Gaussian Limits for Generalized Spacings

Yu. Baryshnikov, [1] M. D. Penrose, [2] and J. E. Yukich [3]

Bell Laboratories, University of Bath, and Lehigh University

October 19, 2007

Abstract

Nearest neighbor cells in $\mathbb{R}^d$, $d \in \mathbb{N}$, are used to define coefficients of divergence (ϕ-divergences) between continuous multivariate samples. For large sample sizes, such distances are shown to be asymptotically normal with a variance depending on the underlying point density. The finite-dimensional distributions of the point measures induced by the coefficients of divergence converge to those of a generalized Gaussian field with a covariance structure determined by the point densities. In $d = 1$, this extends classical central limit theory for sum functions of spacings. The general results yield central limit theorems for logarithmic k-spacings, information gain, log-likelihood ratios, and the number of pairs of sample points within a fixed distance of each other.

[1] Rm 2C-323, Bell Laboratories, Lucent Technologies, 600-700 Mountain Ave, Murray Hill, NJ 07974: ymb@research.bell-labs.com

[2] Department of Mathematical Sciences, University of Bath, Bath BA1 7AY, United Kingdom: m.d.penrose@bath.ac.uk

[3] Department of Mathematics, Lehigh University, Bethlehem PA 18015, USA: joseph.yukich@lehigh.edu

[3] This material is based upon work supported by the NSA under grant H98230-06-1-0052

Key words and phrases: ϕ-divergence, central limit theorems, spacing statistics, logarithmic spacings, information gain, log-likelihood

1 Introduction

Suppose $X_{(i)}$, $1 \leq i \leq n$, are the order statistics drawn from an i.i.d. sample with distribution F on $\mathbb{R}$ and let G be a distribution function. Classical spacing functionals on $\mathbb{R}$ [41] take the form of an *empirical ϕ-divergence*

$$\sum_{i=1}^{n-1} \phi(n[G(X_{(i+1)}) - G(X_{(i)})]), \tag{1.1}$$

where $\phi : \mathbb{R}^+ \to \mathbb{R}$ is a specified function and where typically F is unknown. When F and G have densities f and g, respectively, the functionals (1.1) represent an empirical version of the *ϕ-divergence* of g from f, namely $\int f(x)\phi(\frac{g(x)}{f(x)})dx$. The ϕ-divergence functional, introduced by Ali and Silvey [1, 2, 3] and independently by Csiszár [11, 12, 13], is a measure of the discrepancy of G relative to F. Empirical ϕ-divergences are widely used in non-parametric estimation and are well suited for goodness-of-fit tests [8, 10, 17, 25, 40, 43, 50].

This paper has two main goals. The first is to use kth nearest neighbor cells to establish high-dimensional analogs of the ϕ-divergences (1.1). The nearest neighbor cells are employed to define the statistical discrepancy of a proposed distribution with density g relative to an observed i.i.d. sample drawn from a distribution with density f. We establish a general central limit theorem (CLT) showing that the resulting distance functionals converge to a normal random variable whenever f and g are bounded away from zero and infinity. The limiting variance is given in terms of the $V_{\phi,k}$-divergence and $\Delta_{\phi,k}$-divergence of g from f, where $V_{\phi,k}$ and $\Delta_{\phi,k}$ are certain integral transforms of ϕ.

Our second goal is to use ϕ-divergences based on kth nearest neighbors cells to provide a unifying approach towards proving classical CLTs for sum functions of k-spacings [8, 10, 14, 15, 20, 47, 50]. This yields asymptotic normality for information gain, log-likelihood ratios, and sums of logarithmic spacings whenever the densities of F and G are bounded away from zero and infinity. The methods extend to yield a CLT for the number of pairs of sample points within a fixed distance.

Going beyond univariate central limit theorems, we consider the natural random measures associated with the empirical ϕ-divergences, obtained for $d = 1$ by putting a point mass at each $X_{(i)}$ of size equal to the ith term in (1.1), and analogously for $d > 1$. We show that the finite-dimensional distributions of these point measures, after re-normalization, converge to those of a mean zero finitely additive Gaussian field. We are motivated to consider these measures since in applications it might be useful to compare the discrepancy of f and g over subsets of $\mathbb{R}^d$ and not just the whole of $\mathbb{R}^d$.

