum_{x\in\mathcal{M}\cup\{\hat{y}_1\}} D_{\hat{y}_2}\xi(x, \mathcal{M}\cup\{\hat{y}_1\}) - \xi(\hat{y}_2, \mathcal{M}\cup\{\hat{y}_2\}) - \sum_{x\in\mathcal{M}} D_{\hat{y}_2}\xi(x,\mathcal{M}) \\ &= D_{\hat{y}_1}\xi(\hat{y}_2, \mathcal{M}\cup\{\hat{y}_2\}) + D_{\hat{y}_2}\xi(\hat{y}_1, \mathcal{M}\cup\{\hat{y}_1\}) + \sum_{x\in\mathcal{M}} \big(D_{\hat{y}_2}\xi(x, \mathcal{M}\cup\{\hat{y}_1\}) - D_{\hat{y}_2}\xi(x,\mathcal{M})\big) \\ &= D_{\hat{y}_1}\xi(\hat{y}_2, \mathcal{M}\cup\{\hat{y}_2\}) + D_{\hat{y}_2}\xi(\hat{y}_1, \mathcal{M}\cup\{\hat{y}_1\}) + \sum_{x\in\mathcal{M}} D^2_{\hat{y}_1,\hat{y}_2}\xi(x,\mathcal{M}), \end{aligned}$$

which completes the proof. □

If a point $\hat{y} \in \widehat{\mathbb{X}}$ is inserted into $\mathcal{M} \in \mathbf{N}$ at a distance exceeding the stabilization radius at $\hat{x} \in \mathcal{M}$, then the difference operator $D_{\hat{y}}$ of the score at $\hat{x}$ vanishes, as seen by the next lemma.

Lemma 4.3. *Let $\mathcal{M} \in \mathbf{N}$, $(x, m_x) \in \mathcal{M}$, $\widehat{\mathcal{A}} \subset \widehat{\mathbb{X}}$ with $|\widehat{\mathcal{A}}| \leq 6$, $y, y_1, y_2 \in \mathbb{X}$ and $m_y, m_{y_1}, m_{y_2} \in \mathbb{M}$. Then, for $s \geq 1$,*

$$D_{(y,m_y)}\xi_s((x, m_x), \mathcal{M} \cup \widehat{\mathcal{A}}) = 0 \quad \text{if} \quad R_s((x, m_x), \mathcal{M}\cup\{(x, m_x)\}) < \mathrm{d}(x, y)$$

and

$$D^2_{(y_1,m_{y_1}),(y_2,m_{y_2})}\xi_s((x, m_x), \mathcal{M}) = 0 \quad \text{if} \quad R_s((x, m_x), \mathcal{M}\cup\{(x, m_x)\}) < \max\{\mathrm{d}(x, y_1), \mathrm{d}(x, y_2)\}.$$

Proof. Note that $R := R_s$ and $\xi := \xi_s$. Moreover, we use the abbreviations $\hat{x} := (x, m_x)$, $\hat{y} := (y, m_y)$, $\hat{y}_1 := (y_1, m_{y_1})$ and $\hat{y}_2 := (y_2, m_{y_2})$. Recall that $\widehat{B}(z, r)$ stands for the cylinder $B(z, r) \times \mathbb{M}$ for $z \in \mathbb{X}$ and $r > 0$. It follows from the definitions of the difference operator and of the radius of stabilization that

$$\begin{aligned} D_{\hat{y}}\xi(\hat{x}, \mathcal{M} \cup \widehat{\mathcal{A}}) &= \xi(\hat{x}, \mathcal{M} \cup \widehat{\mathcal{A}} \cup \{\hat{y}\}) - \xi(\hat{x}, \mathcal{M} \cup \widehat{\mathcal{A}}) \\ &= \xi(\hat{x}, (\mathcal{M} \cup \widehat{\mathcal{A}} \cup \{\hat{y}\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))) \\ &\quad - \xi(\hat{x}, (\mathcal{M} \cup \widehat{\mathcal{A}}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))). \end{aligned} \tag{4.3}$$

If $R(\hat{x}, \mathcal{M}) < \mathrm{d}(x, y)$, we have

$$(\mathcal{M} \cup \widehat{\mathcal{A}} \cup \{\hat{y}\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\})) = (\mathcal{M} \cup \widehat{\mathcal{A}}\ ) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))$$

so that the terms on the right-hand side of (4.3) cancel out. For the second order difference operator, we obtain that

$$\begin{aligned} D^2_{\hat{y}_1,\hat{y}_2}\xi(\hat{x}, \mathcal{M}) &= \xi(\hat{x}, (\mathcal{M} \cup \{\hat{y}_1, \hat{y}_2\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))) \\ &\quad - \xi(\hat{x}, (\mathcal{M} \cup \{\hat{y}_1\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))) \\ &\quad - \xi(\hat{x}, (\mathcal{M} \cup \{\hat{y}_2\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))) \\ &\quad + \xi(\hat{x}, \mathcal{M} \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))). \end{aligned} \tag{4.4}$$

Without loss of generality we can assume that $\mathrm{d}(x, y_1) \geq \mathrm{d}(x, y_2)$. If $R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}) < \max\{\mathrm{d}(x, y_1), \mathrm{d}(x, y_2)\} = \mathrm{d}(x, y_1)$, we see that

$$(\mathcal{M} \cup \{y_1, y_2\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\})) = (\mathcal{M} \cup \{y_2\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\}))$$

and

$$(\mathcal{M} \cup \{\hat{y}_1\}) \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\})) = \mathcal{M} \cap \widehat{B}(x, R(\hat{x}, \mathcal{M}\cup\{\hat{x}\})),$$

whence the terms on the right-hand side of (4.4) cancel out. $\square$

We recall that M_x, $x \in \mathbb{X}$, always stands for a random mark distributed according to $\mathbb{Q}_{\mathbb{M}}$ and associated with the point x. Moreover, we tacitly assume that M_x is independent from everything else. For a finite set $\mathcal{A} \subset \mathbb{X}$, $(\mathcal{A}, M_{\mathcal{A}})$ is the shorthand notation for $\{(x, M_x) : x \in \mathcal{A}\}$. The next lemma shows that moments of difference operators of the scores decay exponentially fast. In the following $p \in (0, 1]$ and $C_p > 0$ come from the moment assumptions (2.6) and (2.7), respectively.

Lemma 4.4.

(a) For any $\varepsilon \in (0,p)$ *there are constants* $\hat{C}_\varepsilon, \hat{c}_\varepsilon \in (0,\infty)$ *such that*

$$\mathbb{E}\,|D_{(y,M_y)}\xi_s((x,M_x),\mathcal{P}_s \cup \{(x,M_x)\} \cup (\mathcal{A},M_{\mathcal{A}})\ )|^{4+\varepsilon} \le \hat{C}_\varepsilon \exp(-\hat{c}_\varepsilon\,\mathrm{d}_s(x,y)^\alpha)$$

for all $s \ge 1$*,* $x,y \in \mathbb{X}$ *and* $\mathcal{A} \subset \mathbb{X}$ *with* $|\mathcal{A}| \le 6$*.*

(b) For any $\varepsilon \in (0,p)$ *there are constants* $\hat{C}_\varepsilon, \hat{c}_\varepsilon \in (0,\infty)$ *such that*

$$\mathbb{E}\,|D_{(y,M_y)}\xi_n((x,M_x),\mathcal{X}_{n-8} \cup \{(x,M_x)\} \cup (\mathcal{A},M_{\mathcal{A}})\ )|^{4+\varepsilon} \le \hat{C}_\varepsilon \exp(-\hat{c}_\varepsilon\,\mathrm{d}_n(x,y)^\alpha)$$

for all $n \ge 9$*,* $x,y \in \mathbb{X}$ *and* $\mathcal{A} \subset \mathbb{X}$ *with* $|\mathcal{A}| \le 6$*.*

Proof. It follows from Hölder's inequality, (2.6), Lemma 4.3 and (2.4) in this order that

$$\begin{aligned}
&\mathbb{E}\,|D_{(y,M_y)}\xi_s((x,M_x),\mathcal{P}_s \cup (\mathcal{A},M_{\mathcal{A}} \cup \{(x,M_x)\}))|^{4+\varepsilon}\\
&\le \mathbb{E}\,\mathbf{1}_{\{D_{(y,M_y)}\xi_s((x,M_x),\mathcal{P}_s\cup(\mathcal{A},M_{\mathcal{A}})\cup\{(x,M_x)\})\neq 0\}}\\
&\qquad 2^{3+\varepsilon}(|\xi_s((x,M_x),\mathcal{P}_s \cup (\mathcal{A},M_{\mathcal{A}}) \cup \{(y,M_y)\} \cup \{(x,M_x)\})|^{4+\varepsilon}+\\
&\qquad |\xi_s((x,M_x),\mathcal{P}_s \cup (\mathcal{A},M_{\mathcal{A}} \cup \{(x,M_x)\}))|^{4+\varepsilon})\\
&\le 2^{4+\varepsilon}\ C_p^{\frac{4+\varepsilon}{4+p}}\ \mathbb{P}(D_{(y,M_y)}\xi_s((x,M_x),\mathcal{P}_s \cup \{(x,M_x)\} \cup (\mathcal{A},M_{\mathcal{A}})) \neq 0)^{\frac{p-\varepsilon}{4+p}}\\
&\le 2^{4+\varepsilon}\ C_p^{\frac{4+\varepsilon}{4+p}}\ \mathbb{P}(R_s((x,M_x),\mathcal{P}_s \cup \{(x,M_x)\}) \ge \mathrm{d}(x,y))^{\frac{p-\varepsilon}{4+p}}\\
&\le 2^{4+\varepsilon}\ C_p^{\frac{4+\varepsilon}{4+p}}\ C^{\frac{p-\varepsilon}{4+p}}\ \exp\bigg(-\frac{c(p-\varepsilon)}{4+p}\,\mathrm{d}_s(x,y)^\alpha\bigg),
\end{aligned}$$

which proves (a). Part (b) follows in the same way from (2.7), Lemma 4.3 and (2.5). □

Lemma 4.5. *For any* $\varepsilon \in (0,p)$*, there is a constant* $C_\varepsilon \in (0,\infty)$ *only depending on the constants in* (2.1)*,* (2.4)*, and* (2.6) *such that*

$$\mathbb{E}\,|D_{(y,M_y)}h_s(\mathcal{P}_s \cup \{(\mathcal{A},M_{\mathcal{A}})\})|^{4+\varepsilon} \le C_\varepsilon$$

for $y \in \mathbb{X}$*,* $\mathcal{A} \subset \mathbb{X}$ *with* $|\mathcal{A}| \le 1$ *and* $s \ge 1$*.*

Proof. Fix $y \in \mathbb{X}$. We start with the case $\mathcal{A} = \emptyset$. It follows from Lemma 4.2 and Jensen's inequality that

$$\begin{aligned}
&\mathbb{E}\,|D_{(y,M_y)}h_s(\mathcal{P}_s)|^{4+\varepsilon}\\
&= \mathbb{E}\,\bigg|\xi_s((y,M_y),\mathcal{P}_s \cup \{(y,M_y)\}) + \sum_{x\in\mathcal{P}_s} D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)\bigg|^{4+\varepsilon}\\
&\le 2^{3+\varepsilon}\mathbb{E}\,|\xi_s((y,M_y),\mathcal{P}_s \cup \{(y,M_y)\})|^{4+\varepsilon} + 2^{3+\varepsilon}\mathbb{E}\,\bigg|\sum_{x\in\mathcal{P}_s} D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)\bigg|^{4+\varepsilon}.
\end{aligned}$$

Here, the first summand is bounded by $2^{3+\varepsilon}(C_p+1)$ by assumption (2.6). The second summand is a sum of $Z := \sum_{x\in\mathcal{P}_s} \mathbf{1}\{D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)\neq 0\}$ terms distinct from zero. A further application of Jensen's inequality to the function $x\mapsto x^{4+\varepsilon}$ leads to

$$\begin{aligned}\Big|\sum_{x\in\mathcal{P}_s} D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)\Big|^{4+\varepsilon} &\le Z^{4+\varepsilon}\Big|\sum_{x\in\mathcal{P}_s} Z^{-1}D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)\Big|^{4+\varepsilon}\\ &\le Z^{4+\varepsilon}\sum_{x\in\mathcal{P}_s} Z^{-1}|D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)|^{4+\varepsilon}\\ &\le Z^{4}\sum_{x\in\mathcal{P}_s} |D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)|^{4+\varepsilon}.\end{aligned}$$

By deciding whether points in different sums are identical or distinct, we obtain that

$$\mathbb{E}\, Z^4\sum_{x\in\mathcal{P}_s}|D_{(y,M_y)}\xi_s(x,\mathcal{P}_s)|^{4+\varepsilon} = I_1+15I_2+25I_3+10I_4+I_5,$$

where, for $i\in\{1,\dots,5\}$,

$$I_i=\mathbb{E}\sum_{(x_1,\dots,x_i)\in\mathcal{P}^i_{s,\neq}} \mathbf{1}\{D_{(y,M_y)}\xi_s(x_j,\mathcal{P}_s)\neq 0, j=1,\dots,i\}\,|D_{(y,M_y)}\xi_s(x_1,\mathcal{P}_s)|^{4+\varepsilon}.$$

By $\mathcal{P}^i_{s,\neq}$ we denote the set of i-tuples of distinct points of $\mathcal{P}_s$. It follows from the multivariate Mecke formula and Hölder's inequality that

$$\begin{aligned}I_i &= s^i\int_{\widehat{\mathbb{X}}^i}\mathbb{E}\,\mathbf{1}\{D_{(y,M_y)}\xi_s(x_j,\mathcal{P}_s\cup\{x_1,\dots,x_i\})\neq 0, j=1,\dots,i\}\\ &\qquad\qquad |D_{(y,M_y)}\xi_s(x_1,\mathcal{P}_s\cup\{x_1,\dots,x_i\})|^{4+\varepsilon}\,\widehat{\mathbb{Q}}^i(\mathrm{d}(x_1,\dots,x_i))\\ &\le s^i\int_{\mathbb{X}^i}\prod_{j=1}^{i}\Big[\mathbb{P}(D_{(y,M_y)}\xi_s(x_j,\mathcal{P}_s\cup\{(x_1,M_{x_1}),\dots,(x_i,M_{x_i})\})\neq 0)^{\frac{p-\varepsilon}{4i+pi}}\Big]\\ &\qquad (\mathbb{E}\,|D_{(y,M_y)}\xi_s(x_1,\mathcal{P}_s\cup\{(x_1,M_{x_1}),\dots,(x_i,M_{x_i})\})|^{4+p})^{\frac{4+\varepsilon}{4+p}}\,\mathbb{Q}^i(\mathrm{d}(x_1,\dots,x_i)).\end{aligned}$$

Combining this with (2.6), Lemma 4.3 and (2.4) leads to

$$\begin{aligned}I_i &\le 2^{4+\varepsilon}C_p^{\frac{4+\varepsilon}{4+p}}s^i\int_{\mathbb{X}^i}C^{\frac{p-\varepsilon}{4+p}}\prod_{j=1}^{i}\exp\Big(-\frac{c(p-\varepsilon)}{4i+pi}\,\mathrm{d}_s(x_j,y)^\alpha\Big)\,\mathbb{Q}^i(\mathrm{d}(x_1,\dots,x_i))\\ &= 2^{4+\varepsilon}C_p^{\frac{4+\varepsilon}{4+p}}\Big(sC^{\frac{p-\varepsilon}{4i+pi}}\int_{\mathbb{X}}\exp\Big(-\frac{c(p-\varepsilon)}{4i+pi}\,\mathrm{d}_s(x,y)^\alpha\Big)\,\mathbb{Q}(\mathrm{d}x)\Big)^i.\end{aligned}$$

Now (4.2) with $r=0$ yields that the integrals on the right-hand side are uniformly bounded and thus the first asserted moment bound holds.

Next we assume that $\mathcal{A} = \{z\}$ with $z \in \mathbb{X}$. Lemma 4.2 and a further application of Jensen's inequality show that

$$\begin{aligned}
&\mathbb{E}\,|D_{(y,M_y)}h_s(\mathcal{P}_s \cup \{(z,M_z)\})|^{4+\varepsilon} \\
&= \mathbb{E}\,|\xi_s((y,M_y),\mathcal{P}_s \cup \{(y,M_y),(z,M_z)\}) + D_{(y,M_y)}\xi_s((z,M_z),\mathcal{P}_s \cup \{(z,M_z)\}) \\
&\qquad + \sum_{x\in\mathcal{P}_s} D_{(y,M_y)}\xi_s(x,\mathcal{P}_s \cup \{(z,M_z)\})|^{4+\varepsilon} \\
&\le 3^{3+\varepsilon}\mathbb{E}\,|\xi_s((y,M_y),\mathcal{P}_s \cup \{(y,M_y),(z,M_z)\})|^{4+\varepsilon} \\
&\quad + 3^{3+\varepsilon}\mathbb{E}\,|D_{(y,M_y)}\xi_s((z,M_z),\mathcal{P}_s \cup \{(z,M_z)\})|^{4+\varepsilon} \\
&\quad + 3^{3+\varepsilon}\mathbb{E}\left|\sum_{x\in\mathcal{P}_s} D_{(y,M_y)}\xi_s(x,\mathcal{P}_s \cup \{(z,M_z)\})\right|^{4+\varepsilon}.
\end{aligned}$$

The last term on the right-hand side can be now bounded by exactly the same arguments as above since these still hold true if one adds an additional point. As the other terms are bounded by (2.6) and Lemma 4.4, this completes the proof. □

Lemma 4.6. *For any $\varepsilon \in (0,p)$, there is a constant $C_\varepsilon \in (0,\infty)$ only depending on the constants in* (2.1), (2.5) *and* (2.7) *such that*

$$\mathbb{E}\,|D_{(y,M_y)}h_n(\mathcal{X}_{n-1-|\mathcal{A}|} \cup (\mathcal{A},M_{\mathcal{A}}))|^{4+\varepsilon} \le C_\varepsilon$$

for $y \in \mathbb{X}$, $\mathcal{A} \subset \mathbb{X}$ with $|\mathcal{A}| \le 2$ and $n \ge 9$.

Proof. Let $\mathcal{X}_{n,\mathcal{A}} = \mathcal{X}_{n-1-|\mathcal{A}|} \cup (\mathcal{A},M_{\mathcal{A}})$. It follows from Lemma 4.2 and Jensen's inequality that

$$\begin{aligned}
&\mathbb{E}\,|D_{(y,M_y)}h_n(\mathcal{X}_{n,\mathcal{A}})|^{4+\varepsilon} \\
&= \mathbb{E}\left|\xi_n((y,M_y),\mathcal{X}_{n,\mathcal{A}} \cup \{(y,M_y)\}) + \sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}\cup(\mathcal{A},M_{\mathcal{A}})} D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}})\right|^{4+\varepsilon} \\
&\le 4^{3+\varepsilon}\mathbb{E}\,|\xi_n((y,M_y),\mathcal{X}_{n,\mathcal{A}} \cup \{(y,M_y)\})|^{4+\varepsilon} + 4^{3+\varepsilon}\sum_{x\in\mathcal{A}}\mathbb{E}\,|D_{(y,M_y)}\xi_n((x,M_x),\mathcal{X}_{n,\mathcal{A}})|^{4+\varepsilon} \\
&\quad + 4^{3+\varepsilon}\mathbb{E}\left|\sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}} D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}})\right|^{4+\varepsilon}.
\end{aligned}$$

On the right-hand side, the first summand is bounded by $4^{3+\varepsilon}(C_p+1)$ by assumption (2.7) (after conditioning on the points of $\mathcal{X}_{n-1-|\mathcal{A}|} \setminus \mathcal{X}_{n-8}$) and the second summand is bounded by $2 \cdot 4^{3+\varepsilon}\hat{C}_\varepsilon$ by Lemma 4.4. A further application of Jensen's inequality with

$Z := \sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}} \mathbf{1}\{D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}}) \neq 0\}$ leads to

$$\Big| \sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}} D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}}) \Big|^{4+\varepsilon} \leq Z^{3+\varepsilon} \sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}} |D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}})|^{4+\varepsilon}$$
$$\leq Z^4 \sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}} |D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}})|^{4+\varepsilon}.$$

By deciding whether points in different sums are identical or distinct, we obtain that

$$\mathbb{E}\, Z^4 \sum_{x\in\mathcal{X}_{n-1-|\mathcal{A}|}} |D_{(y,M_y)}\xi_n(x,\mathcal{X}_{n,\mathcal{A}})|^{4+\varepsilon} = I_1 + 15 I_2 + 25 I_3 + 10 I_4 + I_5,$$

where, for $i \in \{1,\dots,5\}$,

$$I_i = \mathbb{E} \sum_{(x_1,\dots,x_i)\in\mathcal{X}^i_{n-1-|\mathcal{A}|,\neq}} \mathbf{1}\{D_{(y,M_y)}\xi_n(x_j,\mathcal{X}_{n,\mathcal{A}}) \neq 0, j=1,\dots,i\}\, |D_{(y,M_y)}\xi_n(x_1,\mathcal{X}_{n,\mathcal{A}})|^{4+\varepsilon}.$$

It follows from Hölder's inequality that

$$\begin{aligned} I_i &= \frac{(n-1-|\mathcal{A}|)!}{(n-1-|\mathcal{A}|-i)!} \int_{\widehat{\mathbb{X}}^i} \mathbb{E}\,\mathbf{1}\{D_{(y,M_y)}\xi_n(x_j,\mathcal{X}_{n-i,\mathcal{A}}\cup\{x_1,\dots,x_i\}) \neq 0, j=1,\dots,i\} \\ &\qquad |D_{(y,M_y)}\xi_n(x_1,\mathcal{X}_{n-i,\mathcal{A}}\cup\{x_1,\dots,x_i\})|^{4+\varepsilon}\, \widehat{\mathbb{Q}}^i(\mathrm{d}(x_1,\dots,x_i)) \\ &\leq n^i \int_{\mathbb{X}^i} \prod_{j=1}^{i} \mathbb{P}(D_{(y,M_y)}\xi_n((x_j,M_{x_j}),\mathcal{X}_{n-i,\mathcal{A}}\cup\{(x_1,M_{x_1}),\dots,(x_i,M_{x_i})\}) \neq 0)^{\frac{p-\varepsilon}{4i+pi}} \\ &\qquad (\mathbb{E}\,|D_{(y,M_y)}\xi_n((x_1,M_{x_1}),\mathcal{X}_{n-i,\mathcal{A}}\cup\{(x_1,M_{x_1}),\dots,(x_i,M_{x_i})\})|^{4+p})^{\frac{4+\varepsilon}{4+p}} \\ &\qquad \mathbb{Q}^i(\mathrm{d}(x_1,\dots,x_i)). \end{aligned}$$

Combining this with (2.7), Lemma 4.3 and (2.5) leads to

$$\begin{aligned} I_i &\leq 2^{4+\varepsilon} C_p^{\frac{4+\varepsilon}{4+p}} n^i \int_{\mathbb{X}^i} C^{\frac{p-\varepsilon}{4+p}} \prod_{j=1}^{i} \exp\Big(-\frac{c(p-\varepsilon)}{4i+pi}\,\mathrm{d}_n(x_j,y)^\alpha\Big)\, \mathbb{Q}^i(\mathrm{d}(x_1,\dots,x_i)) \\ &= 2^{4+\varepsilon} C_p^{\frac{4+\varepsilon}{4+p}} \Big(n C^{\frac{p-\varepsilon}{4i+pi}} \int_{\mathbb{X}} \exp\Big(-\frac{c(p-\varepsilon)}{4i+pi}\,\mathrm{d}_n(x,y)^\alpha\Big)\, \mathbb{Q}(\mathrm{d}x)\Big)^i. \end{aligned}$$

Now (4.2) yields that the integrals on the right-hand side are uniformly bounded. □

Lemma 4.7.

(a) For $x,z\in\mathbb{X}$ and $s\geq 1$,

$$\mathbb{P}(D_{(x,M_x)}\xi_s((z,M_z),\mathcal{P}_s\cup\{z,M_z\}) \neq 0) \leq 2C\exp(-c\max\{\mathrm{d}_s(x,z),\mathrm{d}_s(z,K)\}^\alpha).$$

(b) For $x_1, x_2, z \in \mathbb{X}$ and $s \geq 1$,

$$\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_s((z,M_z),\mathcal{P}_s \cup \{(z,M_z)\}) \neq 0)$$
$$\leq 4C\exp(-c\max\{\mathrm{d}_s(x_1,z),\mathrm{d}_s(x_2,z),\mathrm{d}_s(z,K)\}^\alpha).$$

Proof. We prove part (b). By Lemma 4.3 the event

$$D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_s((z,M_z),\mathcal{P}_s \cup \{(z,M_z)\}) \neq 0$$

is a subset of the event that the points x_1, x_2 belong to the ball centered at z with radius $R_s((z,M_z),\mathcal{P}_s\cup\{(z,M_z)\})$, i.e., $R_s((z,M_z),\mathcal{P}_s\cup\{(z,M_z)\}) \geq \max\{\mathrm{d}(x_1,z),\mathrm{d}(x_2,z)\}$. By (2.4) this gives

$$\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_s((z,M_z),\mathcal{P}_s\cup\{(z,M_z)\}) \neq 0) \leq C\exp(-c\max\{\mathrm{d}_s(x_1,z),\mathrm{d}_s(x_2,z)\}^\alpha).$$

By (2.8) we also have

$$\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_s((z,M_z),\mathcal{P}_s \cup \{(z,M_z)\}) \neq 0) \leq 4C\exp(-c\,\mathrm{d}_s(z,K)^\alpha).$$

This gives the proof of part (b). Part (a) is proven in a similar way. □

Lemma 4.8.

(a) For $x, z \in \mathbb{X}$ and $n \geq 9$,

$$\sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \mathbb{P}(D_{(x,M_x)}\xi_n((z,M_z),\mathcal{X}_{n-2-|\mathcal{A}|} \cup \{(z,M_z)\} \cup (\mathcal{A},M_\mathcal{A})) \neq 0)$$
$$\leq 2C\exp(-c\max\{\mathrm{d}_n(x,z),\mathrm{d}_n(z,K)\}^\alpha).$$

(b) For $x_1, x_2, z \in \mathbb{X}$ and $n \geq 9$,

$$\sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 1} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_n((z,M_z),\mathcal{X}_{n-3-|\mathcal{A}|} \cup \{(z,M_z)\} \cup (\mathcal{A},M_\mathcal{A})) \neq 0)$$
$$\leq 4C\exp(-c\max\{\mathrm{d}_n(x_1,z),\mathrm{d}_n(x_2,z),\mathrm{d}_n(z,K)\}^\alpha).$$

Proof. By Lemma 4.3 and (2.5) together with similar arguments as in the proof of Lemma 4.7, we obtain

$$\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_n((z,M_z),\mathcal{X}_{n-3-|\mathcal{A}|} \cup \{(z,M_z)\} \cup (\mathcal{A},M_\mathcal{A})) \neq 0)$$
$$\leq \mathbb{P}(R_n((z,M_z),\mathcal{X}_{n-8} \cup \{(z,M_z)\}) > \max\{\mathrm{d}_s(x_1,z),\mathrm{d}_s(x_2,z)\})$$
$$\leq C\exp(-c\max\{\mathrm{d}_s(x_1,z),\mathrm{d}_s(x_2,z)\}^\alpha).$$

It follows from (2.9) that

$$\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_n((z,M_z),\mathcal{X}_{n-3-|\mathcal{A}|} \cup \{(z,M_z)\} \cup (\mathcal{A},M_\mathcal{A})) \neq 0)$$
$$\leq 4C\exp(-c\,\mathrm{d}_s(z,K)^\alpha),$$

which completes the proof of part (b). Part (a) is proven similarly. □

Lemma 4.9.

(a) Let $\beta \in (0,\infty)$ be fixed. Then there is a constant $C_\beta \in (0,\infty)$ such that

$$s \int_{\mathbb{X}} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_s(\mathcal{P}_s) \neq 0)^\beta \, \mathbb{Q}(\mathrm{d}x_2) \leq C_\beta \exp(-c\beta \, \mathrm{d}_s(x_1,K)/4^{\alpha+1})$$

for all $x_1 \in \mathbb{X}$ and $s \in [1,\infty)$.

(b) Let $\beta \in (0,\infty)$ be fixed. Then there is a constant $C_\beta \in (0,\infty)$ such that

$$\begin{aligned} &n \int_{\mathbb{X}} \sup_{\mathcal{A} \subset \mathbb{X}, |\mathcal{A}| \leq 1} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_n(\mathcal{X}_{n-2-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0)^\beta \, \mathbb{Q}(\mathrm{d}x_2) \\ &\leq C_\beta \exp(-c\beta \, \mathrm{d}_s(x_1,K)/4^{\alpha+1}) \end{aligned}$$

for all $x_1 \in \mathbb{X}$ and $n \geq 9$.

Proof. By Lemma 4.2 we have

$$\begin{aligned} D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_s(\mathcal{P}_s) &= D_{(x_1,M_{x_1})} \xi_s((x_2,M_{x_2}), \mathcal{P}_s \cup \{(x_2,M_{x_2})\}) \\ &\quad + D_{(x_2,M_{x_2})} \xi_s((x_1,M_{x_1}), \mathcal{P}_s \cup \{(x_1,M_{x_1}\}) \\ &\quad + \sum_{z \in \mathcal{P}_s} D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} \xi_s(z, \mathcal{P}_s) \end{aligned}$$

so that the Slivnyak-Mecke formula leads to

$$\begin{aligned} &\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_s(\mathcal{P}_s) \neq 0) \\ &\leq \mathbb{P}(D_{(x_1,M_{x_1})} \xi_s((x_2,M_{x_2}), \mathcal{P}_s \cup \{(x_2,M_{x_2})\}) \neq 0) \\ &\quad + \mathbb{P}(D_{(x_2,M_{x_2})} \xi_s((x_1,M_{x_1}), \mathcal{P}_s \cup \{(x_1,M_{x_1})\}) \neq 0) \\ &\quad + \underbrace{s \int_{\mathbb{X}} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} \xi_s((z,M_z), \mathcal{P}_s \cup \{(z,M_z)\}) \neq 0) \, \mathbb{Q}(\mathrm{d}z)}_{=:T_{x_1,x_2,s}} . \end{aligned}$$

Here, we use part (a) of Lemma 4.7 to bound each of the first two terms on the right-hand side. We may bound the first term by

$$\mathbb{P}(D_{(x_1,M_{x_1})} \xi_s((x_2,M_{x_2}), \mathcal{P}_s \cup \{(x_2,M_{x_2})\}) \neq 0) \leq 2C \exp\left(-c \max\{\mathrm{d}_s(x_1,x_2), \mathrm{d}_s(x_2,K)\}^\alpha\right).$$

Observing that $\mathrm{d}_s(x_1,K) \leq 2\max\{\mathrm{d}_s(x_2,K), \mathrm{d}_s(x_1,x_2)\}$ we obtain

$$\begin{aligned} &\mathbb{P}(D_{(x_1,M_{x_1})} \xi_s((x_2,M_{x_2}), \mathcal{P}_s \cup \{(x_2,M_{x_2})\}) \neq 0) \\ &\leq 2C \exp\left(-c \max\{\mathrm{d}_s(x_1,x_2), \mathrm{d}_s(x_1,K), \mathrm{d}_s(x_2,K)\}^\alpha / 2^\alpha\right). \end{aligned}$$

We bound $\mathbb{P}(D_{(x_2,M_{x_2})}\xi_s((x_1,M_{x_1}),\mathcal{P}_s\cup\{(x_1,M_{x_1})\})\neq 0)$ in the same way. It follows from part (b) of Lemma 4.7 that

$$T_{x_1,x_2,s}\leq 4Cs\int_{\mathbb{X}}\exp(-c\max\{\mathrm{d}_s(x_1,z),\mathrm{d}_s(x_2,z),\mathrm{d}_s(z,K)\}^\alpha)\,\mathbb{Q}(\mathrm{d}z).$$

Assume that $\mathrm{d}_s(x_1,K)\geq \mathrm{d}_s(x_2,K)$ (the reasoning is similar if $\mathrm{d}_s(x_2,K)\geq\mathrm{d}_s(x_1,K)$) and let $r=\max\{\mathrm{d}(x_1,K),\mathrm{d}(x_1,x_2)\}/2$. For any $z\in B(x_1,r)$ the triangle inequality leads to $\max\{\mathrm{d}(z,x_2),\mathrm{d}(z,K)\}\geq r$. This implies that

$$\begin{aligned}T_{x_1,x_2,s}\leq 4Cs\int_{B(x_1,r)}&\exp(-c\underbrace{\max\{\mathrm{d}_s(z,x_2),\mathrm{d}_s(z,K)\}^\alpha}_{\geq(s^{1/\gamma}r)^\alpha})\,\mathbb{Q}(\mathrm{d}z)\\&+4Cs\int_{\mathbb{X}\setminus B(x_1,r)}\exp(-c\,\mathrm{d}_s(x_1,z)^\alpha)\,\mathbb{Q}(\mathrm{d}z).\end{aligned}$$

Recalling from (4.1) that $\mathbb{Q}(B(x_1,r))\leq\kappa r^\gamma$, we have

$$4Cs\int_{B(x_1,r)}\exp(-c(s^{1/\gamma}r)^\alpha)\,\mathbb{Q}(\mathrm{d}z)\leq 4Cs\kappa r^\gamma\exp(-c(s^{1/\gamma}r)^\alpha)\leq C_1\exp(-c(s^{1/\gamma}r)^\alpha/2)$$

with a constant $C_1\in(0,\infty)$ only depending on C, c, γ and α. On the other hand, (4.2) yields

$$4Cs\int_{\mathbb{X}\setminus B(x_1,r)}\exp(-c\,\mathrm{d}_s(x_1,z)^\alpha)\,\mathbb{Q}(\mathrm{d}z)\leq C_2\exp(-c(s^{1/\gamma}r)^\alpha/2)$$

with a constant $C_2\in(0,\infty)$ only depending on C, c, γ and α. Hence, we obtain

$$T_{x_1,x_2,s}\leq(C_1+C_2)\exp(-c\max\{\mathrm{d}_s(x_1,K),\mathrm{d}_s(x_2,K),\mathrm{d}_s(x_1,x_2)\}^\alpha/2^{\alpha+1})$$

and

$$\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}h_s(\mathcal{P}_s)\neq 0)\leq C_3\exp(-c\max\{\mathrm{d}_s(x_1,K),\mathrm{d}_s(x_2,K),\mathrm{d}_s(x_1,x_2)\}^\alpha/2^{\alpha+1})$$

with $C_3:=C_1+C_2+4C$. Consequently, we have

$$\begin{aligned}&s\int_{\mathbb{X}}\mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}h_s(\mathcal{P}_s)\neq 0)^\beta\,\mathbb{Q}(\mathrm{d}x_2)\\&\leq C_3^\beta s\int_{B(x_1,\mathrm{d}(x_1,K)/2)}\exp(-c\beta\,\mathrm{d}_s(x_2,K)^\alpha/2^{\alpha+1})\,\mathbb{Q}(\mathrm{d}x_2)\\&\quad+C_3^\beta s\int_{\mathbb{X}\setminus B(x_1,\mathrm{d}(x_1,K)/2)}\exp(-c\beta\,\mathrm{d}_s(x_2,x_1)^\alpha/2^{\alpha+1})\,\mathbb{Q}(\mathrm{d}x_2).\end{aligned}$$

Using the same arguments as above, the right-hand side can be bounded by

$$C_\beta\exp(-c\beta\,\mathrm{d}_s(K,x_1)^\alpha/4^{\alpha+1})$$

with a constant $C_\beta \in (0,\infty)$ only depending on β, C_3, c, γ and α. This completes the proof of (a).

Similar arguments show for the binomial case that

$$\begin{aligned}
&\sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\le 1} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_n(\mathcal{X}_{n-2-|\mathcal{A}|}\cup(\mathcal{A},M_{\mathcal{A}}))\neq 0)\\
&\le \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\le 1} \mathbb{P}(D_{(x_1,M_{x_1})}\xi_n((x_2,M_{x_2}),\mathcal{X}_{n-2-|\mathcal{A}|}\cup\{(x_2,M_{x_2})\}\cup(\mathcal{A},M_{\mathcal{A}}))\neq 0)\\
&\quad + \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\le 1} \mathbb{P}(D_{(x_2,M_{x_2})}\xi_n((x_1,M_{x_1}),\mathcal{X}_{n-2-|\mathcal{A}|}\cup\{(x_1,M_{x_1})\}\cup(\mathcal{A},M_{\mathcal{A}}))\neq 0)\\
&\quad + n\int_{\mathbb{X}} \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\le 1} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}\xi_n((z,M_z),\mathcal{X}_{n-3-|\mathcal{A}|}\cup\{(z,M_z)\}\cup(\mathcal{A},M_{\mathcal{A}}))\neq 0)\,\mathbb{Q}(\mathrm{d}z).
\end{aligned}$$

Now similar computations as for the Poisson case conclude the proof of part (b). □

For $\alpha,\tau\ge 0$ put

$$I_{K,s}(\alpha,\tau) := s\int_{\mathbb{X}} \exp(-\tau\,\mathrm{d}_s(x,K)^\alpha)\,\mathbb{Q}(\mathrm{d}x),\quad s\ge 1.$$

Lemma 4.10. *Let $\beta\in(0,\infty)$ be fixed. There is a constant $\tilde{C}_\beta\in(0,\infty)$ such that for all $s\ge 1$ we have*

$$s\int_{\mathbb{X}}\left(s\int_{\mathbb{X}} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}h_s(\mathcal{P}_s)\neq 0)^\beta\,\mathbb{Q}(\mathrm{d}x_2)\right)^2\mathbb{Q}(\mathrm{d}x_1)\le \tilde{C}_\beta I_{K,s}(\alpha,c\beta/2^{2\alpha+1}), \tag{4.5}$$

$$s^2\int_{\mathbb{X}^2} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})}h_s(\mathcal{P}_s)\neq 0)^\beta\,\mathbb{Q}^2(\mathrm{d}(x_1,x_2))\le \tilde{C}_\beta I_{K,s}(\alpha,c\beta/4^{\alpha+1}) \tag{4.6}$$

and

$$s\int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)}h_s(\mathcal{P}_s)\neq 0)^\beta\,\mathbb{Q}(\mathrm{d}x)\le \tilde{C}_\beta I_{K,s}(\alpha,c\beta/2^{\alpha+1}). \tag{4.7}$$

Proof. By Lemma 4.9 a) the integrals in (4.5) and (4.6) are bounded by

$$C_\beta^2 s\int_{\mathbb{X}}\exp(-c\beta\,\mathrm{d}_s(x_1,K)^\alpha/2^{2\alpha+1})\,\mathbb{Q}(\mathrm{d}x_1) = C_\beta^2 I_{k,s}(\alpha,c\beta/2^{2\alpha+1})$$

and

$$C_\beta s\int_{\mathbb{X}}\exp(-c\beta\,\mathrm{d}_s(x_1,K)^\alpha/4^{\alpha+1})\,\mathbb{Q}(\mathrm{d}x_1) = C_\beta I_{k,s}(\alpha,c\beta/4^{\alpha+1}),$$

respectively. In order to derive the bound in (4.7), we compute $\mathbb{P}(D_{(x,M_x)}h_s(\mathcal{P}_s)\neq 0)$ as follows. By Lemma 4.2 and the Slivnyack-Mecke formula we obtain that

$$\begin{aligned}
&\mathbb{P}(D_{(x,M_x)}h_s(\mathcal{P}_s)\neq 0)\\
&\le \mathbb{P}(\xi_s((x,M_x),\mathcal{P}_s\cup\{(x,M_x)\})\neq 0) + \mathbb{E}\sum_{z\in\mathcal{P}_s}\mathbf{1}\{D_{(x,M_x)}\xi_s(z,\mathcal{P}_s)\neq 0\}
\end{aligned}$$

$$= \mathbb{P}(\xi_s((x, M_x), \mathcal{P}_s \cup \{(x, M_x)\}) \neq 0)$$
$$+ s \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} \xi_s((z, M_z), \mathcal{P}_s \cup \{(z, M_z)\}) \neq 0) \, \mathbb{Q}(\mathrm{d}z).$$