Our approach uses *stabilization* methods, a tool [5], [35]-[38], for establishing general limit theorems for sums of weakly dependent terms in geometric probability. These methods quantify local dependence in ways useful for establishing thermodynamic and Gaussian limits and they also show that locally defined functionals of Poisson points on large bounded sets can be well approximated by globally defined functionals of homogeneous Poisson points on all of $\mathbb{R}^d$. This latter feature conveniently often leads to explicit thermodynamic and variance asymptotics.

Existing general limit results cannot be applied directly to the high-dimensional analogs of (1.1). However, these empirical ϕ-divergences are nonetheless defined in terms of nearest neighbor cells and one might thus expect that the underlying ideas and methods at the heart of stabilization are applicable and lead to Gaussian limits for the high-dimensional analogs of (1.1). This paper shows that this is indeed the case. The approach bypasses the need to treat ϕ-divergences of non-uniform samples as a limiting case of statistics over samples having step densities; such an approach requires interchanging limits, an obstacle to rigorous analysis even in dimension $d = 1$.

By establishing that the high-dimensional analogs of (1.1) are sums of stabilizing functionals, we establish convergence of the associated pair correlation functional, thus identifying limiting variances in the setting of Poisson samples. Further, by adapting the methods of [36] to the present setting, we may prove central limit theorems over point sets with a fixed (non-Poisson) number of points.

2 Main Results

2.1 Preliminaries

Notation. We use the following notation throughout. If B is a Borel subset of $\mathbb{R}^d$, then $|B|$ denotes its Lebesgue measure. Given $\mathcal{X} \subset \mathbb{R}^d$, $a \geq 0$, and $y \in \mathbb{R}^d$, let $y + a\mathcal{X} := \{y + ax : x \in \mathcal{X}\}$. For $x \in \mathbb{R}^d$, let $|x|$ be its Euclidean modulus and for

$r > 0$, let $B_r(x)$ denote the open Euclidean ball $\{y \in \mathbb{R}^d : |y-x| < r\}$. Let $\mathbf{0}$ denote the origin of $\mathbb{R}^d$, and let $\omega_d := |B_1(\mathbf{0})| = \pi^{d/2}/\Gamma((d/2)+1)$. We use $\log x$ to denote the natural logarithm of x.

We let f and g denote two probability density functions on $\mathbb{R}^d$ ($d \in \mathbb{N}$) with common compact support, which we assume is convex and which is denoted by A. *We assume once and for all that f and g are bounded and that they are bounded away from zero on A.* Abusing notation, we let $F(\cdot)$ (respectively $G(\cdot)$) denote the probability measure on $\mathbb{R}^d$ with density f (respectively g), i.e. $F(B) := \int_B f(x)dx$ and $G(B) := \int_B g(x)dx$.

Throughout $X_1, X_2, ...$ denotes a sequence of independent random d-vectors with common density f. Let $\mathcal{X}_n := \{X_1, ..., X_n\}$. Also, given $\lambda > 0$, let $\mathcal{P}_\lambda$ be a Poisson point process in A with intensity function $\lambda f : A \to \mathbb{R}^+$. For all $a > 0$, let $\mathcal{H}_a$ denote a homogeneous Poisson point process on $\mathbb{R}^d$ with intensity a. We write $\mathcal{H}$ for $\mathcal{H}_1$.

Given a Borel subset $E \subset \mathbb{R}^d$, let $\mathcal{B}(E)$ denote the class of bounded Borel-measurable real-valued functions on E. Given $h \in \mathcal{B}(\mathbb{R}^d)$, we write $\|h\|_\infty$ for $\sup_{x\in\mathbb{R}^d}(|h(x)|)$ and given also μ a Borel measure on $\mathbb{R}^d$, we let $\langle h, \mu\rangle$ denote the integral of h with respect to μ.