By (2.8) and Lemma 4.7 a) we obtain that for all $x \in \mathbb{X}$ and $s \geq 1$,

$$\mathbb{P}(D_{(x,M_x)} h_s(\mathcal{P}_s) \neq 0) \leq C \exp(-c \, \mathrm{d}_s(x, K)^\alpha) + 2Cs \int_{\mathbb{X}} \exp(-c \max\{\mathrm{d}_s(x, z), \mathrm{d}_s(z, K)\}^\alpha) \, \mathbb{Q}(\mathrm{d}z).$$

Letting $r := \mathrm{d}(x, K)/2$, partitioning $\mathbb{X}$ into the union of $\mathbb{X} \setminus B(x, r)$ and $B(x, r)$, and following the discussion in the proof of Lemma 4.9, we obtain

$$2Cs \int_{\mathbb{X}} \exp(-c \max\{\mathrm{d}_s(x, z), \mathrm{d}_s(z, K)\}^\alpha) \, \mathbb{Q}(\mathrm{d}z) \leq C_1 \exp(-c \, \mathrm{d}_s(x, K)^\alpha / 2^{\alpha+1})$$

with a constant $C_1 \in (0, \infty)$. Consequently, for all $x \in \mathbb{X}$ and $s \geq 1$ we have

$$\mathbb{P}(D_{(x,M_x)} h_s(\mathcal{P}_s) \neq 0) \leq (C + C_1) \exp(-c \, \mathrm{d}_s(x, K)^\alpha / 2^{\alpha+1})$$

and

$$s \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)} h_s(\mathcal{P}_s) \neq 0)^\beta \, \mathbb{Q}(\mathrm{d}x) \leq (C + C_1) I_{K,s}(\alpha, c\beta/2^{\alpha+1}),$$

which was to be shown. □

Lemma 4.11. *Let $\beta \in (0, \infty)$ be fixed. There is a constant $C_\beta \in (0, \infty)$ such that for all $n \geq 9$ we have*

$$n \int_{\mathbb{X}} \left(n \int_{\mathbb{X}} \sup_{\mathcal{A} \subset \mathbb{X}, |\mathcal{A}| \leq 1} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_n(\mathcal{X}_{n-2-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0)^\beta \, \mathbb{Q}(\mathrm{d}x_2) \right)^2 \mathbb{Q}(\mathrm{d}x_1) \leq C_\beta I_{K,n}(\alpha, c\beta/2^{2\alpha+1}), \tag{4.8}$$

$$n^2 \int_{\mathbb{X}^2} \sup_{\mathcal{A} \subset \mathbb{X}, |\mathcal{A}| \leq 1} \mathbb{P}(D^2_{(x_1,M_{x_1}),(x_2,M_{x_2})} h_n(\mathcal{X}_{n-2-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}}))) \neq 0)^\beta \, \mathbb{Q}^2(\mathrm{d}(x_1, x_2) \leq C_\beta I_{K,n}(\alpha, c\beta/4^{\alpha+1}) \tag{4.9}$$

and

$$n \int_{\mathbb{X}} \sup_{\mathcal{A} \subset \mathbb{X}, |\mathcal{A}| \leq 2} \mathbb{P}(D_{(x,M_x)} h_n(\mathcal{X}_{n-1-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0)^\beta \, \mathbb{Q}(\mathrm{d}x) \leq C_\beta I_{K,n}(\alpha, c\beta/2^{\alpha+1}). \tag{4.10}$$

Proof. The bounds in (4.8) and (4.9) follow immediately from Lemma 4.9 b) and the definition of $I_{K,n}(\alpha,\tau)$. By Lemma 4.2 we obtain that, for $x \in \mathbb{X}$,

$$
\begin{aligned}
&\sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \mathbb{P}(D_{(x,M_x)} h_n(\mathcal{X}_{n-1-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \\
&\leq \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \mathbb{P}(\xi_n((x,M_x), \mathcal{X}_{n-1-|\mathcal{A}|} \cup \{(x,M_x)\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \\
&\quad + \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \mathbb{E} \sum_{z\in\mathcal{X}_{n-1-|\mathcal{A}|}\cup(\mathcal{A},M_{\mathcal{A}})} \mathbf{1}(D_{(x,M_x)}\xi_n(z, \mathcal{X}_{n-1-|\mathcal{A}|} \cup \{z\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \\
&\leq \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \mathbb{P}(\xi_n((x,M_x), \mathcal{X}_{n-1-|\mathcal{A}|} \cup \{(x,M_x)\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \\
&\quad + \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \sum_{z\in\mathcal{A}} \mathbb{P}(D_{(x,M_x)}\xi_n((z,M_z), \mathcal{X}_{n-1-|\mathcal{A}|} \cup \{(z,M_z)\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \\
&\quad + \sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} n \int_{\mathbb{X}} \mathbb{P}(D_{(x,M_x)}\xi_n((z,M_z), \mathcal{X}_{n-2-|\mathcal{A}|} \cup \{(z,M_z)\} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0)\, \mathbb{Q}(\mathrm{d}z).
\end{aligned}
$$

Combining the bound from Lemma 4.8 a) with the computations from the proof of Lemma 4.10 and (2.9), we see that there is a constant $C_1 \in (0,\infty)$ such that for all $x \in \mathbb{X}$ and $s \geq 1$ we have

$$
\sup_{\mathcal{A}\subset\mathbb{X},|\mathcal{A}|\leq 2} \mathbb{P}(D_{(x,M_x)} h_n(\mathcal{X}_{n-1-|\mathcal{A}|} \cup (\mathcal{A}, M_{\mathcal{A}})) \neq 0) \leq C_1 \exp(-c\, \mathrm{d}_s(x,K)^\alpha / 2^{\alpha+1}).
$$

Now (4.10) follows from the definition of $I_{K,n}(\alpha,\tau)$. □

Proof of Theorem 2.1. We start with the proof of the Poisson case (2.11). It follows from Lemma 4.5 that the condition (3.1) with the exponent $4 + p/2$ in Theorem 3.1 is satisfied for all $s \geq 1$ with the constant $C_{p/2}$. In the following we use the abbreviation

$$
I_{K,s} = I_{K,s}(\alpha, cp/(36 \cdot 4^{\alpha+1})).
$$

Together with Lemma 4.10 (with $\beta = p/36$) it follows from Theorem 3.1 that there is a constant $\tilde{C} \in (0,\infty)$ depending on $\tilde{C}_{p/36}$,$C_{p/2}$ and p such that

$$
d_K\left(\frac{H_s - \mathbb{E}\, H_s}{\sqrt{\operatorname{Var} H_s}}, N\right) \leq \tilde{C}\left(\frac{\sqrt{I_{K,s}}}{\operatorname{Var} H_s} + \frac{I_{K,s}}{(\operatorname{Var} H_s)^{3/2}} + \frac{I_{K,s}^{5/4} + I_{K,s}^{3/2}}{(\operatorname{Var} H_s)^2}\right),
$$

which completes the proof of the Poisson case.

For the binomial case (2.12) it follows from Lemma 4.6 that the condition (3.3) in Theorem 3.2 is satisfied with the exponent $4 + p/2$ for all $n \geq 9$ with the same constant $C_{p/2} \geq 1$. Using the abbreviation

$$
I_{K,n} = I_{K,n}(\alpha, cp/(18 \cdot 4^{\alpha+1})),
$$

we obtain from Lemma 4.10 (with $\beta = p/18$) and Theorem 3.2 that there is a constant $\tilde{C} \in (0, \infty)$ depending on $\tilde{C}_{p/18}$, $C_{p/2}$ and p such that

$$d_K\left(\frac{H_n' - \mathbb{E}\, H_n'}{\sqrt{\operatorname{Var} H_n'}}, N\right) \le \tilde{C}\left(\frac{\sqrt{I_{K,n}}}{\operatorname{Var} H_n'} + \frac{I_{K,n}}{(\operatorname{Var} H_n')^{3/2}} + \frac{I_{K,n} + (I_{K,n})^{3/2}}{(\operatorname{Var} H_n)^2}\right),$$

which completes the proof. □

Before giving the proof of Theorem 2.3 we require a lemma. We thank Steffen Winter for providing the essential argument. For $K \subset \mathbb{R}^d$, recall that $K_r := \{y \in \mathbb{R}^d : \ \mathrm{d}(y, K) \le r\}$ and that $\overline{\mathcal{M}}^{d-1}(K)$ is defined at (2.16).

Lemma 4.12. *If K is either a full-dimensional subset of $\mathbb{R}^d$ with $\overline{\mathcal{M}}^{d-1}(\partial K) < \infty$ or a $(d-1)$-dimensional compact subset of $\mathbb{R}^d$ with $\overline{\mathcal{M}}^{d-1}(K) < \infty$, then there exists a constant C such that*

$$\mathcal{H}^{d-1}(\partial K_r) \le C(1 + r^{d-1}), \ r > 0. \tag{4.11}$$

Proof. By Corollary 3.6 in [34], the hypotheses yield $r_0, C_1 \in (0, \infty)$ such that

$$\mathcal{H}^{d-1}(\partial K_r) \le C_1, \ r \in (0, r_0).$$

By Lemma 4.2 in [34], for any $0 < a < b$ there is a constant $C_2 := C_2(a, b, K)$ such that

$$\mathcal{H}^{d-1}(\partial K_r) \le C_2, \ r \in (a, b).$$

Fixing $a \in (0, r_0)$ and $b \in (\operatorname{diam}(K), \infty)$ it follows that

$$\mathcal{H}^{d-1}(\partial K_r) \le \max(C_1, C_2), \quad r \in (0, b). \tag{4.12}$$

By Kneser's lemma for parallel sets (cf. [47]), we have for any $t \in (1, \infty)$ and $\varepsilon > 0$,

$$\frac{\operatorname{Vol}_d(K_{tb}) - \operatorname{Vol}_d(K_{t(b-\varepsilon)})}{t\varepsilon} \le t^{d-1} \frac{\operatorname{Vol}_d(K_b) - \operatorname{Vol}_d(K_{b-\varepsilon})}{\varepsilon}.$$

Letting $\varepsilon \downarrow 0$ gives for all $t \in (1, \infty)$ that $(\operatorname{Vol}_d)'_-(K_{tb}) \le t^{d-1}(\operatorname{Vol}_d)'_-(K_b)$, where we use that left-derivatives of the volume function $r \to \operatorname{Vol}_d(K_r)$, here denoted $(\operatorname{Vol}_d)'_-$, exist everywhere.

The derivative of the function $t \mapsto \operatorname{Vol}_d(K_{tb})$ is the limit of

$$\frac{\operatorname{Vol}_d(K_{tb}) - \operatorname{Vol}_d(K_{t(b-\varepsilon)})}{t\varepsilon},$$

as ε tends to zero. Fix $t \in (1, \infty)$. As noted in Section 2 of [34], since $tb \in (\operatorname{diam}(K), \infty)$, it follows that the stated left-derivatives coincide with derivatives, and are respectively equal to $\mathcal{H}^{d-1}(\partial K_{tb})$ and $\mathcal{H}^{d-1}(\partial K_b)$ (cf. Corollary 2.5 in [34]). Putting $r = tb$ gives

$$\mathcal{H}^{d-1}(\partial K_r) \le r^{d-1} b^{-d+1} \mathcal{H}^{d-1}(\partial K_b), \ r \in (b, \infty). \tag{4.13}$$

Putting $C_3 := b^{-d+1}\mathcal{H}^{d-1}(\partial K_b)$ and combining (4.12) and (4.13) gives (4.11) with $C := \max(C_1, C_2, C_3)$. □

Proof of Theorem 2.3. Note that we have the same situation as described in Example 1 in Section 2. In the following we evaluate $I_{K,s}$, which allows us to apply Corollary 2.2. It suffices to show that if K is a full d-dimensional subset of $\mathbb{X}$, then $I_{K,s} = O(s)$, while $I_{K,s} = O(s^{1-1/d})$ for lower dimensional K. Indeed, put $c := \min\{c_{stab}, c_K\}p/(36 \cdot 4^{\alpha+1})$, so that

$$\begin{aligned}
I_{K,s} &= s\int_{\mathbb{X}} \exp(-c\,\mathrm{d}_s(x,K)^\alpha) g(x)\,\mathrm{d}x \\
&\le \|g\|_\infty s \int_K \exp(-cs^{\alpha/d}\,\mathrm{d}(x,K)^\alpha)\,\mathrm{d}x + \|g\|_\infty s\int_{\mathbb{X}\setminus K} \exp(-cs^{\alpha/d}\,\mathrm{d}(x,K)^\alpha)\,\mathrm{d}x \\
&= \|g\|_\infty \mathrm{Vol}_d(K)s + \|g\|_\infty s \int_0^\infty \int_{\partial K_r} \exp(-cs^{\alpha/d} r^\alpha)\,\mathcal{H}^{d-1}(\mathrm{d}y)\,\mathrm{d}r \\
&\le \|g\|_\infty \mathrm{Vol}_d(K)s + C\|g\|_\infty s \int_0^\infty \exp(-cs^{\alpha/d} r^\alpha)\,(1 + r^{d-1})\,\mathrm{d}r \\
&\le \|g\|_\infty \mathrm{Vol}_d(K)s + C\|g\|_\infty s^{(d-1)/d} \int_0^\infty \exp(-cu^\alpha)\,(1 + s^{-(d-1)/d}u^{d-1})\,\mathrm{d}u,
\end{aligned}$$

where the second equality follows by the co-area formula and where the second inequality follows by Lemma 4.12. If K is a full d-dimensional subset of $\mathbb{X}$, then the first integral dominates and is $O(s)$. Otherwise, $\mathrm{Vol}_d(K)$ vanishes and the second integral is $O(s^{(d-1)/d})$. □

5 Applications

By appropriately choosing the measure space $(\mathbb{X}, \mathcal{F}, \mathbb{Q})$, the scores $(\xi_s)_{s\ge1}$ and $(\xi_n)_{n\in\mathbb{N}}$, and the set $K \subset \mathbb{X}$, we may use the general results of Theorem 2.1, Corollary 2.2 and Theorem 2.3 to deduce presumably optimal rates of normal convergence for statistics in geometric probability. For example, in the setting $\mathbb{X} = \mathbb{R}^d$, we expect that all of the statistics H_s and H'_n described in [7, 26, 29, 30, 31] consist of sums of scores ξ_s and ξ_n satisfying the conditions of Theorem 2.3, showing that the statistics in these papers enjoy rates of normal convergence (in the Kolmogorov distance) given by the reciprocal of the standard deviation of H_s and H'_n, respectively. Previously, the rates in these papers either contained extraneous logarithmic factors, as in the case of Poisson input, or the rates were sometimes non-existent, as in the case of binomial input. In the following we do this in detail for some prominent statistics featuring in the stochastic geometry literature, including the k-face and intrinsic volume functionals of convex hulls of random samples. In some instances the rates of convergence are subject to variance lower bounds, a separate problem not addressed here.

We believe that one could use our approach to also deduce presumably optimal rates of normal convergence for statistics of random sequential packing problems as in [42], set approximation via Delaunay triangulations as in [19], generalized spacings as in [8], and general proximity graphs as in [16]. We mention here additional potential applications of our general results. The list is far from exhaustive and is meant to illustrate the wide applicability of Theorem 2.1 and the relative simplicity of Corollary 2.2 and Theorem 2.3.

5.1 Nearest neighbor graphs and statistics of high-dimensional data sets

a. *Total edge length of nearest neighbor graphs.* Let $(\mathbb{X}, \mathcal{F}, \mathbb{Q})$ be equipped with a semi-metric d such that (2.1) is satisfied for some γ and κ. We equip $\mathbb{X}$ with a fixed linear order, which is possible by the well-ordering principle. Given $\mathcal{X} \in \mathbf{N}$, $k \in \mathbb{N}$, and $x \in \mathcal{X}$, let $V_k(x, \mathcal{X})$ be the set of k nearest neighbors of x, i.e., the k closest points of x in $\mathcal{X} \setminus \{x\}$. In case that that these k points are not unique, we break the tie via the fixed linear order on $\mathbb{X}$. The (undirected) nearest neighbor graph $NG_1(\mathcal{X})$ is the graph with vertex set $\mathcal{X}$ obtained by including an edge $\{x, y\}$ if $y \in V_1(x, \mathcal{X})$ and/or $x \in V_1(y, \mathcal{X})$. More generally, the (undirected) k-nearest neighbors graph $NG_k(\mathcal{X})$ is the graph with vertex set $\mathcal{X}$ obtained by including an edge $\{x, y\}$ if $y \in V_k(x, \mathcal{X})$ and/or $x \in V_k(y, \mathcal{X})$. For all $q \in \mathbb{R}$ define

$$\xi^{(q)}(x, \mathcal{X}) := \sum_{y \in V_k(x, \mathcal{X})} \rho^{(q)}(x, y), \tag{5.1}$$

where $\rho^{(q)}(x, y) := \mathrm{d}(x, y)^q/2$ if x and y are mutual k-nearest neighbors, i.e., $x \in V_k(y, \mathcal{X})$ and $y \in V_k(x, \mathcal{X})$, and otherwise $\rho^{(q)}(y, x) := \mathrm{d}(x, y)^q$. The total edge length of the undirected k-nearest neighbors graph on $\mathcal{X}$ with qth power weighted edges is

$$L^{(q)}_{NG_k}(\mathcal{X}) = \sum_{x \in \mathcal{X}} \xi^{(q)}(x, \mathcal{X}).$$

As usual $\mathcal{P}_s$ is a Poisson point process on $\mathbb{X}$ with intensity measure $s\mathbb{Q}$ and $\mathcal{X}_n$ is a binomial point process of n points in $\mathbb{X}$ distributed according to $\mathbb{Q}$. We assume in the following that $(\mathbb{X}, \mathcal{F}, \mathbb{Q})$ satisfies (2.1) and

$$\inf_{x \in \mathbb{X}} \mathbb{Q}(B(x, r)) \geq c r^{\gamma}, \quad r \in [0, \mathrm{diam}(\mathbb{X})], \tag{5.2}$$

where γ is the constant from (2.1) and $c > 0$.