We shall consider ϕ-divergences and related quantities for a general class $\mathcal{F}$ of functions ϕ, which we now describe. Let $\mathbb{R}^+ := (0, \infty)$. Given a continuous function $\phi : \mathbb{R}^+ \to \mathbb{R}$, define the function $\phi^* : \mathbb{R}^+ \to [0, \infty)$ by

$$\phi^*(t) := \begin{cases} \sup\{|\phi(u)| : t \le u \le 1\} & \text{if } \ 0 < t \le 1 \\ \sup\{|\phi(s)| : 1 \le s \le t\} & \text{if } \ t \ge 1. \end{cases} \tag{2.1}$$

In other words, ϕ^* is the minimal function on $\mathbb{R}^+$ with the properties that (i) $-\phi^*(\cdot)$ is unimodal with a maximum at 1, and (ii) $\phi^*(\cdot)$ dominates $|\phi(\cdot)|$ pointwise.

Let $\mathcal{F}$ be the class of continuous functions $\phi : \mathbb{R}^+ \to \mathbb{R}$ such that the restriction to $(0, 1)$ of the function ϕ^* defined by (2.1) is square-integrable on $(0, 1)$, and such that $\log(\max(\phi(t), 1)) = o(t)$ as $t \to \infty$. Let $\mathcal{F}_0$ be the class of functions in $\mathcal{F}$ which are bounded on $(0, 1]$.

Let Γ_1 denote a gamma(1,1) random variable i.e., exponentially distributed with mean one. Letting $\Gamma_{1,i}$, $i \ge 1$, be independent copies of Γ_1, we put $\Gamma_k := \sum_{i=1}^k \Gamma_{1,i}$, a gamma random variable with parameters k and 1. For $\sigma^2 > 0$ let $N(0, \sigma^2)$ denote a normal random variable with mean zero and variance σ^2. Given random variables

X, Y we write $X \prec Y$ (or $Y \succ X$) if Y dominates X stochastically, i.e. if $P[X \leq t] \geq P[Y \leq t]$ for all $t \in \mathbb{R}$.

2.2 High-dimensional ϕ-divergence based on k-nearest neighbor cells

Let $\mathcal{K}$ be an open convex cone in $\mathbb{R}^d$ (a cone is a set that is invariant under dilations). For all $r > 0$ let $B_r^{\mathcal{K}}(x) := x + (\mathcal{K} \cap B_r(\mathbf{0}))$. Recall that the aspect ratio of a subset E of $\mathbb{R}^d$ is the ratio of the radius of the smallest ball containing E and the radius of the largest ball contained in E. For $d \geq 2$ we assume that $\mathcal{K}$ is 'regular' with respect to A, that is $\mathcal{K}$ is chosen such that the sets $B_r^{\mathcal{K}}(x) \cap A$ have bounded aspect ratio uniformly over $x \in A,\ r > 0$. When $\mathcal{K} = \mathbb{R}^d$ this condition is trivially satisfied. If A is the unit cube then $\mathcal{K}$ may be either a tilted orthant or a right circular cone not tangent to any coordinate subspace.

Given the cone $\mathcal{K}$, $x \in \mathbb{R}^d$, a finite set $\mathcal{X} \subset \mathbb{R}^d$, and $k \in \mathbb{N}$, put

$$C_k(x, \mathcal{X}) := C_k^{\mathcal{K}}(x, \mathcal{X}) := \bigcup_{t>0:\mathrm{card}(B_t^{\mathcal{K}}(x) \cap \mathcal{X} \setminus \{x\}) < k} B_t^{\mathcal{K}}(x). \tag{2.2}$$

Here $\mathrm{card}(\mathcal{Y})$ denotes the cardinality of the finite set $\mathcal{Y}$. If $\mathrm{card}((x+\mathcal{K}) \cap \mathcal{X} \setminus \{x\}) \geq k$ then $C_k^{\mathcal{K}}(x, \mathcal{X})$ is the largest set of the form $B_t^{\mathcal{K}}(x)$ containing fewer than k points of $\mathcal{X} \setminus \{x\}$; otherwise, $C_k^{\mathcal{K}}(x, \mathcal{X})$ is the whole 'wedge' $x + \mathcal{K}$. When $\mathcal{K} = \mathbb{R}^d$, $C_k^{\mathcal{K}}(x, \mathcal{X})$ is a ball whose radius is the distance between x and its kth nearest neighbor in $\mathcal{X} \setminus x$.