Theorem 5.1. *If* $q \geq 0$ *and* $\mathrm{Var}(L^{(q)}_{NG_k}(\mathcal{P}_s)) = \Omega(s^{1-2q/\gamma})$*, then there is a* $\tilde{C} := \tilde{C}(k) \in (0, \infty)$ *such that*

$$d_K\left(\frac{L^{(q)}_{NG_k}(\mathcal{P}_s) - \mathbb{E}\, L^{(q)}_{NG_k}(\mathcal{P}_s)}{\sqrt{\mathrm{Var}\, L^{(q)}_{NG_k}(\mathcal{P}_s)}}, N\right) \leq \frac{\tilde{C}}{\sqrt{s}}, \quad s \geq 1, \tag{5.3}$$

whereas if $\mathrm{Var}(L^{(q)}_{NG_k}(\mathcal{X}_n)) = \Omega(n^{1-2q/\gamma})$*, then*

$$d_K\left(\frac{L^{(q)}_{NG_k}(\mathcal{X}_n) - \mathbb{E}\, L^{(q)}_{NG_k}(\mathcal{X}_n)}{\sqrt{\mathrm{Var}\, L^{(q)}_{NG_k}(\mathcal{X}_n)}}, N\right) \leq \frac{\tilde{C}}{\sqrt{n}}, \quad n \geq 9. \tag{5.4}$$

Remarks. (i) *Comparison with previous work.* Research has focused on central limit theorems for $L^{(q)}_{NG_k}(\mathcal{P}_s), s \to \infty$, and $L^{(q)}_{NG_k}(\mathcal{X}_n), n \to \infty$, when $\mathbb{X}$ is a subset of $\mathbb{R}^d$ and where d is the usual Euclidean distance. This includes the seminal work [9], the paper [1] and the more recent works [28, 29, 31]. When $\mathbb{X}$ is a sub-manifold of $\mathbb{R}^d$ equipped with the Euclidean metric on $\mathbb{R}^d$, the paper [32] develops the limit theory for $L^{(q)}_{NG_k}(\mathcal{P}_s), s \to \infty$, and $L^{(q)}_{NG_k}(\mathcal{X}_n), n \to \infty$. When $\mathbb{X}$ is a subset of $\mathbb{R}^d$ the paper [23] establishes the presumably optimal $O(s^{-1/2})$ rate of normal convergence for $L^{(q)}_{NG_k}(\mathcal{P}_s)$. However these papers neither provide the presumably optimal $O(n^{-1/2})$ rate of normal convergence for $L^{(q)}_{NG_k}(\mathcal{X}_n)$ in the d_K distance, nor do they consider input on arbitrary metric spaces. Theorem 5.1 rectifies this and also allows $\mathbb{Q}$ to have an unbounded density.

(ii) *Binomial input.* The rate for binomial input (5.4) improves upon the rate of convergence in the Wasserstein distance d_W given by

$$d_W\left(\frac{L^{(q)}_{NG_k}(\mathcal{X}_n) - \mathbb{E}\, L^{(q)}_{NG_k}(\mathcal{X}_n)}{\sqrt{\mathrm{Var}\, L^{(q)}_{NG_k}(\mathcal{X}_n)}}, N\right) = O\left(\frac{k^4 \tilde{\gamma}_p^{2/p}}{n^{(p-8)/2p}} + \frac{k^3 \tilde{\gamma}_p^{3/p}}{n^{(p-6)/2p}}\right), \tag{5.5}$$

as in Theorem 3.4 of [11] as well as the same rate in the Kolmogorov distance as in Section 6.3 of [22]. Here $\tilde{\gamma}_p := \mathbb{E}\, |n^{q/\gamma} \xi^{(q)}(X_1, \mathcal{X}_n)|^p$ and $p > 8$. For all $\varepsilon > 0$ we have $\mathbb{P}(n^{q/\gamma} \xi^{(q)}(X_1, \mathcal{X}_n) > \varepsilon) = (1 - C\varepsilon^\gamma / n)^n$ and it follows that $\tilde{\gamma}_p^{1/p} \uparrow \infty$ as $p \to \infty$. Thus by letting $p \to \infty$, we do not recover the $O(n^{-1/2})$ rate in (5.5), but only achieve the rate $O(n^{-1/2} (\log n)^\tau)$ with some $\tau > 0$.

(iii) *Variance bounds.* When $\mathbb{X}$ is a full-dimensional compact convex subset of $\mathbb{R}^d$, then $\gamma = d$, $\mathrm{Var}(L^{(q)}_{NG_k}(\mathcal{P}_s)) = \Theta(s^{1-2q/\gamma})$, and $\mathrm{Var}(L_{NG_k}(\mathcal{X}_n)) = \Theta(n^{1-2q/\gamma})$, which follows from Theorem 2.1 and Lemma 6.3 of [29] (these results treat the case $q = 1$ but the proofs easily extend to arbitrary $q \in (0, \infty)$). Thus we obtain the required variance lower bounds of Theorem 5.1. If $\mathrm{Var}(L^{(q)}_{NG_k}(\mathcal{P}_s)) = \Omega(s^{1-2q/\gamma})$ does not hold, then the convergence rate in (5.3) is replaced by (2.11) with $I_{K,s}$ set to s, with a similar statement if $\mathrm{Var}(L_{NG_k}(\mathcal{X}_n)) = \Omega(n^{1-2q/\gamma})$ does not hold.

(iv) *Extension of Theorem 5.1.* The *directed* k-nearest neighbors graph, denoted $NG'_k(\mathcal{X})$, is the directed graph with vertex set $\mathcal{X}$ obtained by including a directed edge from each point to each of its k nearest neighbors. The total edge length of the directed k-nearest neighbors graph on $\mathcal{X}$ with qth power-weighted edges is

$$L^{(q)}_{NG'_k}(\mathcal{X}) = \sum_{x \in \mathcal{X}} \tilde{\xi}^{(q)}(x, \mathcal{X})$$

where

$$\tilde{\xi}^{(q)}(x, \mathcal{X}) := \sum_{y \in V_k(x, \mathcal{X})} \mathrm{d}(x, y)^q.$$

The proof of Theorem 5.1 given below shows that the analogs of (5.3) and (5.4) hold for $L^{(q)}_{NG'_k}(\mathcal{P}_s)$ and $L^{(q)}_{NG'_k}(\mathcal{X}_n)$ as well.

Proof. In the following we prove (5.3). We deduce this from Corollary 2.2 with $\xi_s(x, \mathcal{P}_s)$ set to $s^{q/\gamma}\xi^{(q)}(x, \mathcal{P}_s)$, with $\xi^{(q)}$ as at (5.1) and with K set to $\mathbb{X}$. Recalling the terminology of Corollary 2.2, we have $H_s := s^{q/\gamma} L^{(q)}_{NG_k}(\mathcal{P}_s) = \sum_{x \in \mathcal{P}_s} \xi_s(x, \mathcal{P}_s)$, with $\operatorname{Var} H_s = \operatorname{Var}(s^{q/\gamma} \sum_{x \in \mathcal{P}_s} \xi^{(q)}(x, \mathcal{P}_s)) = \Omega(s)$, by assumption. Recall from (2.13) that $I_{K,s} = \Theta(s)$. We claim that $R_s(x, \mathcal{X} \cup \{x\}) := 3\,\mathrm{d}(x, x_{kNN}(x, \mathcal{X} \cup \{x\}))$ is a radius of stabilization for $\xi_s(x, \mathcal{X} \cup \{x\})$, where $x_{kNN}(x, \mathcal{X} \cup \{x\})$ is the point of $V_k(x, \mathcal{X} \cup \{x\})$ with the maximal distance to x. Indeed, if a point y is a k-nearest neighbor of x, then all of its k-nearest neighbors must belong to $B(y, 2\,\mathrm{d}(x, x_{kNN}(x, \mathcal{X} \cup \{x\})))$, since this ball contains with x and its $k-1$ nearest neighbors enough potential k-nearest neighbors for y.

We now show that $R_s(x, \mathcal{P}_s \cup \{x\})$ satisfies exponential stabilization (2.4). Notice that

$$\mathbb{P}(R_s(x, \mathcal{P}_s \cup \{x\}) > r) = \mathbb{P}(\mathcal{P}_s(B(x, r/3) < k)), \quad r \geq 0.$$

The number of points from $\mathcal{P}_s$ in $B(x, r/3)$ follows a Poisson distribution with parameter $s\mathbb{Q}(B(x, r/3))$. By (5.2) this exceeds $cs(r/3)^\gamma$ if $r \in [0, 3\,\mathrm{diam}(\mathbb{X})]$. By a Chernoff bound for the Poisson distribution (e.g. Lemma 1.2 of [25]), there is another constant $\tilde{c} \in (0, \infty)$ such that

$$\mathbb{P}(R_s(x, \mathcal{P}_s \cup \{x\}) > r) \leq k \exp(-\tilde{c} s r^\gamma), \quad r \in [0, 3\,\mathrm{diam}(\mathbb{X})].$$

This also holds for $r > 3\,\mathrm{diam}(\mathbb{X})$, since $\mathbb{P}(R_s(x, \mathcal{P}_s \cup \{x\}) \geq r) = 0$ in this case. This gives (2.4), with $\alpha_{stab} = \gamma$, $c_{stab} = \tilde{c}$, and $C_{stab} = k$. We may modify this argument to obtain exponential stabilization with respect to binomial input as at (2.5).

For all $q \in (0, \infty)$, the scores $(\xi_s)_{s \geq 1}$ also satisfy the $(4+p)$-moment condition (2.6) for all $p \in [0, \infty)$ since

$$\xi_s(x, \mathcal{P}_s \cup \{x\} \cup \mathcal{A}) \leq k s^{q/\gamma}\,\mathrm{d}(x, x_{kNN}(x, \mathcal{P}_s \cup \{x\}))^q$$

for all $\mathcal{A} \subset \mathbb{X}$ with $|\mathcal{A}| \leq 7$, and the above computation shows that $s^{q/\gamma}\,\mathrm{d}(x, x_{kNN}(x, \mathcal{P}_s \cup \{x\}))^q$ has an exponentially decaying tail. The bound (5.3) follows by Corollary 2.2. The proof of (5.4) is similar. This completes the proof of Theorem 5.1. □

b. *Statistics of high-dimensional data sets.* In the case that $\mathbb{X}$ is an m-dimensional C^1 sub-manifold of $\mathbb{R}^d$, with d the Euclidean distance in $\mathbb{R}^d$, the *directed* nearest neighbors graph version of Theorem 5.1 (cf. Remark (iii) above) may be refined to give rates of normal convergence for statistics of high-dimensional non-linear data sets. This goes as follows. Recall that high-dimensional non-linear data sets are typically modeled as the realization of $\mathcal{X}_n := \{X_1, ..., X_n\}$, with $X_i, 1 \leq i \leq n$, i.i.d. copies of a random variable X having support on an unknown (non-linear) manifold $\mathbb{X}$ embedded in $\mathbb{R}^d$. Typically the coordinate representation of X_i is unknown, but the interpoint distances are known. Given this information, the goal is to establish estimators of global characteristics of $\mathbb{X}$, including intrinsic dimension, as well as global properties of the distribution of X, such as Rényi entropy. Recall that if the distribution of the random variable X has a Radon-Nikodym derivative f_X with respect to the uniform measure on $\mathbb{X}$, then given $\rho \in (0, \infty), \rho \neq 1$, the Rényi ρ-entropy of X is

$$H_\rho(f_X) := (1 - \rho)^{-1} \log \int_{\mathbb{X}} f_X(x)^\rho \,\mathrm{d}x.$$

Let $\mathbb{X}$ be an m-dimensional subset of $\mathbb{R}^d$, $m \leq d$, equipped with the Euclidean metric d on $\mathbb{R}^d$. Let $\mathbb{Q}$ be a measure on $\mathbb{X}$ with a bounded density f_X with respect to the uniform surface measure on $\mathbb{X}$. We assume $\mathbb{X}$ satisfies condition (2.1) with $\gamma := m$ (recall that Example 2 (Section 2) gives conditions guaranteeing (2.1)). Henceforth, *assume* $\mathbb{X}$ *is an m-dimensional* C^1 *submanifold-with-boundary* (see Section 2.1 of [32] for details and precise definitions). Assume f_X is bounded away from zero and infinity, and

$$\inf_x \mathbb{Q}(B(x, r)) \geq cr^m, \quad r \in [0, \mathrm{diam}(\mathbb{X})],$$

with some constant $c \in (0, \infty)$. The latter condition is called the 'locally conic' condition in [32] (cf. (2.3) in [32]).

Under the above conditions and given Poisson input $\mathcal{P}_s$ with intensity sf_X, the main results of [32] establish rates of normal convergence for estimators of intrinsic dimension, estimators of Rényi entropy, and for Vietoris-Rips clique counts (see Section 2 of [32] for precise statements). However these rates contain extraneous logarithmic factors and [32] also stops short of establishing rates of normal convergence when Poisson input is replaced by binomial input. In what follows we rectify this for estimators of Rényi entropy. The methods potentially apply to yield rates of normal convergence for estimators of Shannon entropy and intrinsic dimension, but this lies beyond the scope of this paper.

When f_X satisfies the assumptions stated above and is also continuous on $\mathbb{X}$, then $n^{q/m-1} L^{(q)}_{NG'_1}(\mathcal{X}_n)$ is a consistent estimator of a multiple of $\int_{\mathbb{X}} f_X(x)^{1-q/m} \, \mathrm{d}x$, as shown in Theorem 2.2 of [32]. The following result establishes a rate of normal convergence for $L^{(q)}_{NG'_k}(\mathcal{X}_n)$ and, in particular, for the estimator $n^{q/m-1} L^{(q)}_{NG'_1}(\mathcal{X}_n)$.

Theorem 5.2. *If $k \in \mathbb{N}$ and $q \in (0, \infty)$, then there is a constant $c \in (0, \infty)$ such that*

$$d_K \left(\frac{L^{(q)}_{NG'_k}(\mathcal{X}_n) - \mathbb{E}\, L^{(q)}_{NG'_k}(\mathcal{X}_n)}{\sqrt{\operatorname{Var} L^{(q)}_{NG'_k}(\mathcal{X}_n)}}, N \right) \leq \frac{c}{\sqrt{n}}, \quad n \geq 9. \tag{5.6}$$

A similar result holds if the binomial input $\mathcal{X}_n$ is replaced by Poisson input.

Remarks. (i) We have to exclude the case $q = 0$ since $L^{(0)}_{NG'_1}(\mathcal{X}_n) = kn$ if $n > k$. For the Poisson case a central limit theorem still holds, but becomes trivial since we have $L^{(0)}_{NG'_1}(\mathcal{P}_s) = k|\mathcal{P}_s|$ if $|\mathcal{P}_s| \geq k + 1$.
(ii) In the same vein as Remark (ii) following Theorem 5.1, Theorem 3.4 of [11] yields a rate of normal convergence for $L^{(q)}_{NG'_1}(\mathcal{X}_n)$ in the Wasserstein distance d_W given by the right-hand side of (5.5). However, the bound (5.6) is superior and is moreover expressed in the Kolmogorov distance d_K.When the input $\mathcal{X}_n$ is replaced by Poisson input $\mathcal{P}_s$, we obtain the rate of normal convergence $O(s^{-1/2})$, improving upon the rates of [31, 32].

Proof. Appealing to the method of proof in Theorem 2.3 of [32] and the variance lower bounds of Theorem 6.1 of [29], we see that $\operatorname{Var} L^{(q)}_{NG'_k}(\mathcal{X}_n) = \Theta(n^{1-2q/m})$ and $\operatorname{Var} L^{(q)}_{NG'_k}(\mathcal{P}_s) = \Theta(s^{1-2q/m})$. The proof follows now the proof of Theorem 5.1. □

5.2 Maximal points

Consider the cone $\mathrm{Co} = (\mathbb{R}^+)^d$ with apex at the origin of $\mathbb{R}^d$. Given $\mathcal{X} \in \mathbf{N}$, $x \in \mathcal{X}$ is called maximal if $(\mathrm{Co} \oplus x) \cap \mathcal{X} = x$. In other words, a point $x = (x_1, ..., x_d) \in \mathcal{X}$ is maximal if there is no other point $(z_1, ..., z_d) \in \mathcal{X}$ with $z_i \geq x_i$ for all $1 \leq i \leq d$. The maximal layer $m_{\mathrm{Co}}(\mathcal{X})$ is the collection of maximal points in $\mathcal{X}$. Let $M_{\mathrm{Co}}(\mathcal{X}) := \mathrm{card}(m_{\mathrm{Co}}(\mathcal{X}))$. Maximal points are of broad interest in computational geometry and economics; see the books [33], [12], and the survey [46].

Put

$$\mathbb{X} := \{x \in [0, \infty)^d : \; F(x) \leq 1\}$$

where $F : [0, \infty)^d \to \mathbb{R}^+$ is a strictly increasing function of each coordinate variable, satisfies $F(0) < 1$, is continuously differentiable, and has continuous partials F_i, $1 \leq i \leq d$, bounded away from zero and infinity. Let $\mathbb{Q}$ be a measure on $\mathbb{X}$ with Radon-Nikodym derivative g with respect to Lebesgue measure on $\mathbb{X}$, with g bounded away from zero and infinity. As usual, $\mathcal{P}_s$ is the Poisson point process with intensity $s\mathbb{Q}$ and $\mathcal{X}_n$ is a binomial point process of n i.i.d. points distributed according to $\mathbb{Q}$.

Theorem 5.3. *We have*

$$d_K\left(\frac{M_{\mathrm{Co}}(\mathcal{P}_s) - \mathbb{E}\, M_{\mathrm{Co}}(\mathcal{P}_s)}{\sqrt{\mathrm{Var}\, M_{\mathrm{Co}}(\mathcal{P}_s)}}, N\right) \le c s^{\frac{-1}{2}+\frac{1}{2d}}, \quad s \ge 1. \tag{5.7}$$

Assuming $\mathrm{Var}\, M_{\mathrm{Co}}(\mathcal{X}_n) = \Omega(n^{(d-1)/d})$, *the binomial counterpart to* (5.7) *holds, with* $\mathcal{P}_s$ *replaced by* $\mathcal{X}_n$.

Remarks. (i) *Existing results.* The rates of normal convergence given by Theorem 5.3 improve upon those given in [5] for Poisson and binomial input for the bounded Wasserstein distance and in [6] and [50] for Poisson input for the Kolmogorov distance. While these findings are also proved via the Stein method, the local dependency methods employed there all incorporate extraneous logarithmic factors. Likewise, when $d = 2$, the paper [2] provides rates of normal convergence in the d_K distance for binomial input, but aside from the special case that F is linear, the rates incorporate extraneous logarithmic factors. The precise approximation bounds of Theorem 5.3 remove the logarithmic factors in [2, 5, 6, 50].

(ii) We have taken $\mathrm{Co} = (\mathbb{R}^+)^{\mathrm{d}}$ to simplify the presentation, but the results extend to general cones which are subsets of $(\mathbb{R}^+)^d$ and which have apex at the origin.

Proof of Theorem 5.3. We only prove (5.7) for Poisson input, as the proof for binomial input is similar. We deduce this theorem from Theorem 2.3(b) and consider score functions

$$\zeta(x, \mathcal{X}) := \begin{cases} 1 & \text{if } ((\mathrm{Co} \oplus x) \cap \mathbb{X}) \cap \mathcal{X} = x, \\ 0 & \text{otherwise.} \end{cases}$$

Notice that $M_{\mathrm{Co}}(\mathcal{P}_s) = \sum_{x \in \mathcal{P}_s} \zeta(x, \mathcal{P}_s)$.

Put $K := \{x \in [0,\infty)^d : \ F(x) = 1\}$. The assumptions on F imply $\overline{\mathcal{M}}^{d-1}(K) < \infty$. Taking $\alpha = d$, we get $I_{K,s} = O(s^{(d-1)/d})$ as shown in the proof of Theorem 2.3. Thus we only need to show that the scores $\zeta_s \equiv \zeta$ satisfy the conditions of Theorem 2.3(b). First, ζ is bounded and so satisfies the $(4+p)$-moment condition (2.6) for all $p \in [0, \infty)$. As shown in [50] (see proof of Theorem 2.5 in Section 6), ζ_s also satisfy exponential stabilization (2.4) with $\alpha_{stab} = d$ with respect to Poisson input $\mathcal{P}_s$. Also, we assert that the scores decay exponentially fast with the distance to K with $\alpha_K = d$. To see this, let $r(x) := \mathrm{d}(x, K)$ be the distance between x and K and note that $(\mathrm{Co} \oplus x) \cap \mathbb{X}$ contains the set $S(x) := B(x, r(x)) \cap (\mathrm{Co} \oplus x)$. It follows that

$$\mathbb{P}(\zeta(x, \mathcal{P}_s \cup \{x\}) \neq 0) = \mathbb{P}((\mathrm{Co} \oplus x) \cap \mathbb{X}) \cap \mathcal{P}_s = x) = \exp\left(-s \int_{(\mathrm{Co} \oplus x) \cap \mathbb{X}} d\mathbb{Q}\right)$$
$$\leq \exp(-s\mathbb{Q}(S(x))) \leq \exp(-c \, \mathrm{d}_s(x, K)^d)$$

with some constant $c := c(\mathbb{Q}) \in (0, \infty)$, and thus (2.8) holds with $\alpha_K = d$.