For each $n \geq 2$ and $X_i,\ 1 \leq i \leq n$, we use the *directed nearest neighbor cells* $C_k^{\mathcal{K}}(X_i, \mathcal{X}_n)$ to define high-dimensional spacing functionals analogous to the classical one-dimensional functionals (1.1). Define for $1 \leq i \leq n$ the transformed kth nearest neighbor spacings

$$D_{i,n,k}^g := G(C_k^{\mathcal{K}}(X_i, \mathcal{X}_n)).$$

Given $\phi \in \mathcal{F}$, define the random point measure $\nu_{n,\phi,k}^g$, with total measure $N_{n,\phi,k}^g$, as follows:

$$\nu_{n,\phi,k}^g := \sum_{i=1}^n \phi(nD_{i,n,k}^g)\delta_{X_i}; \quad N_{n,\phi,k}^g := \sum_{i=1}^n \phi(nD_{i,n,k}^g). \tag{2.3}$$

Here δ_x denotes the unit point mass at x. Let $\overline{\nu}_{n,\phi,k}^g := \nu_{n,\phi,k}^g - \mathbb{E}\,[\nu_{n,\phi,k}^g]$ be the centered version of the measure $\nu_{n,\phi,k}^g$.

Henceforth we call $N^g_{n,\phi,k}$ the 'k-nearest neighbors spacing statistic', or 'empirical nearest neighbor ϕ-divergence'; it provides a high-dimensional analog of the statistic (1.1). Our main concern is with the limit theory of $\nu^g_{n,\phi,k}$ and $N^g_{n,\phi,k}$.

The statistic $N^g_{n,\phi,k}$ provides an empirical measure of the discrepancy of the proposed distribution G from the (typically unknown) true distribution F. For example, if $k = 1$, then equating $D^f_{i,n,1}$ with its approximate expected value of $1/n$ yields the approximation $N^g_{n,\phi,1} \approx \sum_i \phi\left(D^g_{i,n,1}/D^f_{i,n,1}\right)$, and thus $N^g_{n,\phi,1}$ provides a naive empirical estimate for the so-called *ϕ-divergence* [1, 2, 3, 11, 12, 13] of g from f which is defined by

$$I_\phi(g,f) := \int_A \phi\left(\frac{g(x)}{f(x)}\right) f(x)dx. \tag{2.4}$$

In general $I_\phi(g,f)$ is possibly negative, and $I_\phi(g,f) = I_{\phi^*}(f,g)$ where $\phi^*(x) := x\phi(x^{-1})$. Also,

$$I_\phi(f,f) = \phi(1); \tag{2.5}$$

$$I_\phi(g,f) \geq I_\phi(f,f) \quad \text{if } \phi \text{ is convex.} \tag{2.6}$$

Choices of $\phi \in \mathcal{F}$ figuring prominently in estimation and decision theory include:

- $\phi_0(x) := -\log x$ defines Kullback-Leibler information (also called the modified log-likelihood ratio statistic or relative entropy) and is used in maximum spacing methods,
- $\phi_{1/2}(x) := 2(1-\sqrt{x})^2$ yields the square of the Hellinger distance,
- $\phi_1(x) := x\log x$ yields the log-likelihood ratio statistic or I-divergence of Kullback-Leibler,
- $\phi_2(x) := (x-1)^2/2$ yields the chi-squared divergence, and
- $\phi^{(r)}(x) := x^r$ yields information gain of order r, $r > 0$.

The ϕ-divergences $N^g_{n,\phi,k}$ and $I_\phi(g,f)$ ('coefficients of divergence') are used heavily in goodness-of-fit tests [43] and are useful in characterizing the amount of information of one distribution contained in another [43, 44]. Note that (2.6) shows

$I_{\phi_0}(g,f)$ and $I_{\phi_1}(g,f)$ are nonnegative, and that $I_{\phi^{(1/2)}}(g,f)$ is symmetric in f and g.

The following integral transforms of ϕ (defined for $\beta > 0$) arise naturally in the asymptotic analysis of $\nu^g_{n,\phi,k}$ (the random variables Γ_k were defined in Section 2.1):

$$
\begin{aligned}
M_{\phi,k}(\beta) &:= \mathbb{E}\,[\phi(\beta\Gamma_k)], & (2.7)\\
\Delta_{\phi,k}(\beta) &:= (k+1)M_{\phi,k}(\beta) - kM_{\phi,k+1}(\beta), & (2.8)\\
V_{\phi,k}(\beta) &:= M_{\phi^2,k}(\beta)\\
&+ \int_{\mathbb{R}^d} \left[\mathbb{E}\,[\phi(\beta|C_k(\mathbf{0},\mathcal{H}\cup y)|)\phi(\beta|C_k(y,\mathcal{H}\cup\mathbf{0})|)] - M_{\phi,k}(\beta)^2\right] dy & (2.9)
\end{aligned}
$$

Note that $M_{\phi,1}(x) = (1/x)\hat{\phi}(1/x)$, where $\hat{\phi}$ denotes the Laplace transform of ϕ.

2.3 A general CLT for ϕ-divergences

The following general CLT, our main result, establishes convergence of the finite-dimensional distributions of $n^{-1/2}\overline{\nu}^g_{n,\phi,k}$ as $n \to \infty$ (i.e., the convergence of the m-vector $n^{-1/2}\left(\langle h_1, \overline{\nu}^g_{n,\phi,k}\rangle, ..., \langle h_m, \overline{\nu}^g_{n,\phi,k}\rangle\right)$ for all $h_1, ..., h_m$ in $\mathcal{B}(A)$) to the finite-dimensional distributions of a mean zero finitely additive Gaussian field whose covariance kernel is a weighted average of the functions $V_{\phi,k}$ and $\Delta_{\phi,k}$. For $h \in \mathcal{B}(A)$, we define the *h-weighted ϕ-divergence* of g from f by

$$I_\phi(g,f,h) := \int_A f(x)\phi\left(\frac{g(x)}{f(x)}\right)h(x)dx,$$

which in the case $h \equiv 1$ reduces to the ϕ-divergence $I_\phi(f,g)$ defined at (2.4). Also, for h, h_1, h_2 in $\mathcal{B}(A)$ and $\phi \in \mathcal{F}$ we define the functions h^2, h_1h_2, ϕ^2 pointwise, i.e. $h^2(x) = (h(x))^2$ and so on.

Henceforth, by 'convergence in law' we shall mean convergence of finite-dimensional distributions. The following general CLT, proved in Sections 4 and 5, establishes convergence in law of $n^{-1/2}\overline{\nu}^g_{n,\phi,k}$ as $n \to \infty$. In the theorem below, since the formula (2.10) is rather concise we expand it in (2.11).

Theorem 2.1 *Suppose that either $\phi \in \mathcal{F}_0$, $d = 1$, or $\mathcal{K} = \mathbb{R}^d$. As $n \to \infty$, it is the case that for $h \in \mathcal{B}(A)$,*

$$n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi,k}\rangle] \to I_{V_{\phi,k}}(g,f,h^2) - (I_{\Delta_{\phi,k}}(g,f,h))^2 \qquad (2.10)$$

$$= \int_A h^2(x)V_{\phi,k}\left(\frac{g(x)}{f(x)}\right)f(x)dx - \left(\int_A h(x)\Delta_{\phi,k}\left(\frac{g(x)}{f(x)}\right)f(x)dx\right)^2 \qquad (2.11)$$

and $(n^{-1/2}\langle h, \overline{\nu}^g_{n,\phi,k}\rangle, h \in \mathcal{B}(A))$, *converges in law as* $n \to \infty$ *to a mean zero Gaussian field with covariance kernel*

$$(h_1, h_2) \mapsto I_{V_{\phi,k}}(g, f, h_1 h_2) - I_{\Delta_{\phi,k}}(g, f, h_1) I_{\Delta_{\phi,k}}(g, f, h_2). \tag{2.12}$$