We now show $\mathrm{Var}\, M_{\mathrm{Co}}(\mathcal{P}_s) = \Theta(s^{(d-1)/d})$. The hypotheses on F imply that there are $N = \Theta(s^{(d-1)/d})$ disjoint sets $S_i := (\mathrm{Co} \oplus x_i) \cap \mathbb{X}, i = 1, ..., N$, with $x_i \in \mathbb{X}$, such that $Q_i := [0, \delta s^{-1/d}]^d \oplus x_i \subset S_i$ Given x_i, for $1 \leq j \leq d$ define d sub-cubes of Q_i

$$Q_{ij} := \Big(\frac{2\delta}{3} s^{-1/d}, \delta s^{-1/d}\Big]^{j-1} \times [0, \frac{\delta}{3} s^{-1/d}) \times \Big(\frac{2\delta}{3} s^{-1/d}, \delta s^{-1/d}\Big]^{d-j} \oplus x_i,$$

as well as the central cube $\tilde{Q}_i := \Pi_{j=1}^d [\frac{\delta}{3} s^{-1/d}, \frac{2\delta}{3} s^{-1/d}] \oplus x_i$. All cubes thus constructed are disjoint. Say that $S_i, 1 \leq i \leq N$, is admissible if there are points

$$p_{ij} \in \mathcal{P}_s \cap Q_{ij}, \ 1 \leq j \leq d,$$

which are maximal and $Q_i \backslash \tilde{Q}_i$ contains no other points in $\mathcal{P}_s$. Given that S_i is admissible, we assert that the maximality status of points in $\mathcal{P}_s \cap \tilde{Q}_i^c$ is unaffected by the (possibly empty) configuration of $\mathcal{P}_s$ inside $\tilde{Q}_i$. Indeed, if $x \in \mathcal{P}_s \cap \tilde{Q}_i^c \cap Q_i$, then $x \in \{p_{ij}\}_{j=1}^d$ and so $(\mathrm{Co} \oplus x) \cap \tilde{Q}_i = \emptyset$, showing the assertion in this case. On the other hand, if $x \in \mathcal{P}_s \cap \tilde{Q}_i^c \cap Q_i^c$ and if $(\mathrm{Co} \oplus x) \cap \tilde{Q}_i \neq \emptyset$, then $\mathrm{Co} \oplus x$ must contain at least one of the cubes Q_{ij}, thus $\mathrm{Co} \oplus x$ contains at least one of the points $\{p_{ij}\}_{j=1}^d$ and hence $\zeta(x, \mathcal{P}_s)$ vanishes. Let I be the indices $i \in \{1, ..., N\}$ such that S_i is admissible.

Let $\mathcal{F}_s$ be the sigma algebra generated by I and $\mathcal{P}_s \cap (\mathbb{X} \setminus \cup_{i \in I} \tilde{Q}_i)$, including the maximal points $\{\{p_{ij}\}_{j=1}^d\}_{i \in I}$. Conditional on $\mathcal{F}_s$, note that $\zeta(x, \mathcal{P}_s)$ is deterministic for $x \in \mathcal{P}_s \cap (\mathbb{X} \setminus \cup_{i \in I} \tilde{Q}_i)$. The conditional variance formula gives

$$\begin{aligned}
\mathrm{Var}\, M_{\mathrm{Co}}(\mathcal{P}_s) &\geq \mathbb{E} \ \mathrm{Var}[\sum_{x \in \mathcal{P}_s \cap \cup_{i \in I} \tilde{Q}_i} \zeta(x, \mathcal{P}_s) + \sum_{x \in \mathcal{P}_s \cap (\mathbb{X} \setminus \cup_{i \in I} \tilde{Q}_i))} \zeta(x, \mathcal{P}_s) |\, \mathcal{F}_s] \\
&= \mathbb{E} \ \mathrm{Var}[\sum_{i \in I} \sum_{x \in \mathcal{P}_s \cap \tilde{Q}_i} \zeta(x, \mathcal{P}_s) |\, \mathcal{F}_s] \\
&= \mathbb{E} \sum_{i \in I} \mathrm{Var}[\sum_{x \in \mathcal{P}_s \cap \tilde{Q}_i} \zeta(x, \mathcal{P}_s) |\, \mathcal{F}_s]
\end{aligned}$$

where the last equality follows by independence of $\sum_{x \in \mathcal{P}_s \cap \tilde{Q}_i} \zeta(x, \mathcal{P}_s), i \in I$. The bounds on g imply that there is a set of positive probability, so that on this set and conditional on $\mathcal{F}_s$, there are either one or two maximal points in $\tilde{Q}_i$. In other words, on this set the random variable $\sum_{x \in \mathcal{P}_s \cap \tilde{Q}_i} \zeta(x, \mathcal{P}_s)$ exhibits non-zero variability, showing

$\mathrm{Var}[\sum_{x \in \mathcal{P}_s \cap \tilde{Q}_i} \zeta(x, \mathcal{P}_s) | \mathcal{F}_s] \geq c > 0$ uniformly in $i \in I$. Since $\mathbb{E}\,\mathrm{card}(I) \geq cs^{(d-1)/d}$ the asserted variance lower bound follows. Theorem 2.3(b) gives (5.7).

The proof method in [50] is for Poisson input $\mathcal{P}_s$, but it may be easily extended to show that the $(\zeta_n)_{n \geq 1}$ are exponentially stabilizing with respect to binomial input and that $(\zeta_n)_{n \geq 1}$ decay exponentially fast with the distance to K. Thus the conditions of Theorem 2.3(b) are satisfied, and so (5.7) follows from (2.14), concluding the proof of Theorem 5.3. □

5.3 Set approximation via Voronoi tessellations

Throughout this subsection let $\mathbb{X} := [-1/2, 1/2]^d, d \geq 2$, and let $A \subset \mathrm{int}(\mathbb{X})$ be a full-dimensional subset of $\mathbb{R}^d$. Let $\mathbb{Q}$ be the uniform measure on $\mathbb{X}$. For $\mathcal{X} \in \mathcal{M}$ and $x \in \mathcal{X}$ the Voronoi cell $C(x, \mathcal{X})$ is the set of all $z \in \mathbb{X}$ such that the distance between z and x is at most equal to the distance between z and any other point of $\mathcal{X}$. The collection of all $C(x, \mathcal{X})$ with $x \in \mathcal{X}$ is called the Voronoi tessellation of $\mathbb{X}$. The Voronoi approximation of A with respect to $\mathcal{X}$ is the union of all Voronoi cells $C(x, \mathcal{X}), x \in \mathcal{X}$, with $x \in A$, i.e.,

$$A(\mathcal{X}) := \bigcup_{x \in \mathcal{X} \cap A} C(x, \mathcal{X}).$$

In the following we let $\mathcal{X}$ be either the Poisson point process $\mathcal{P}_s$, $s \geq 1$, with intensity measure $s\mathbb{Q}$ or a binomial point process $\mathcal{X}_n$ of $n \in \mathbb{N}$ points distributed according to $\mathbb{Q}$. We are now interested in the behavior of the random approximations

$$A_s := A(\mathcal{P}_s), \quad s \geq 1, \quad \text{and} \quad A'_n := A(\mathcal{X}_n), \quad n \in \mathbb{N},$$

of A. Note that A_s is also called the Poisson-Voronoi approximation.

Typically A is an unknown set having unknown geometric characteristics such as volume and surface area. Notice that A_s and A_n are random polyhedral approximations of A, with volumes closely approximating that of A as s and n become large. There is a large literature devoted to quantifying this approximation and we refer to [22, 50] for further discussion and references. One might also expect that $\mathcal{H}^{d-1}(\partial A_s)$ closely approximates a scalar multiple of $\mathcal{H}^{d-1}(\partial A)$, provided the latter quantity exists and is finite. This has been shown in [50]. Using Theorem 2.3(b) we deduce rates of normal convergence for the volume and surface area statistics of A_s and A_n as well as $\mathrm{Vol}(A_s \Delta A)$ and $\mathrm{Vol}(A_n \Delta A)$. Here and elsewhere in this section we abbreviate Vol_d by Vol. The symmetric difference $U \Delta V$ of two sets $U, V \subset \mathbb{R}^d$ is given by $U \Delta V := (U \setminus V) \cup (V \setminus U)$.

Theorem 5.4. *Let $A \subset (-1/2, 1/2)^d$ be such that ∂A satisfies $\overline{\mathcal{M}}^{d-1}(\partial A) < \infty$ and let $F \in \{\mathrm{Vol}, \mathrm{Vol}(\cdot \Delta A), \mathcal{H}^{d-1}(\partial \cdot)\}$. If ∂A contains a $(d-1)$-dimensional C^2 submanifold* *then there is a constant* $\tilde{C} \in (0, \infty)$ *such that*

$$d_K\left(\frac{F(A_s) - \mathbb{E}\, F(A_s)}{\sqrt{\operatorname{Var} F(A_s)}}, N\right) \leq \tilde{C} s^{-\frac{(d-1)}{2d}}, \quad s \geq 1, \tag{5.8}$$

and

$$d_K\left(\frac{F(A'_n) - \mathbb{E}\, F(A'_n)}{\sqrt{\operatorname{Var} F(A'_n)}}, N\right) \leq \tilde{C} n^{-\frac{(d-1)}{2d}}, \quad n \geq 9. \tag{5.9}$$

Additionally we have

$$d_K\left(\frac{\operatorname{Vol}(A_s) - \operatorname{Vol}(A)}{\sqrt{\operatorname{Var}\operatorname{Vol}(A_s)}}, N\right) \leq \tilde{C} s^{-\frac{(d-1)}{2d}}, \quad s \geq 1, \tag{5.10}$$

and

$$d_K\left(\frac{\operatorname{Vol}(A'_n) - \operatorname{Vol}(A)}{\sqrt{\operatorname{Var}\operatorname{Vol}(A'_n)}}, N\right) \leq \tilde{C} n^{-\frac{(d-1)}{2d}}, \quad n \geq 9. \tag{5.11}$$

Additionally, if $F = \operatorname{Vol}$ *and* A *is compact and convex, then all of the above inequalities are in force.*

Remarks. (i) The bound (5.10) provides a rate of convergence for the main result of [44] (see Theorem 1.1 there), which establishes asymptotic normality for $\operatorname{Vol}(A_s)$, A convex. The bound (5.10) also improves upon Corollary 2.1 of [50] which shows

$$d_K\left(\frac{\operatorname{Vol}(A_s) - \mathbb{E}\operatorname{Vol}(A_s)}{\sqrt{\operatorname{Var}\operatorname{Vol}(A_s)}}, N\right) = O\left((\log s)^{3d+1} s^{\frac{-(d-1)}{2d}}\right).$$

Recall that the normal convergence of $\mathcal{H}^{d-1}(\partial A_s)$ is given in Remark (i) after Theorem 2.4 of [50] and the bound (5.8) for $F = \mathcal{H}^{d-1}(\partial \cdot)$ provides a rate for this normal convergence.
(ii) The bound (5.11) improves upon the bound of Theorem 6.1 of [22], which contains extra logarithmic factors, and, thus, addresses an open problem raised in Remark 6.9 of [22].
(iii) We may likewise deduce identical rates of normal convergence for other geometric statistics of A_s, including the total number of k-dimensional faces of A_s, $k \in \{0, 1, ..., d-1\}$, as well as the k-dimensional Hausdorff measure of the union of the k-dimensional faces of A_s (thus when $k = d - 1$, this gives $\mathcal{H}^{d-1}(\partial A_s)$). Second order asymptotics, including the requisite variance lower bounds for these statistics, are established in [48]. In the case of geometric statistics of A'_n, we expect similar variance lower bounds and central limit theorems.
(iv) Lower bounds for $\operatorname{Var} F(A_s)$ and $\operatorname{Var} F(A'_n)$ are essential to showing (5.8)-(5.11). We expect the order of these bounds to be unchanged if $\mathbb{Q}$ has a density bounded away from zero and infinity. We thus expect Theorem 5.5 to remain valid in this context because all other arguments in our proof hold for such $\mathbb{Q}$.

Proof. We first prove (5.8) for $F = \mathrm{Vol}$ and $F = \mathrm{Vol}(\cdot\Delta A)$. The proof method extends easily to the case when Poisson input is replaced by binomial input and we sketch the details as needed. To deduce (5.8) from Theorem 2.3(b), we need to (a) express $sF(A_s)$ as a sum of stabilizing score functions and (b) define $K \subset \mathbb{X}$ and show that the scores decay exponentially fast with respect to K.

(a) *Definition of scores.* As in [50], for $\mathcal{X} \in \mathbf{N}$, $x \in \mathcal{X}$, and a fixed subset A of $\mathbb{X}$, define the scores

$$\nu^{\pm}(x,\mathcal{X}) := \begin{cases} \mathrm{Vol}(C(x,\mathcal{X}) \cap A^c) & \text{if } x \in A \\ \pm\mathrm{Vol}(C(x,\mathcal{X}) \cap A) & \text{if } x \in A^c. \end{cases} \tag{5.12}$$

Define $\nu_s^{\pm}(x,\mathcal{X}) := s\nu^{\pm}(x,\mathcal{X})$. By the definition of $\nu^{\pm}$ at (5.12) we have

$$s\mathrm{Vol}(A_s) = \sum_{x\in\mathcal{P}_s} \nu_s^{-}(x,\mathcal{P}_s) + s\mathrm{Vol}(A) \quad \text{and} \quad s\mathrm{Vol}(A\Delta A_s) = \sum_{x\in\mathcal{P}_s} \nu_s^{+}(x,\mathcal{P}_s).$$

The arguments of Section 5.1 of [27] show that the scores $\nu_s^{\pm}$ have a radius of stabilization $R_s(x,\mathcal{P}_s \cup \{x\})$ with respect to $\mathcal{P}_s$ which satisfies (2.4) with $\gamma = d$ and $\alpha_{stab} = d$. The scores $\nu_s^{\pm}$ also satisfy the $(4+p)$-moment condition (2.7) for all $p \in [0,\infty)$.

As remarked in [50] and as shown in Lemma 5.1 of [27], the scores $\nu_n^{\pm}$ have a radius of stabilization $R_n(x,\mathcal{X}_{n-8} \cup \{x\})$ with respect to binomial input $\mathcal{X}_n$ which satisfies (2.5) with $\gamma = d$ and $\alpha_{stab} = d$.

(b) *Definition of* K. We set K to be ∂A. As noted in the proof of Theorem 2.1 of [50], we assert that the scores $\nu_n^{\pm}$ decay exponentially fast with their distance to ∂A, i.e. they satisfy (2.9) and (2.8) when K is set to ∂A and with $\alpha_K = d$. To see this for Poisson input, note that

$$\mathbb{P}(\nu_s(x,\mathcal{P}_s \cup \{x\}) \neq 0) \le \mathbb{P}(\mathrm{diam}(C(x,\mathcal{P}_s \cup \{x\})) \ge \mathrm{d}(x,K)).$$

Since $\mathrm{diam}(C(x,\mathcal{P}_s \cup \{x\})) \le 2R_s(x,\mathcal{P}_s \cup \{x\})$ and since $R_s(x,\mathcal{P}_s \cup \{x\})$ has exponentially decaying tails, the assertion follows.

We deduce (5.8) from the bound (2.18) of Theorem 2.3(b) as follows. If either ∂A contains a $(d-1)$-dimensional C^2 submanifold or A is compact and convex then $s^2 \operatorname{Var} \mathrm{Vol}(A_s) = \Theta(s^{(d-1)/d})$; see Theorem 1.2 of [44], Theorem 1.1 of [48] and Theorem 2.2 of [50]. All conditions of Theorem 2.3 are satisfied and so (5.8) follows for $F = \mathrm{Vol}$. Replacing $\mathrm{Vol}(A_s)$ with $\mathrm{Vol}(A\Delta A_s)$, the analog of (5.10) holds if ∂A contains a $(d-1)$-dimensional C^2 submanifold. This assertion follows since the stated conditions imply $s^2 \operatorname{Var} \mathrm{Vol}(A\Delta A_s) = \Theta(s^{(d-1)/d})$, as shown in Theorem 2.2 of [50]. See also [48]. We may similarly deduce (5.9) from the bound (2.18) of Theorem 2.3(b). If either ∂A contains a $(d-1)$-dimensional C^2 submanifold or A is compact and convex then $n^2 \operatorname{Var} \mathrm{Vol}(A_n') = \Theta(n^{(d-1)/d})$ as shown in Theorem 2.3 of [50]. Thus (5.9) follows for $F = \mathrm{Vol}$. Considering now $F = \mathrm{Vol}(\cdot\Delta A)$, and appealing to the variance lower bounds of

Theorem 2.3 of [50], we see that when ∂A contains a $(d-1)$-dimensional C^2 submanifold, all conditions of Theorem 2.3(b) are satisfied in the context of binomial input, and so the bound (5.9) follows for $F = \mathrm{Vol}(\cdot \Delta A)$.

To deduce (5.10) from (5.8), we need to replace $\mathbb{E}\,\mathrm{Vol}(A_s)$ with $\mathrm{Vol}(A)$. As shown in [21, Theorem 2], if the random input consists of n i.i.d. uniformly distributed random variables then $|\mathbb{E}\,\mathrm{Vol}(A'_n) - \mathrm{Vol}(A)| \leq c^n$ for some $c \in (0,1)$. A similar statement holds for Poisson input $\mathcal{P}_s$: If $|\mathcal{P}_s|$ is the cardinality of $\mathcal{P}_s$, then

$$|\mathbb{E}\,\mathrm{Vol}(A_s) - \mathrm{Vol}(A)| = \sum_{n\in\mathbb{N}} \mathbb{P}(|\mathcal{P}_s| = n)\,|\mathbb{E}\,\mathrm{Vol}(A'_n) - \mathrm{Vol}(A)| \leq \exp(s(c-1)).$$

Given either Poisson or binomial input, the variance and centered expectations thus have a subpolynomial decay. This exponential bias allows one to replace $\mathbb{E}\,\mathrm{Vol}(A_s)$ by $\mathrm{Vol}(A)$ in (5.8) and similarly for $\mathbb{E}\,\mathrm{Vol}(A'_n)$. This gives (5.10) and (5.11).

We now show (5.8) for $F = \mathcal{H}^{d-1}(\partial\cdot)$ and that it also holds when Poisson input is replaced by binomial input. Given $\mathcal{X} \in \mathbf{N}$ and a Borel subset $A \subset \mathbb{R}^d$, define for $x \in \mathcal{X} \cap A$ the score $\alpha(x, \mathcal{X})$ to be the $\mathcal{H}^{d-1}$ measure of the $(d-1)$-dimensional faces of $C(x, \mathcal{X})$ belonging to the boundary of $\bigcup_{w\in\mathcal{X}\cap A} C(w, \mathcal{X})$; if there are no such faces or if $x \notin \mathcal{X} \cap A$, then set $\alpha(x, \mathcal{X})$ to be zero.

Put $\alpha_s(x, \mathcal{X}) := s^{(d-1)/d}\alpha(x, \mathcal{X})$. Recalling the notation (1.1), the surface area of $s^{1/d}A_s$ is then given by

$$s^{(d-1)/d}\mathcal{H}^{d-1}(\partial A_s) = h_s(\mathcal{P}_s) = \sum_{x\in\mathcal{P}_s} \alpha_s(x, \mathcal{P}_s).$$

We want to deduce (5.8) and (5.9) for $F = \mathcal{H}^{d-1}$ from Theorem 2.3(b) with K set to ∂A. Note that $I_{K,s} = \Theta(s^{(d-1)/d})$.