Notice that putting $h \equiv 1$, Theorem 2.1 yields the following CLT for the nearest neighbors spacing statistic $N^g_{n,\phi,k}$:

$$n^{-1/2}(N^g_{n,\phi,k} - \mathbb{E}\, N^g_{n,\phi,k}) \xrightarrow{\mathcal{D}} N(0, I_{V_{\phi,k}}(g, f) - (I_{\Delta_{\phi,k}}(g, f))^2).$$

For practical purposes, it is of use to compute $V_{\phi,k}$, and the next two results show how to simplify the expression (2.9) in some special cases. Using these simplifications we may explicitly identify $V_{\phi,k}$ for certain choices of ϕ, as shown in Section 3.

The first of our simplifications applies when $\mathcal{K} \neq \mathbb{R}^d$, and either $k = 1$ or $d = 1$. The latter case is particularly relevant to the study of spacings (see Section 2.4).

Proposition 2.1 *If* $\mathcal{K} \neq \mathbb{R}^d$, *and either* $d = 1$ *or* $k = 1$, *then for all* $\beta > 0$ *we have*

$$\begin{aligned} V_{\phi,k}(\beta) = M_{\phi^2,k}(\beta) &+ 2k M_{\phi,k}(\beta)(M_{\phi,k}(\beta) - M_{\phi,k+1}(\beta)) \\ &+ 2 \sum_{j=1}^{k-1} \mathrm{Cov}[\phi(\beta\Gamma_k), \phi(\beta(\Gamma_{k+j} - \Gamma_j))], \end{aligned} \tag{2.13}$$

the sum being interpreted as zero for $k = 1$.

Our second simplifying formula for $V_{\phi,k}$ is applicable when $k = 1$, $\mathcal{K} = \mathbb{R}^d$, and ϕ is differentiable with $\lim_{t \downarrow 0} \phi(t) = 0$. This will provide limiting distributions for some statistical distances of interest, including information gain and log-likelihood in high dimensions (see Section 3.2). For $s, t, u \in \mathbb{R}^+$, let $I(s, t, u)$ be the volume of the intersection of two balls in $\mathbb{R}^d$, with respective volumes s and t, at a distance u apart. Set

$$J_d(s, t) := \int_{\max(s,t)}^{\infty} [e^{I(s,t,(u/\omega_d)^{1/d})} - 1] du. \tag{2.14}$$

Proposition 2.2 *Suppose that* $\mathcal{K} = \mathbb{R}^d$ *and that* $\phi \in \mathcal{F}$ *is differentiable with* $\lim_{t \downarrow 0} \phi(t) = 0$. *Then for all* $\beta > 0$,

$$V_{\phi,1}(\beta) = M_{\phi^2,1}(\beta) \;+\; \beta^2 \int_0^\infty \int_0^\infty \phi'(\beta s)\phi'(\beta t) e^{-(s+t)} [J_d(s, t) - \max(s, t)] ds dt$$

provided that the integral exists.

Remarks. (i) (*Related work*) Bickel and Breiman [7], and subsequently Schilling [45], consider the functionals $N^g_{n,\phi,1}$ when $\phi(x) = \exp(-x)$ and $\mathcal{K} = \mathbb{R}^d$. Using the approximation $D^g_{i,n,1} \approx g(X_i)|C_1(X_i, \mathcal{X}_n)|$, they establish a CLT for the empirical process of nearest neighbor distances, but do not consider convergence of the associated random measures. Strong limit theorems for multivariate spacings using general 'shapes' are given by Deheuvels et al. [16]. Penrose [33] finds a CLT for k-nearest neighbor distances and a strong law [31] for the largest nearest neighbor link. Henze [23] establishes the limit distribution for the maxima of weighted nearest neighbor distances.