As shown in the proof of Theorem 2.5 of [50], the scores α_s are exponentially stabilizing with respect to Poisson and binomial input. In other words they satisfy (2.4) and (2.5) with $\gamma = d$ and $\alpha_{stab} = d$. They also satisfy the $(4+p)$-moment conditions (2.6) and (2.7) for all $p \in [0, \infty)$. As noted in the proof of Theorem 2.5 of [50], the scores α_s decay exponentially fast with their distance to ∂A, i.e. they satisfy (2.8) and (2.9) when K is set to ∂A. We note that

$$\mathrm{Var}\,\mathcal{H}^{d-1}(\partial A_s) = \Theta(s^{-(d-1)/d}), \tag{5.13}$$

as shown in Theorem 1.1 of [48]. The conditions of Corollary 2.2(b) are satisfied and (5.8) for $F = \mathcal{H}^{d-1}$ follows from (2.15).

Recalling the notation (1.2), we also have

$$n^{(d-1)/d}\mathcal{H}^{d-1}(\partial A'_n) = h_n(\mathcal{X}_n) = \sum_{x\in\mathcal{X}_n} \alpha_n(x, \mathcal{X}_n).$$

We assert that $\operatorname{Var} \mathcal{H}^{d-1}(\partial A'_n) = \Theta(n^{-(d-1)/d})$. This may be proved by mimicking the methods to prove (5.13) or, alternatively, with $Z(n)$ denoting an independent Poisson random variable with mean n, we could use Lemma 6.1 of [50] to show $|\operatorname{Var} h_n(\mathcal{X}_{Z(n)}) - \operatorname{Var} h_n(\mathcal{X}_n)| = o(n^{(d-1)/d})$. This gives (5.9) for $F = \mathcal{H}^{d-1}$, as desired. □

5.4 Statistics of convex hulls of random point samples

In the following let A be a compact convex subset of $\mathbb{R}^d$ with non-empty interior, C^2-boundary and positive Gaussian curvature. By $\mathbb{Q}$ we denote the uniform measure on A. Let $\mathcal{P}_s$, $s \geq 1$, be a Poisson point process with intensity measure $s\mathbb{Q}$ and let $\mathcal{X}_n$, $n \in \mathbb{N}$, be a binomial point process of n independent points distributed according to $\mathbb{Q}$. From now on $\operatorname{Conv}(\mathcal{X})$ stands for the convex hull of a set $\mathcal{X} \subset \mathbb{R}^d$. The aim of this subsection is to establish rates of normal convergence for statistics of the random polytopes $\operatorname{Conv}(\mathcal{P}_s)$ and $\operatorname{Conv}(\mathcal{X}_n)$. We denote the number of k-faces of a polytope P by $f_k(P)$, $k \in \{0, \dots, d-1\}$, and its intrinsic volumes by $V_i(P)$, $i \in \{1, \dots, d\}$.

Theorem 5.5. *For any $h \in \{f_0, \dots, f_{d-1}, V_1, \dots, V_d\}$, there is a constant C_h also depending on A such that*

$$d_K\left(\frac{h(\operatorname{Conv}(\mathcal{P}_s)) - \mathbb{E}\, h(\operatorname{Conv}(\mathcal{P}_s))}{\sqrt{\operatorname{Var} h(\operatorname{Conv}(\mathcal{P}_s))}}, N\right) \leq C_h s^{-\frac{d-1}{2(d+1)}}, \quad s \geq 1, \tag{5.14}$$

and

$$d_K\left(\frac{h(\operatorname{Conv}(\mathcal{X}_n)) - \mathbb{E}\, h(\operatorname{Conv}(\mathcal{X}_n))}{\sqrt{\operatorname{Var} h(\operatorname{Conv}(\mathcal{X}_n))}}, N\right) \leq C_h n^{-\frac{d-1}{2(d+1)}}, \quad n \geq 9. \tag{5.15}$$

Remarks. (i) *Previous work.* The asymptotic study of the statistics $h(\operatorname{Conv}(\mathcal{P}_s))$ and $h(\operatorname{Conv}(\mathcal{X}_n))$, $h \in \{f_0, \dots, f_{d-1}, V_1, \dots, V_d\}$, has a long and rich history, starting with the seminal work of Rényi and Sulanke [37, 38]. Reitzner's breakthrough paper [35], which relies on dependency graph methods and Voronoi cells, establishes rates of normal convergence for Poisson input and $h \in \{f_0, \dots, f_{d-1}, V_d\}$ of the order $s^{-\frac{d-1}{2(d+1)}} \ln(s)^{2+\frac{2}{d+1}}$ (Theorems 1 and 2). Still in the setting $h \in \{f_0, \dots, f_{d-1}, V_d\}$, but with binomial input, for $d \geq 3$, Theorem 1.3 of Vu's paper [49] provides a rate of convergence $n^{-1/(d+1)+o(1)}$ in (5.15), which contains extraneous powers of n. For $d = 2$, both Reitzner (Poisson) and Vu (binomial) obtain the rate $n^{-1/6+o(1)}$, which is still slower that (5.15). When $h \in \{f_0, \dots, f_{d-1}, V_1, \dots, V_d\}$, Theorem 7.1 of [10] gives a central limit theorem for $h(\operatorname{Conv}(\mathcal{P}_s))$, with convergence rates involving extra logarithmic factors. We are unaware of central limit theorem results for intrinsic volume functionals over binomial input.
(ii) *Extensions.* Lower bounds for $\operatorname{Var} h(\operatorname{Conv}(\mathcal{X}_n))$ are essential to showing (5.14) and (5.15). We expect the order of these bounds to be unchanged if $\mathbb{Q}$ has a density bounded away from zero and infinity. Consequently we anticipate that Theorem 5.5 remains valid

in this context because all other arguments in our proof below also work for such a density.

In the following we may assume without loss of generality that $\mathbf{0}$ is in the interior of A. The proof of Theorem 5.5 is divided into several lemmas and we prepare it by recalling some geometric facts and introducing some notation.

For a boundary point $z \in \partial A$ we denote by T_z the tangent space parametrized by $\mathbb{R}^{d-1}$ in such a way that z is the origin. The boundary of A in a neighborhood of z may be identified with the graph of a function $f_z : T_z \to \mathbb{R}$. It may be deduced from [35, Section 5] that there are constants $\underline{c} \in (0,1)$, $\overline{c} \in (1,\infty)$ and $r_0 \in (0,\infty)$ such that uniformly for all $z \in \partial A$,

$$\underline{c}^2 \|v\|^2 \le f_z(v) \le \overline{c}^2 \|v\|^2, \quad v \in T_z \cap B^{d-1}(\mathbf{0}, r_0). \tag{5.16}$$

For $u > 0$ we define

$$A_{-u} := \{y \in A : \mathrm{d}(y, A^c) \le u\},$$

where $A^c := \mathbb{R}^d \setminus A$. It follows from (5.16) that there is a $\varrho > 0$ such that all points $x \in A_{-3\varrho}$ have a unique projection $\Pi_{\partial A}(x)$ to ∂A. For $3\varrho \ge \overline{r} \ge r \ge \underline{r} \geqslant 0$ it also holds that

$$\partial A_{-\overline{r}} \subset (\partial A_{-r} \oplus (\overline{r} - r) B^d(0,1)) \quad \text{and} \quad \partial A_{-\underline{r}} \subset (\partial A_{-r} \oplus (r - \underline{r}) B^d(0,1)). \tag{5.17}$$

We denote by d_{max} the metric

$$\mathrm{d}_{max}(x,y) := \max\{\|x-y\|, \sqrt{|\,\mathrm{d}(x,A^c) - \mathrm{d}(y,A^c)|}\}, \quad x, y \in A,$$

and define for $x \in A_{-2\varrho}$ and $r > 0$

$$B_{\mathrm{d}_{max}}(x,r) := \{y \in A_{-\varrho} : \mathrm{d}_{max}(x,y) \le r\}.$$

The following lemma ensures that the space $(A, \mathcal{B}(A), \mathbb{Q})$ and the metric d_{max} satisfy condition (2.1) for $x \in A_{-2\varrho}$, with $\gamma = d + 1$.

Lemma 5.6. *There is a constant $\kappa > 0$ such that for all $x \in A_{-2\varrho}$ and $r > 0$*

$$\limsup_{\varepsilon \to \infty} \frac{\mathbb{Q}(B_{\mathrm{d}_{max}}(x, r+\varepsilon)) - \mathbb{Q}(B_{\mathrm{d}_{max}}(x,r))}{\varepsilon} \le \kappa (d+1) r^d.$$

Proof. We define for $x \in A_{-2\varrho}$ and $r > 0$

$$U_{x,r} := B_{\mathrm{d}_{max}}(x,r) \cap \{y \in A_{-\varrho} : \|x - y\| = r\}$$

and

$$V_{x,r} := B_{\mathrm{d}_{max}}(x,r) \cap \{y \in A_{-\varrho} : |\,\mathrm{d}(x,A^c) - \mathrm{d}(y,A^c)| = r^2\}.$$

It follows from (5.17) that

$$\begin{aligned}
&\limsup_{\varepsilon\to\infty} \frac{\mathbb{Q}(B_{\mathrm{d}_{max}}(x,r+\varepsilon)) - \mathbb{Q}(B_{\mathrm{d}_{max}}(x,r))}{\varepsilon} \\
&\le \limsup_{\varepsilon\to\infty} \frac{\mathbb{Q}(U_{x,r} \oplus \varepsilon B^d(0,1)) + \mathbb{Q}(V_{x,r} \oplus (2r\varepsilon+\varepsilon^2)B^d(0,1))}{\varepsilon} \\
&\le 2\mathcal{H}^{d-1}(U_{x,r}) + 4r\mathcal{H}^{d-1}(V_{x,r}).
\end{aligned}$$

For r sufficiently small, we obtain sub- and supersets for $A_{-(\mathrm{d}(x,A^c)-r^2)} \cap B^d(x,r)$ and $A_{-(\mathrm{d}(x,A^c)+r^2)} \cap B^d(x,r)$ by taking the inner parallel sets with respect to the paraboloids given in (5.16). Consequently, $U_{x,r}$ is contained in a strip whose Euclidean thickness is of the order r^2. This implies that $\mathcal{H}^{d-1}(U_{x,r}) \le c_K r^d$ for all $r > 0$ with some constant $c_A \in (0,\infty)$ only depending on A.

Since $V_{x,r}$ is the union of the intersection of the boundaries of the convex sets $A_{-(\mathrm{d}(x,A^c)+r^2)}$ and $A_{-(\mathrm{d}(x,A^c)-r^2)}$ with $B^d(x,r)$ we have that $\mathcal{H}^{d-1}(V_{x,r}) \le 2d\kappa_d r^{d-1}$, which completes the proof. □

We let $u_x := (\Pi_{\partial A}(x) - x)/\|\Pi_{\partial A}(x) - x\|$ be the unit normal at $x \in A_{-3\varrho}$ pointing in the direction of the boundary. For $x \in A$ and $r > 0$ we define the hyperplanes

$$H_x := \{y \in A : \langle \Pi_{\partial A}(x) - x, y\rangle = \langle \Pi_{\partial A}(x) - x, x\rangle\}$$

and the parametrized family of sets

$$A_{x,r} := \begin{cases} \mathrm{Conv}((H_x \cap B^d(x,r/\overline{c})) \cup \{x + r^2 u_x\}), & r \le \sqrt{\mathrm{d}(x,A^c)}, \\ A \setminus \mathrm{Conv}((A \setminus B^d(x,r/\underline{c})) \cup \{x\}), & r > \sqrt{\mathrm{d}(x,A^c)}. \end{cases}$$

The sets $A_{x,r}$ have the following important properties. When $r > \sqrt{\mathrm{d}(x,A^c)}$ we note that x is an extreme point of $A_{x,r}$.

Lemma 5.7. *a) There is a constant $c_{\mathbb{Q}} \in (0,\infty)$ such that*

$$\mathbb{Q}(A_{x,r}) \ge c_{\mathbb{Q}} r^{d+1}, \quad x \in A_{-\varrho}, r \in [0,1].$$

b) There is a constant $c_{max} \in (0,\infty)$ such that $A_{x,r} \subset B_{\mathrm{d}_{max}}(x, c_{max} r)$ for any $r > 0$ and $x \in A_{-\tilde{\varrho}}$ with $\tilde{\varrho} := \min\{1/(8\overline{c}^2), \varrho\}$.

Proof. We denote the epigraphs of $v \mapsto \underline{c}^2\|v\|^2$ and $v \mapsto \overline{c}^2\|v\|^2$ by $\underline{P}_z$ and $\overline{P}_z$. For $r \le \sqrt{\mathrm{d}(x,A^c)}$ we have $\mathbb{Q}(A_{x,r}) = \kappa_{d-1} r^{d+1}/(d\overline{c}^{d-1})$. For $x \in A$ and $r > \sqrt{\mathrm{d}(x,A^c)}$ let $z := \Pi_{\partial A}(x)$. Since

$$\mathrm{Conv}((A \setminus B^d(z,r/\underline{c})) \cup \{x\}) \subseteq \mathrm{Conv}((A \setminus B^d(z,r/\underline{c})) \cup \{z\}),$$

it follows that $A_{x,r} \supset A_{z,r}$. Additionally

$$A_{z,r} \supset \overline{P}_z \setminus \operatorname{Conv}(\{z\} \cup (\underline{P}_z \setminus B^d(z, r/\underline{c}))).$$

A longer computation shows that the volume of the set on the right-hand side can be bounded below by a non-negative scalar multiple of r^{d+1}, which gives part a).

To prove part b) it suffices to consider only the situation $r \in [0,1]$. It follows immediately from the definition of $A_{x,r}$ that $A_{x,r} \subset B^d(x, r/\underline{c})$ for $r \in [0,1]$. For $x \in A$ with $\mathrm{d}(x, A) \leq 1/(8\overline{c}^2)$, $r \leq \sqrt{\mathrm{d}(x, A^c)}$ and $y \in A_{x,r}$, we obtain by a direct but longer computation that

$$\mathrm{d}(x, A^c) \geq \mathrm{d}(y, A^c) \geq \mathrm{d}(x, A^c) - 4\overline{c}^2 r^2.$$

On the other hand, for $r > \sqrt{\mathrm{d}(x, A^c)}$ and $y \in A_{x,r}$, we have with $z = \Pi_{\partial A}(x)$ that

$$\begin{aligned} \mathrm{d}(y, A^c) &\leq \sup_{z \in \operatorname{Conv}(\partial A \cap B^d(x, r/\underline{c}))} \mathrm{d}(z, A^c) \leq \sup_{z \in \operatorname{Conv}(\overline{P}_z \cap B^d(z, r/\underline{c} + \mathrm{d}(x, A^c)))} \mathrm{d}(z, A^c) \\ &\leq \sup_{z \in \operatorname{Conv}(\overline{P}_z \cap B^d(z, (1/\underline{c}+1)r))} \mathrm{d}(z, A^c) \leq \overline{c}^2 (1/\underline{c} + 1)^2 r^2. \end{aligned}$$

This implies that $A_{x,r} \subset \{y \in B^d(0,1) : \sqrt{|\,\mathrm{d}(x, A^c) - \mathrm{d}(y, A^c)|} \leq \overline{c}(1/\underline{c} + 2)r\}$, which completes the proof of part b). □

For $k \in \{0, \dots, d-1\}$ and $\mathcal{X} \in \mathbf{N}$ let $\mathcal{F}_k(\operatorname{Conv}(\mathcal{X}))$ be the set of k-dimensional faces of $\operatorname{Conv}(\mathcal{X})$. To cast $f_k(\operatorname{Conv}(\mathcal{X}))$, $k \in \{1, \dots, d-1\}$, in the form of (1.1) and (1.2), we define

$$\xi_k(x, \mathcal{X}) := \frac{1}{k+1} \sum_{F \in \mathcal{F}_k(\operatorname{Conv}(\mathcal{X}))} \mathbf{1}\{x \in F\}, \quad x \in \mathcal{X}.$$

Note that $f_k(\operatorname{Conv}(\mathcal{X})) = \sum_{x \in \mathcal{X}} \xi_k(x, \mathcal{X})$.

To cast the intrinsic volumes $V_j(\operatorname{Conv}(\mathcal{X}))$, $j \in \{1, \dots, d-1\}$, in the form of (1.1) and (1.2), we need some more notation. Given the convex set A and a linear subspace E, denote by $A|E$ the orthogonal projection of A onto E. For $x \in \mathbb{R}^d \setminus \{0\}$, let $L(x)$ the line spanned by x. Given a line $N \subset \mathbb{R}^d$ through the origin, and for $1 \leq j \leq d$, let $G(N, j)$ be the set of j-dimensional linear subspaces of $\mathbb{R}^d$ containing N. Let then $\nu_j^N(\cdot)$ be the Haar probability measure on $G(N, j)$. Let $M \subset A$ be convex. For $j \in \{0, \dots, d-1\}$, $x \in \mathbb{R}^d \setminus \{\mathbf{0}\}$, and $L \in G(L(x), j)$ define

$$f^L(x) := \mathbf{1}_{\{x \in (A|L) \setminus (M|L)\}}$$

and, as in [10], define the *projection avoidance function* $\theta_j^{A,M} : \mathbb{R}^d \setminus \{\mathbf{0}\} \mapsto [0,1]$ by

$$\theta_j^{A,M}(x) := \int_{G(L(x),j)} f^L(x)\, \nu_j^{L(x)}(\mathrm{d}L).$$

The following result generalizes [10, (2.7)] to non-spherical compact sets, with arguments similar to Lemma A1 from [17]. The proof is in the appendix.

Lemma 5.8. *Let $M \subset A$ be convex subsets of $\mathbb{R}^d$. For all $j \in \{0, \dots, d-1\}$ there is a constant $\kappa_{d,j}$ depending on d, j such that*

$$V_j(A) - V_j(M) = \kappa_{d,j} \int_{A \setminus M} \theta_j^{A,M}(x) \|x\|^{-(d-j)} \, \mathrm{d}x \leq \kappa_{d,j} r(M)^{-(d-j)} (V_d(A) - V_d(M)), \tag{5.18}$$

where $r(M)$ is the radius of the largest ball centered at $\mathbf{0}$ and contained in M.

For $\mathcal{X} \in \mathbf{N}$ and $F \in \mathcal{F}_{d-1}(\mathrm{Conv}(\mathcal{X}))$ put $\mathrm{cone}(F) := \{ry : \ y \in F, r > 0\}$. Define for $j \in \{1, ..., d-1\}$

$$\xi_{j,s}(x, \mathcal{X}) = \frac{s\kappa_{d,j}}{d} \sum_{F \in \mathcal{F}_{d-1}(\mathrm{Conv}(\mathcal{X}))} \mathbf{1}_{\{x \in F\}} \int_{\mathrm{Cone}(F) \cap (A \setminus \mathrm{Conv}(\mathcal{X}))} \|x\|^{-(d-j)} \theta_j^{A,\mathrm{Conv}(\mathcal{X})}(x) \, \mathrm{d}x$$

for $x \in \mathcal{X}$, $s \geq 1$. Lemma 5.8 yields

$$s(V_j(A) - V_j(\mathrm{Conv}(\mathcal{X}))) = \sum_{x \in \mathcal{X}} \xi_{j,s}(x, \mathcal{X}) \tag{5.19}$$

if $\mathbf{0}$ is in the interior of $\mathrm{Conv}(\mathcal{X})$ and if all points of $\mathcal{X}$ are in general position. For $x \in \mathcal{X}$ and $s \geq 1$ define

$$\xi_{d,s}(x, \mathcal{X}) := \frac{s}{d} \sum_{F \in \mathcal{F}_{d-1}(\mathrm{Conv}(\mathcal{X}))} \mathbf{1}_{\{x \in F\}} \int_{\mathrm{Cone}(F) \cap (A \setminus \mathrm{Conv}(\mathcal{X}))} \mathrm{d}x.$$

Here we include the multiplicative factor s, since when $\mathcal{X}$ is replaced by $\mathcal{P}_s$ we get that $\xi_{d,s}(x, \mathcal{X})$ is of constant order. Without multiplying by s, the defect volume at a vertex is of order $s^{-1} = s^{-2/(d+1)} s^{-(d-1)/(d+1)}$, i.e., roughly the product of its width and its $(d-1)$-dimensional surface area. The same comment applies to $\xi_{j,s}$.