(ii)(*Poisson CLT*) For all $\lambda > 0$, $k \in \mathbb{N}$, the Poisson analog of the point measures (2.3) is

$$\mu^g_{\lambda,\phi,k} := \sum_{x \in \mathcal{P}_\lambda} \phi\left(\lambda G(C_k(x, \mathcal{P}_\lambda))\right) \delta_x, \tag{2.15}$$

and its total measure is a Poissonized version of $N^g_{n,\phi,k}$. A byproduct of the approach used here (see Proposition 4.1 below) is a proof that if $\phi \in \mathcal{F}$ and $h \in \mathcal{B}(A)$, then as $\lambda \to \infty$,

$$\lambda^{-1}\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi,k}\rangle] \to I_{V_{\phi,k}}(g, f, h^2) \tag{2.16}$$

and $\lambda^{-1/2}\overline{\mu}^g_{\lambda,\phi,k}$ converges in law to a mean zero Gaussian field with covariance kernel $(h_1, h_2) \mapsto I_{V_{\phi,k}}(g, f, h_1 h_2)$ (here $\overline{\mu}^g_{\lambda,\phi,k} := \mu^g_{\lambda,\phi,k} - \mathbb{E}\,[\mu^g_{\lambda,\phi,k}]$).

(iii) (*Law of large numbers, limits are distribution free*) A further consequence of the approach used here (see also [35]) is the following weak law of large numbers for all $h \in \mathcal{B}(A)$ and $\phi \in \mathcal{F}$, namely

$$n^{-1}\langle h, \nu^g_{n,\phi,k}\rangle \xrightarrow{L^2} I_{M_{\phi,k}}(g, f, h).$$

By taking $h \equiv 1$, we obtain a weak law of large numbers for the k-nearest neighbors spacing statistic $N^g_{n,\phi,k}$. Combining this with Theorem 2.1 and taking $g = f$, we see from (2.5) that the limiting mean of $n^{-1}\langle h, \nu^f_{n,\phi,k}\rangle$ and the limiting variance and distribution of $n^{-1/2}\langle h, \overline{\nu}^f_{n,\phi,k}\rangle$ do not depend on f for $h \equiv 1$ (and in fact for any h). Therefore the nearest neighbor functionals are *asymptotically distribution free under the null hypothesis* $g = f$ and have asymptotic variance $V_{\phi,k}(1) - (\Delta_{\phi,k}(1))^2$. A possible goodness-of-fit test would be to take the density g to be tested, compute the functional $N^g_{n,\phi,1}$, and see whether the cumulative distribution function is close to the $N(0, V_{\phi,1}(1) - (\Delta_{\phi,1}(1))^2)$ cumulative distribution function.

(iv) (*Voronoi cells*) Volumes of nearest neighbor cells are computationally attractive and have correlations decaying exponentially with the distance between cell centers. Defining point measures analogous to (2.3) based on cells generated by any locally defined Euclidean graph (e.g. Voronoi cells) leads to similar CLTs.

(v) (*Properties of limiting variance*) In most of our examples, $\Delta_{\phi,k}$ is strictly positive, showing that Poissonization contributes extra randomness which manifests itself in a larger limiting variance. We note that when $V_{\phi,k}$ is convex, which is the case when $k = 1$, $\phi(x) = x^r$, $r \in [1, \infty)$, or when $\phi(x) = x \log x$ (see Section 3.1), then the inequality (2.6) implies that the limiting variance over Poisson samples is minimized when $g = f$.

(vi) (*Independence of limit over disjoint sets*) Theorem 2.1 (respectively Remark (ii)) says that for $h_1, ..., h_m$ in $\mathcal{B}(A)$ the m-vector $n^{-1/2}(\langle h_1, \overline{\nu}^g_{n,\phi,k}\rangle, ..., \langle h_m, \overline{\nu}^g_{n,\phi,k}\rangle)$ (respectively, the m-vector $\lambda^{-1/2}(\langle h_1, \overline{\mu}^g_{\lambda,\phi,k}\rangle, ..., \langle h_m, \overline{\nu}^g_{\lambda,\phi,k}\rangle)$ tends to a Gaussian random m-vector with covariance kernel (2.12) (respectively $I_{V_{\phi,k}}(g, f, h_1 h_2)$).