If $\mathbf{0}$ is in the interior of $\mathrm{Conv}(\mathcal{X})$ and all points of $\mathcal{X}$ are in general position, we have as well

$$sV_d(A \setminus \mathrm{Conv}(\mathcal{X})) = \sum_{x \in \mathcal{X}} \xi_{d,s}(x, \mathcal{X}).$$

The definitions of the scores and (5.18) show that for $\mathcal{X} \in \mathbf{N}$, $x \in \mathcal{X}$, $s \geq 1$ and $j \in \{0, ..., d-1\}$

$$\xi_{j,s}(x, \mathcal{X}) \leq r(\mathrm{Conv}(\mathcal{X}))^{-(d-j)} \xi_{d,s}(x, \mathcal{X}). \tag{5.20}$$

Since $\mathbf{0} \in \mathrm{int}(A)$, we can choose $\rho_0 \in (0, \tilde{\varrho})$ such that $B(\mathbf{0}, 2\rho_0) \subset A$. For a score ξ we denote by $\tilde{\xi}$ the modified score

$$\tilde{\xi}(x, \mathcal{X}) := \mathbf{1}\{x \in A_{-\rho_0}\} \xi(x, (\mathcal{X} \cap A_{-\varrho_0}) \cup \{\mathbf{0}\})$$

for $\mathcal{X} \in \mathbf{N}$ and $x \in \mathcal{X}$. Our strategy of proof for Theorem 5.5 is to apply in a first step Corollary 2.2 to these modified scores, putting $\mathbb{X} := A_{-\varrho_0}$ and K set to ∂A. Thereafter we show that the result remains true without truncating and without adding the origin as an additional point.

For a score ξ and $\mathcal{X} \in \mathbf{N}$ we define

$$S_\xi(\mathcal{X}) := \sum_{x \in \mathcal{X}} \xi(x, \mathcal{X}).$$

Lemma 5.9. *For any $\xi_s \in \{\xi_0, \ldots, \xi_{d-1}, \xi_{1,s}, \ldots, \xi_{d,s}\}$ there are constants $C_0, c_0 \in (0, \infty)$ such that*

$$\begin{aligned}&\max\{\mathbb{P}(S_{\xi_s}(\mathcal{P}_s) \neq S_{\tilde{\xi}_s}(\mathcal{P}_s)), \mathbb{P}(B^d(\mathbf{0}, \rho_0) \not\subset \operatorname{Conv}(\mathcal{P}_s)),\\ &\qquad |\mathbb{E}\, S_{\xi_s}(\mathcal{P}_s) - \mathbb{E}\, S_{\tilde{\xi}_s}(\mathcal{P}_s)|, |\operatorname{Var} S_{\xi_s}(\mathcal{P}_s) - \operatorname{Var} S_{\tilde{\xi}_s}(\mathcal{P}_s)|\}\\ &\leq C_0 \exp(-c_0 s)\end{aligned}$$

for $s \geq 1$ and

$$\begin{aligned}&\max\{\mathbb{P}(S_{\xi_n}(\mathcal{X}_n) \neq S_{\tilde{\xi}_n}(\mathcal{X}_n)), \mathbb{P}(B^d(\mathbf{0}, \rho_0) \not\subset \operatorname{Conv}(\mathcal{X}_n)),\\ &\qquad |\mathbb{E}\, S_{\xi_n}(\mathcal{X}_n) - \mathbb{E}\, S_{\tilde{\xi}_n}(\mathcal{X}_n)|, |\operatorname{Var} S_{\xi_n}(\mathcal{X}_n) - \operatorname{Var} S_{\tilde{\xi}_n}(\mathcal{X}_n)|\}\\ &\leq C_0 \exp(-c_0 n)\end{aligned}$$

for $n \geq 1$.

Proof. One can choose sets $A_1, \ldots, A_m \subset \{x \in A : \mathrm{d}_{max}(x, A^c) \leq \rho_0\}$ with non-empty interior such that, for any $\mathcal{X} \in \mathbf{N}$ with $A_i \cap \mathcal{X} \neq \emptyset$, $i \in \{1, \ldots, m\}$,

$$\operatorname{conv}(\mathcal{X}) \supset \{x \in A : \mathrm{d}_{max}(x, A^c) > \rho_0\}.$$

Using $B(\mathbf{0}, 2\rho_0) \subset A$, this inclusion yields $B(\mathbf{0}, \rho_0) \subset \operatorname{Conv}(\mathcal{X})$. The event $S_{\tilde{\xi}_s}(\mathcal{X}) \neq S_{\xi_s}(\mathcal{X})$ is also a subset of the event $A_i \cap \mathcal{X} = \emptyset$ for some $i \in \{1, \ldots, m\}$. These observations yield the probability bounds.

Combining these observations with the generous upper bounds $\max_{k \in \{0,\ldots,d-1\}} f_k(\mathcal{X}) \leq |\mathcal{X}|^{d-1}$, $|S_{\tilde{\xi}_s}(\mathcal{X}) - S_{\xi_s}(\mathcal{X})| \leq C_d s |\mathcal{X}|^d$ for some universal constant $C_d \in (0, \infty)$, and Hölder's inequality, we obtain the asserted expectation and variance bounds. □

The results of [35] show that for $\xi_s \in \{\xi_0, \ldots, \xi_{d-1}, \xi_{d,s}\}$ one has

$$\operatorname{Var} S_{\xi_s}(\mathcal{P}_s) = \Theta(s^{\frac{d-1}{d+1}}) \quad \text{and} \quad \operatorname{Var} S_{\xi_n}(\mathcal{X}_n) = \Theta(n^{\frac{d-1}{d+1}}). \tag{5.21}$$

For $\xi_s \in \{\xi_{1,s}, \ldots, \xi_{d-1,s}\}$ and taking into account scaling (5.19), we know from Corollary 7.1 of [10] and from Theorems 1 and 2 of [3] that

$$\operatorname{Var} S_{\xi_s}(\mathcal{P}_s) = \Theta(s^{\frac{d-1}{d+1}}) \quad \text{and} \quad \operatorname{Var} S_{\xi_n}(\mathcal{X}_n) = \Theta(n^{\frac{d-1}{d+1}}). \tag{5.22}$$

Hence, Lemma 5.9 implies that for $\xi_s \in \{\xi_0, \dots, \xi_{d-1}, \xi_{1,s}, \dots, \xi_{d,s}\}$

$$\operatorname{Var} S_{\tilde{\xi}_s}(\mathcal{P}_s) = \Theta(s^{\frac{d-1}{d+1}}) \quad \text{and} \quad \operatorname{Var} S_{\tilde{\xi}_n}(\mathcal{X}_n) = \Theta(n^{\frac{d-1}{d+1}}). \tag{5.23}$$

For a point $x \in A$ let $\tilde{H}_{x,1}, \dots, \tilde{H}_{x,2^{d-1}}$ be a decomposition of H_x into solid orthants with x as joint point and let $H_{x,i} := \tilde{H}_{x,i} + \operatorname{Span}(x)$ for $i \in \{1, \dots, 2^{d-1}\}$.

Lemma 5.10. *Let $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ and let $x \in A$, $r > 0$ and $\mathcal{X} \in \mathbf{N}$ be such that $x \in \mathcal{X}$ and $\mathcal{X} \cap A_{x,r} \cap H_{x,i} \neq \emptyset$ for $i \in \{1, \dots, 2^{d-1}\}$. Then, $\tilde{\xi}_s(x, \mathcal{X})$ is completely determined by $\mathcal{X} \cap A_{x,r}$, i.e., thus by $\mathcal{X} \cap B_{\mathrm{d}_{max}}(x, c_{max} r)$ with c_{max} as in Lemma 5.7.*

Proof. Let $\mathrm{d}(x, A^c) \leq \varrho_0 \leq 1/(8\overline{c}^2)$ since, otherwise, the assertion is trivial. By assumption there are $y_1, \dots, y_{2^{d-1}}$ such that $y_i \in \mathcal{X} \cap A_{x,r} \cap H_{x,i}$ for $i \in \{1, \dots, 2^{d-1}\}$. Let $C_{x,y_1,\dots,y_{2^{d-1}}}$ be the cone with apex x generated by the points $\mathbf{0}, y_1, \dots, y_{2^{d-1}}$. If $C_{x,y_1,\dots,y_{2^{d-1}}} = \mathbb{R}^d$, we have $x \in \operatorname{Conv}(\{\mathbf{0}, y_1, \dots, y_{2^{d-1}}\})$, whence $\tilde{\xi}_s(x, \mathcal{X}) = 0$. If $C_{x,y_1,\dots,y_{2^{d-1}}} \neq \mathbb{R}^d$ (this implies that $r > \sqrt{\mathrm{d}(x, A^c)}$), no point in the interior of $C_{x,y_1,\dots,y_{2^{d-1}}}$ can be connected with x by an edge. Since $\operatorname{Conv}((A \setminus B^d(x, r/\underline{c})) \cup \{x\}) \subset C_{x,y_1,\dots,y_{2^{d-1}}}$, all points in $A \setminus A_{x,r}$ are irrelevant for the facial structure at x. Consequently the scores $\tilde{\xi}_s$ are completely determined by $\mathcal{X} \cap A_{x,r}$. In view of Lemma 5.7 b) we have $A_{x,r} \subset B_{\mathrm{d}_{max}}(x, c_{max} r)$ so the same is true for $\mathcal{X} \cap B_{\mathrm{d}_{max}}(x, c_{max} r)$. □

We define the map $R : A \times \mathbf{N} \to \mathbb{R}$ which sends $(x, \mathcal{X})$ to

$$R(x, \mathcal{X} \cup \{x\}) := \begin{cases} c_{max} \inf\{r \geq 0 : \mathcal{X} \cap A_{x,r} \cap H_{x,i} \neq \emptyset \text{ for } i \in \{1, \dots, 2^{d-1}\}\}, & x \in A_{-\varrho_0}, \\ 0, & x \notin A_{-\varrho_0}. \end{cases}$$

The next lemma shows that all $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ satisfy (2.4) and (2.5) with $\alpha_{stab} = d + 1$.

Lemma 5.11. *R is a radius of stabilization for any $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ and there are constants $C, c \in (0, \infty)$ such that for $r \geq 0$, $x \in A$*

$$\mathbb{P}(R(x, \mathcal{P}_s \cup \{x\}) \geq r) \leq C \exp(-csr^{d+1}), \quad s \geq 1,$$

whereas

$$\mathbb{P}(R(x, \mathcal{X}_{n-8} \cup \{x\}) \geq r) \leq C \exp(-cnr^{d+1}), \quad n \geq 9.$$

Proof. It follows from Lemma 5.10 that R is a radius of stabilization. Using Lemma 5.7 a), we see that

$$\begin{aligned} \mathbb{P}(R(x, \mathcal{P}_s \cup \{x\}) \geq r) &\leq \mathbb{P}(\exists i \in \{1, \dots, 2^{d-1}\} : \mathcal{P}_s \cap A_{x,r/c_{max}} \cap H_{x,i} = \emptyset) \\ &\leq 2^{d-1} \exp(-s c_{\mathbb{Q}} r^{d+1}/c_{max}^{d+1}). \end{aligned}$$

The proof for the binomial case goes similarly. □

The next lemma shows that all $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ satisfy (2.8) and (2.9) with $\alpha_{A^c} = (d+1)/2$.

Lemma 5.12. *For any $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ there are constants $C_b, c_b \in (0,\infty)$ such that for $x \in A$, $\mathcal{A} \subset A$ with $|\mathcal{A}| \leq 7$*

$$\mathbb{P}(\tilde{\xi}_s(x, \mathcal{P}_s \cup \{x\} \cup \mathcal{A}) \neq 0) \leq C_b \exp(-c_b s \, \mathrm{d}_{max}(x, A^c)^{d+1}), \quad s \geq 1,$$

whereas

$$\mathbb{P}(\tilde{\xi}_n(x, \mathcal{X}_{n-8} \cup \{x\} \cup A) \neq 0) \leq C_b \exp(-c_b n \, \mathrm{d}_{max}(x, A^c)^{d+1}), \quad n \geq 9.$$

Proof. For $x \in A$, $\mathcal{X} \in \mathbf{N}$ and $\mathcal{A} \subset A$ with $|\mathcal{A}| \leq 7$ we have that $\tilde{\xi}_s(x, \mathcal{X} \cup \{x\} \cup \mathcal{A}) = 0$ if $R(x, \mathcal{X} \cup \{x\}) \leq \sqrt{\mathrm{d}(x, A^c)}$. Thus, the assertions follow from Lemma 5.11. □

Lemma 5.13. *For any $q \geq 1$ and $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ there is a constant $C_q \in (0,\infty)$ such that for all $\mathcal{A} \subset A$ with $|\mathcal{A}| \leq 7$,*

$$\sup_{s \geq 1} \sup_{x \in A} \mathbb{E}\, |\tilde{\xi}_s(x, \mathcal{P}_s \cup \{x\} \cup \mathcal{A})|^q \leq C_q \quad \text{and} \quad \sup_{n \in \mathbb{N}, n \geq 9} \sup_{x \in A} \mathbb{E}\, |\tilde{\xi}_n(x, \mathcal{X}_{n-8} \cup \{x\} \cup \mathcal{A})|^q \leq C_q.$$

Proof. The assertion for $\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}$ can be shown similarly as in Lemma 7.1 of [10]. It follows from Lemma 5.11 that the product of $s^{\frac{1}{d+1}}$ and the length of the longest edge emanating from x in any of the $(d-1)$ spatial directions has exponential tails. It also follows from Lemma 5.11 that the product of $s^{\frac{2}{d+1}}$ and the width of the defect volume in the radial direction has exponential tails. These observations prove the assertion for $\tilde{\xi}_{d,s}$. For the intrinsic volumes $\tilde{\xi}_{j,s}, j \in \{0, ..., d-1\}$, the bound (5.20) shows that the q-th moment of $\tilde{\xi}_{j,s}$ is bounded by a constant multiple of the q-th moment of $\tilde{\xi}_{d,s}$ plus $s^q \mathbb{P}(B(\mathbf{0}, \rho_0) \not\subset \mathrm{Conv}(\mathcal{X}_s))$, which by Lemma 5.9 is bounded by $s^q C_0 \exp(-c_0 s)$. This completes the proof. □

Lemma 5.14. *For any $\tilde{\xi}_s \in \{\tilde{\xi}_0, \dots, \tilde{\xi}_{d-1}, \tilde{\xi}_{1,s}, \dots, \tilde{\xi}_{d,s}\}$ there is a constant $\tilde{C} \in (0,\infty)$ such that*

$$d_K\left(\frac{S_{\tilde{\xi}_s}(\mathcal{P}_s) - \mathbb{E}\, S_{\tilde{\xi}_s}(\mathcal{P}_s)}{\sqrt{\mathrm{Var}\, S_{\tilde{\xi}_s}(\mathcal{P}_s)}}, N \right) \leq \tilde{C} s^{-\frac{d-1}{2(d+1)}}, \quad s \geq 1,$$

and

$$d_K\left(\frac{S_{\tilde{\xi}_n}(\mathcal{X}_n) - \mathbb{E}\, S_{\tilde{\xi}_n}(\mathcal{X}_n)}{\sqrt{\mathrm{Var}\, S_{\tilde{\xi}_n}(\mathcal{X}_n)}}, N \right) \leq \tilde{C} n^{-\frac{d-1}{2(d+1)}}, \quad n \geq 9.$$

Proof. By Lemmas 5.6, 5.11, 5.12, and 5.13 all conditions of Corollary 2.2 are satisfied for $\tilde{\mathbb{X}} := A_{-\varrho_0}$ and $K := \partial A$. Note that $I_{\partial A, s} = O(s^{(d-1)/(d+1)})$. These observations and Remark (v) after Theorem 2.3 complete the proof. □

Proof of Theorem 5.5. For any pair $(X, \tilde{X})$ of square integrable random variables satisfying $\operatorname{Var} X, \operatorname{Var} \tilde{X} > 0$, a straightforward computation shows that

$$
\begin{aligned}
& d_K\Big(\frac{X - \mathbb{E}\, X}{\sqrt{\operatorname{Var} X}}, N\Big) \\
& \leq d_K\Big(\frac{\tilde{X} - \mathbb{E}\, X}{\sqrt{\operatorname{Var} X}}, N\Big) + \mathbb{P}(X \neq \tilde{X}) \\
& \leq d_K\Big(\frac{\tilde{X} - \mathbb{E}\, \tilde{X}}{\sqrt{\operatorname{Var} \tilde{X}}}, N(\frac{\mathbb{E}\, X - \mathbb{E}\, \tilde{X}}{\sqrt{\operatorname{Var} X}}, \frac{\operatorname{Var} X}{\operatorname{Var} \tilde{X}})\Big) + \mathbb{P}(X \neq \tilde{X}) \\
& \leq d_K\Big(\frac{\tilde{X} - \mathbb{E}\, \tilde{X}}{\sqrt{\operatorname{Var} \tilde{X}}}, N\Big) + d_K\Big(N, N(\frac{\mathbb{E}\, X - \mathbb{E}\, \tilde{X}}{\sqrt{\operatorname{Var} X}}, \frac{\operatorname{Var} X}{\operatorname{Var} \tilde{X}})\Big) + \mathbb{P}(X \neq \tilde{X}) \\
& \leq d_K\Big(\frac{\tilde{X} - \mathbb{E}\, \tilde{X}}{\sqrt{\operatorname{Var} \tilde{X}}}, N\Big) + \frac{|\mathbb{E}\, X - \mathbb{E}\, \tilde{X}|}{\sqrt{\operatorname{Var} X}} + C\Big|\frac{\operatorname{Var} X}{\operatorname{Var} \tilde{X}} - 1\Big| + \mathbb{P}(X \neq \tilde{X}).
\end{aligned}
$$

Applying this to the pairs $(X, \tilde{X}) := (S_{\xi_s}(\mathcal{P}_s), S_{\tilde{\xi}_s}(\mathcal{P}_s))$ and $(X, \tilde{X}) := (S_{\xi_n}(\mathcal{X}_n), S_{\tilde{\xi}_n}(\mathcal{X}_n))$, respectively, together with Lemma 5.9, Lemma 5.14, (5.21), (5.22), and (5.23) completes the proof. $\square$

5.5 Clique counts in generalized random geometric graphs

Let $(\mathbb{X}, \mathcal{F}, \mathbb{Q})$ be equipped with a semi-metric d such that (2.1) is satisfied for some γ and κ. Moreover, let $\mathbb{M} = [0, \infty)$ be equipped with the Borel sigma algebra $\mathcal{F}_{\mathbb{M}} := \mathcal{B}([0, \infty))$ and a probability measure $\mathbb{Q}_{\mathbb{M}}$ on $([0, \infty), \mathcal{B}([0, \infty)))$. By $\widehat{\mathbb{Q}}$ we denote the the product measure of $\mathbb{Q}$ and $\mathbb{Q}_{\mathbb{M}}$. In the following let $\mathcal{P}_s$ be a marked Poisson point process with intensity measure $s\widehat{\mathbb{Q}}$, $s \geq 1$, and let $\mathcal{X}_n$ be a marked binomial point process of $n \in \mathbb{N}$ points distributed according to $\widehat{\mathbb{Q}}$.

Given $\mathcal{X} \in \mathbf{N}$, recall that $\mathbf{N}$ is the set of point configurations in $\widehat{\mathbb{X}}$, and a scale parameter $\beta \in (0, \infty)$, consider the graph $G(\mathcal{X}, \beta)$ on $\mathcal{X}$ with $(x_1, m_{x_1}) \in \mathcal{X}$ and $(x_2, m_{x_2}) \in \mathcal{X}$ joined with an edge iff $\mathrm{d}(x_1, x_2) \leq \beta \min(m_{x_1}, m_{x_2})$. When $m_x = 1$ for all $x \in \mathcal{X}$, we obtain the familiar geometric graph with parameter β. Alternatively, we could use the connection rule that (x_1, m_{x_1}) and (x_2, m_{x_2}) are joined with an edge iff $\mathrm{d}(x_1, x_2) \leq \beta \max(m_{x_1}, m_{x_2})$. A scale-free random graph based on this connection rule with an underlying marked Poisson point process is studied in [18]. The number of cliques of order $k+1$ in $G(\mathcal{X}, \beta)$, here denoted $\mathcal{C}_k(\mathcal{X}, \beta)$, is a well-studied statistic in random geometric graphs. Recall that $k+1$ vertices of a graph form a clique of order $k+1$ if each pair of them is connected by an edge.

The clique count $\mathcal{C}_k(\mathcal{X}, \beta)$ is also a central statistic in topological data analysis. Consider the simplicial complex $\mathcal{R}^{\beta}(\mathcal{X})$ whose k-simplices correspond to unordered $(k+1)$-tuples of points of $\mathcal{X}$ such that any constituent pair of points (x_1, m_{x_1}) and (x_2, m_{x_2})

satisfies $\mathrm{d}(x_1, x_2) \leq \beta \min(m_{x_1}, m_{x_2})$. When $m_x = 1$ for all $x \in \mathcal{X}$ then $\mathcal{R}^\beta(\mathcal{X})$ coincides with the Vietoris-Rips complex with scale parameter β and $\mathcal{C}_k(\mathcal{X}, \beta)$ counts the number of k-simplices in $\mathcal{R}^\beta(\mathcal{X})$.

When $\mathbb{Q}$ is the uniform measure on a compact set $\mathbb{X} \subset \mathbb{R}^d$ with $\mathrm{Vol}(\mathbb{X}) > 0$ and $\gamma = d$, the ungainly quantity $\mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma})$ studied below is equivalent to the more natural clique count $\mathcal{C}_k(\widetilde{\mathcal{P}}_1 \cap s^{1/d}\mathbb{X}, \beta)$, where $\widetilde{\mathcal{P}}_1$ is a rate one stationary Poisson point process in $\mathbb{R}^d$ and $\widetilde{\mathcal{P}}_1 \cap s^{1/d}\mathbb{X}$ is its restriction to $s^{1/d}\mathbb{X}$.

Theorem 5.15. *Let $k \in \mathbb{N}$ and $\beta \in (0, \infty)$ and assume there are constants $c_1 \in (0, \infty)$ and $c_2 \in (0, \infty)$ such that*

$$\mathbb{P}(M_x \geq r) \leq c_1 \exp\left(-\frac{r^{c_2}}{c_1}\right), \quad x \in \mathbb{X},\ r \in (0, \infty). \tag{5.24}$$

If $\inf_{s \geq 1} \mathrm{Var}\, \mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma})/s > 0$, *then there is a constant $\tilde{C} \in (0, \infty)$ such that*

$$d_K\left(\frac{\mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma}) - \mathbb{E}\, \mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma})}{\sqrt{\mathrm{Var}\, \mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma})}}, N\right) \leq \frac{\tilde{C}}{\sqrt{s}}, \quad s \geq 1. \tag{5.25}$$

Likewise if $\inf_{n \geq 9} \mathrm{Var}\, \mathcal{C}_k(\mathcal{X}_n, \beta n^{-1/\gamma})/n > 0$, *then there is a constant $\tilde{C} \in (0, \infty)$ such that*

$$d_K\left(\frac{\mathcal{C}_k(\mathcal{X}_n, \beta n^{-1/\gamma}) - \mathbb{E}\, \mathcal{C}_k(\mathcal{X}_n, \beta n^{-1/\gamma})}{\sqrt{\mathrm{Var}\, \mathcal{C}_k(\mathcal{X}_n, \beta n^{-1/\gamma})}}, N\right) \leq \frac{\tilde{C}}{\sqrt{n}}, \quad n \geq 9. \tag{5.26}$$

Remarks. (i) When $\mathbb{X}$ is a full-dimensional subset of $\mathbb{R}^d$ and when $M_x \equiv 1$ for all $x \in \mathbb{X}$, i.e., $\mathbb{Q}_{\mathbb{M}}$ is the Dirac measure concentrated at one, a central limit theorem for the Poisson case is shown in [25, Theorem 3.10]. Although the result in [25] is non-quantitative, the method of proof should yield a rate of convergence for the Kolmogorov distance. Rates of normal convergence with respect to the Wasserstein distance d_W are given in [13].
(ii) The contributions of this theorem are three-fold. First, $\mathbb{X}$ may be an arbitrary metric space, not necessarily a subset of $\mathbb{R}^d$. Second, the graphs $G(\mathcal{P}_s, \beta s^{-1/\gamma})$ and $G(\mathcal{X}_n, \beta n^{-1/\gamma})$ are more general than the standard random geometric graph, as they consist of edges having arbitrary (exponentially decaying) lengths. Third, by applying our general findings we obtain presumably optimal rates of convergence for the Poisson and the binomial case at the same time.
(iii) The random variable $\mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma})$ is a so-called Poisson U-statistic. In the case $M_x \equiv 1$, bounds for the normal approximation of such random variables were deduced, for example, in [36] and [22] for the Wasserstein distance and in [45] and [15] for the Kolmogorov distance. These results should also yield bounds similar to those in (5.25).
(iv) The assumption $\inf_{s \geq 1} \mathrm{Var}\, \mathcal{C}_k(\mathcal{P}_s, \beta s^{-1/\gamma})/s > 0$ is satisfied if $\mathbb{X} \subset \mathbb{R}^d$ is a full d-dimensional set and g is a bounded probability density, as noted in the proof of Theorem

2.5 in Section 6 of [32]. If this assumption is not satisfied then we would have instead

$$d_K\left(\frac{\mathcal{C}_k(\mathcal{P}_s,\beta s^{-1/\gamma})-\mathbb{E}\,\mathcal{C}_k(\mathcal{P}_s,\beta s^{-1/\gamma})}{\sqrt{\operatorname{Var}\mathcal{C}_k(\mathcal{P}_s,\beta s^{-1/\gamma})}},N\right)$$
$$\leq \tilde{C}\left(\frac{\sqrt{s}}{\operatorname{Var}\mathcal{C}_k(\mathcal{P}_s,\beta s^{-1/\gamma})}+\frac{s}{(\operatorname{Var}\mathcal{C}_k(\mathcal{P}_s,\beta s^{-1/\gamma}))^{3/2}}+\frac{s^{3/2}}{(\operatorname{Var}\mathcal{C}_k(\mathcal{P}_s,\beta s^{-1/\gamma}))^{2}}\right),\quad s\geq 1.$$

A similar comment applies for an underlying binomial point process in the situation where $\inf_{n\geq 9}\operatorname{Var}\mathcal{C}_k(\mathcal{X}_n,\beta n^{-1/\gamma})/n>0$ does not hold.

Proof. To deduce Theorem 5.15 from Corollary 2.2, we express $\mathcal{C}_k(\mathcal{X},\beta s^{-1/\gamma})$ as a sum of stabilizing score functions, which goes as follows. Fix $\gamma,s,\beta\in(0,\infty)$. For $\mathcal{X}\in\mathbf{N}$ and $x\in\mathcal{X}$ let $\phi_{k,s}^{(\beta)}(x,\mathcal{X})$ be the number of $(k+1)$-cliques containing x in $G(\mathcal{X},\beta s^{-1/\gamma})$ and such that x is the point with the largest mark. This gives the desired identification

$$\mathcal{C}_k(\mathcal{X},\beta s^{-1/\gamma})=\sum_{x\in\mathcal{X}}\phi_{k,s}^{(\beta)}(x,\mathcal{X}).$$

Now we are ready to deduce (5.25) and (5.26) from Corollary 2.2 with the scores ξ_s and ξ_n set to $\phi_{k,s}^{(\beta)}$ and $\phi_{k,n}^{(\beta)}$, respectively, and with K set to $\mathbb{X}$. Notice that $I_{K,s}=\Theta(s)$, as noted in (2.13). It is enough to show that $\phi_{k,s}^{(\beta)}$ and $\phi_{k,n}^{(\beta)}$ satisfy all conditions of Corollary 2.2. Stabilization (2.4) is satisfied with $\alpha_{stab}=a$, with the radius of stabilization

$$R_s((x,M_x),\mathcal{P}_s\cup\{(x,M_x)\})=\beta s^{-1/\gamma}M_x,$$

because M_x has exponentially decaying tails as in (5.24). For any $p>0$ we have

$$\mathbb{E}\,|\phi_{k,s}^{(\beta)}((x,M_x),\mathcal{P}_s\cup\{(x,M_x)\cup(\mathcal{A},M_{\mathcal{A}})\})|^{4+p}$$
$$\leq\mathbb{E}\,|\mathrm{card}\{\mathcal{P}_s\cap B(x,\beta s^{-1/\gamma}M_x)\}+7|^{(4+p)k}\leq C(\beta,p,\gamma)<\infty$$

for all $x\in\mathbb{X}$, $s\geq 1$ and $\mathcal{A}\subset\mathbb{X}$ with $|\mathcal{A}|\leq 7$ and so the $(4+p)$-moment condition (2.6) holds for $p\in(0,\infty)$. The conclusion (5.25) follows from (2.14). The proof of (5.26) is similar. □

6 Appendix

Here we provide the proof of Lemma 5.8.

Proof. The second inequality in 5.18 follows since $\theta_j\leq 1$, and so we only need to prove the first inequality. We need some additional notation. Throughout the proof, κ is a constant depending on d,j, whose value may change from line to line. For L some linear space, let ℓ^L the Lebesgue measure on L, $G(L,q),q<\dim(L)$ its space of q-dimensional subspaces, and ν_q^L the Haar probability measure on $G(L,q)$. Note $G(\mathbb{R}^d,j) = G(d,j)$ and $\nu_j = \nu_j^{\mathbb{R}^d}$. Theorem 6.2.2 from [39] yields

$$\begin{aligned} V_j(A) - V_j(M) &= \kappa \int_{G(d,j)} (V_j(A|L) - V_j(M|L))\, \nu_j(\mathrm{d}L) \\ &= \kappa \int_{G(d,j)} \int_L f^L(x)\, \ell^L(\mathrm{d}x)\, \nu_j(\mathrm{d}L). \end{aligned}$$

The Blaschke-Petkantschin formula (Theorem 7.2.1 in [39]) over the ℓ^L integral shows that the right-hand side equals

$$\kappa \int_{G(d,j)} \int_{G(L,1)} \int_N f^L(x) \|x\|^{j-1}\, \ell^N(\mathrm{d}x)\, \nu_1^L(\mathrm{d}N)\, \nu_j(\mathrm{d}L).$$

Fubini's theorem and Theorem 7.1.1 in [39] yield that the last expression is

$$\begin{aligned} &\kappa \int_{G(d,1)} \int_{G(N,j)} \int_N f^L(x) \|x\|^{j-1}\, \ell^N(\mathrm{d}x)\, \nu_j^N(\mathrm{d}L)\, \nu_1(\mathrm{d}N) \\ =&\kappa \int_{G(d,1)} \int_N \int_{G(N,j)} f^L(x) \|x\|^{j-1}\, \nu_j^N(\mathrm{d}L)\, \ell^N(\mathrm{d}x)\, \nu_1(\mathrm{d}N) \\ =&\kappa \int_{G(d,1)} \int_N \|x\|^{j-1} \int_{G(L(x),j)} f^L(x)\, \nu_j^{L(x)}(\mathrm{d}L)\, \ell^N(\mathrm{d}x)\, \nu_1(\mathrm{d}N) \\ =&\kappa \int_{G(d,1)} \int_N f(x)\, \ell^N(\mathrm{d}x)\, \nu_1(\mathrm{d}N) \end{aligned} \tag{6.1}$$

with $f(x) = \|x\|^{j-1} \int_{G(L(x),j)} f^L(x)\, \nu_j^{L(x)}(\mathrm{d}L)$ because $N = L(x)$ in the second line. An independent application of the Blaschke-Petkantschin formula with $g(x) = f(x)\|x\|^{-(d-1)}$ for each L yields

$$\begin{aligned} \int_{\mathbb{R}^d} g(x)\, \ell^d(\mathrm{d}x) &= \int_{G(d,1)} \int_N g(x) \|x\|^{d-1}\, \ell^N(\mathrm{d}x)\, \nu_1(\mathrm{d}N) \\ &= \int_{G(d,1)} \int_N f(x)\, \ell^N(\mathrm{d}x)\, \nu_1(\mathrm{d}N) \end{aligned}$$

whence (6.1) is equal to $\int_{\mathbb{R}^d} \int_{G(L(x),j)} f^L(x) \|x\|^{(j-1)-(d-1)}\, \nu_j^{L(x)}(\mathrm{d}L)\, \ell^d(\mathrm{d}x)$, which completes the proof. □

References

[1] F. Avram and D. Bertsimas (1993), On central limit theorems in geometrical probability, *Ann. Appl. Probab.* **3**, no. 4, 1033-1046.

[2] Z-D Bai, H-K. Hwang, and T-H. Tsai (2003), Berry Esseen bounds for the number of maxima in planar regions, *Electronic Journal of Prob.* **8**, 1-26.

[3] I. Bárány, F. Fodor, V. Vígh (2010), Intrinsic volumes of inscribed random polytopes in smooth convex bodies, *Adv. Applied Prob.*, **42**, 605-619.

[4] I. Bárány and V. Vu (2007), Central limit theorems for Gaussian polytopes, *Ann. Probab.* **35**, 1593-1621.

[5] A. D. Barbour and A. Xia (2001), The number of two dimensional maxima, *Adv. in Appl. Probab.* **33**, 727-750.

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

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

[8] Yu. Baryshnikov, M. Penrose, and J. E. Yukich (2009), Gaussian limits for generalized spacings, *Ann. Appl. Probab.* **19**, No. 1, 158 - 185.

[9] P. Bickel and L. Breiman (1983), Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test, *Ann. Probab.* **11**, 1, 185-214.

[10] P. Calka, T. Schreiber, J. Yukich (2013), Brownian limits, local limits, extreme value and variance asymptotics for convex hulls in the unit ball, *Ann. Probab.* **41**, 50-108.

[11] S. Chatterjee (2008), A new method of normal approximation *Ann. Probab.* **36**, 1584-1610.

[12] W-M. Chen, H-K. Hwang and T-H. Tsai (2003), Efficient maxima finding algorithms for random planar samples, *Discrete Mathematics and Theor. Comp. Sci.* **6**, 107-122.

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

[14] P. Eichelsbacher, M. Raic, and T. Schreiber (2015), Moderate deviations for stabilizing functionals in geometric probability, *Ann. Inst. Henri Poincaré Probab. Stat.* **51**, 89-128.

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

[16] L. Goldstein, T. Johnson, and R. Lachièze-Rey (2015), Bounds to the normal for proximity region graphs, preprint, arXiv:1510.09188.

[17] J. Grote and C. Thäle (2015), Concentration and moderate deviations for Poisson polytopes and polyhedra, arXiv:1508.04994

[18] C. Hirsch (2014), From heavy-tailed Boolean models to scale-free Gilbert graphs, ArXiv:1411.6824v1.

[19] R. Jiménez and J. E. Yukich (2011), Nonparametric estimation of surface integrals, *Ann. Stat.* **39**, 232-260.

[20] H. Kesten and S. Lee (1996), The central limit theorem for weighted minimal spanning trees on random points, *Ann. Appl. Probab.* **6**, 2, 495-527.

[21] R. Lachièze-Rey and S. Vega (2016), Boundary density and Voronoi set estimation for irregular sets, *Trans. American Math. Soc.*, to appear.

[22] R. Lachièze-Rey and G. Peccati (2016+), New Kolmogorov Berry-Esseen bounds for functionals of binomial point processes, *Ann. Appl. Prob.*, to appear.

[23] G. Last, G. Peccati and M. Schulte (2016+), Normal approximations on the Poisson space: Mehler's formula, second order Poincaré inequalities and stabilization, To appear in Probab. Theory and Related Fields.

[24] P. McMullen (1975), Non-linear angle-sum relations for polyhedral cones and polytopes, *Math. Proc. Camb. Phil. Soc.*, **78**, 247-261.

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

[26] M. D. Penrose (2007), Gaussian limits for random geometric measures, *Electron. J. Probab* **12**, 989-1035.

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

[28] M. D. Penrose and T. Rosoman (2008), Error bounds in stochastic-geometrical normal approximation, Fifth Colloquium on Mathematics and Computer Science, Assoc. Discrete Math. Theor. Comput. Sci., Nancy, 71-94.

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

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

[31] 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. Singapore, **5**, A. D. Barbour and Louis H. Y. Chen, Eds., 37-58.

[32] M. D. Penrose and J. E. Yukich (2013), Limit theory for point processes in manifolds, *Ann. Appl. Probab.* **23**, 2161-2211.

[33] F. P. Preparata and M. I. Shamos (1985), Computational Geometry: An Introduction, Springer-Verlag, New York.

[34] J. Rataj and S. Winter (2010), On volume and surface area of parallel sets, *Indiana University Mathematics Journal* **59**, no. 5, 1661-1685.

[35] M. Reitzner (2005), Central limit theorems for random polytopes, *Probab. Theory Related Fields* **133**, 488-507.

[36] M. Reitzner and M. Schulte (2013), Central limit theorems for U-statistics of Poisson point processes, *Ann. Probab.* **41**(6), 3879-3909.

[37] A. Rényi and R. Sulanke (1963), Über die konvexe Hülle von n zufállig gewählten Punkten. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, **2**, 75-84.

[38] A. Rényi and R. Sulanke (1964), Über die konvexe Hülle von n zufállig gewählten Punkten II. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete*, **3**, 138-147.

[39] R. Schneider and W. Weil (2008), Stochastic and Integral Geometry, Springer.

[40] R. Schoen and S.-T Yau (1994), Lectures on Differential Geometry, Conference Proceedings and Lecture Notes in Geometry and Topology, International Press.

[41] T. Schreiber (2010), Limit theorems in stochastic geometry, New perspectives in stochastic geometry, Oxford University Press, Oxford, 111-144.

[42] T. Schreiber, M. Penrose, and J. E. Yukich (2007), Gaussian limits for multidimensional random sequential packing at saturation, Comm. Math. Physics **272**, 167-183.

[43] T. Schreiber and J. E. Yukich (2008), Variance asymptotics and central limit theorems for generalized growth processes with applications to convex hulls and maximal points, *Ann. Probab.* **36**, 363-396.

[44] M. Schulte (2012), A central limit theorem for the Poisson-Voronoi approximation, *Adv. in Appl. Math.* **49**, 285-306.

[45] M. Schulte (2012), Normal approximation of Poisson functionals in Kolmogorov distance, *J. Theoret. Probab.* **29**, 96-117.

[46] L. A. Sholomov (1983), Survey of estimational results in choice problems. *Engrg. Cybernetics* **21**, 51-75.

[47] L. L. Stacho (1976), On the volume function of parallel sets, *Acta Sci. Math.* **38**, 365-374.

[48] C. Thäle, J. E. Yukich (2016), Asymptotic theory for statistics of the Poisson-Voronoi approximation, Bernoulli **22** (4), 2372 - 2400.

[49] V. Vu (2006), Central limit theorems for random polytopes in a smooth convex set, *Adv. in Math.* **207**, 221-243.

[50] J. E. Yukich (2015), Surface order scaling in stochastic geometry, *Ann. Appl. Probab.* **25**, 177-210.