In particular, if $B_1, ..., B_m$ are disjoint, then in the Poisson setting, $\lambda^{-1/2}(\overline{\mu}^g_{\lambda,\phi,k}(B_1), \ldots, \overline{\mu}^g_{\lambda,\phi,k}(B_m))$ tends in distribution to a mean zero m-variate Gaussian with *independent* components. Knowing the distribution of the latter vector could be useful in situations where the density changes from location to location, as in change point problems.

(vii) *(Asymptotic equivalence to a simpler random field)* It is noteworthy that the limiting random field for $n^{-1/2}\overline{\nu}^g_{n,\phi,k}$ in Theorem 2.1 has the same law as the limiting random field for the point measures

$$n^{-1/2}\left(\left(\sum_{i=1}^{n} \Delta_{\phi,k}\left(\frac{g(X_i)}{f(X_i)}\right)\delta_{X_i}\right) + \sum_{X \in \mathcal{P}_n} \sqrt{V_{\phi,k}\left(\frac{g(X)}{f(X)}\right) - \Delta_{\phi,k}\left(\frac{g(X)}{f(X)}\right)^2} \delta_X\right),$$

where here the Poisson process $\mathcal{P}_n$ is taken to be independent of the sequence $\{X_i\}$, and the quantity inside the square root can be shown to be nonnegative. Note that this random field has no interactions between points.

2.4 Asymptotic normality of sum functions of spacings

In dimension $d = 1$, if g is a probability density with distribution function G on $[c_1, c_2]$, then the generalization to k-spacings of the empirical ϕ-divergence defined at (1.1) is the k-spacing statistic defined by

$$S^g_{n,\phi,k} := \sum_{i=1}^{n-k} \phi(n[G(X_{(i+k)}) - G(X_{(i)})]). \tag{2.17}$$

Developing the limit theory for $S^g_{n,\phi,k}$ over continuous samples is important in goodness-of-fit tests. Modulo some boundary effects, we can apply the general theory of Section 2.3 to these classical spacings statistics, by putting $d = 1$ and $\mathcal{K} = (0, \infty)$ in the definition of $C^{\mathcal{K}}_k(x, \mathcal{X})$. Then the width of $C^{\mathcal{K}}_k(x, \mathcal{X})$ is the distance between x and its kth nearest neighbor in $\mathcal{X}$ 'to the right'. Thus the k-nearest neighbors spacing statistic $N^g_{n,\phi,k}$, defined by (2.3), is the same as $S^g_{n,\phi,k}$ but with the sum in (2.17) extended to n terms and with $X_{(j)} := c_2$ if $j > n$.

For the sake of a better match with the existing literature, we extend the general theory by considering a modified version of Theorem 2.1 in which we redefine $C^{\mathcal{K}}_k(x, \mathcal{X})$ to be the empty set whenever $\text{card}(\mathcal{X} \cap (x + \mathcal{K}) \setminus x) < k$, and set $\phi(0) = 0$. Denote by $\nu^*_{n,\phi,k}$ the analog of $\nu^g_{n,\phi,k}$ under this modification (here we suppress the dependence on g). The corresponding centered measure is then denoted $\bar{\nu}^*_{n,\phi,k}$. If $d = 1$ and $\mathcal{K} = (0, \infty)$, the total measure of $\nu^*_{n,\phi,k}$ is indeed equal to $S^g_{n,\phi,k}$.

We assert that the limit theory of Section 2.3 is unaffected by the change from $\nu^g_{n,\phi,k}$ to $\nu^*_{n,\phi,k}$. Since we are principally interested in this modification for the case $d = 1$ and $\mathcal{K} = \mathbb{R}^+$, we restrict the formal statement of this assertion to that case.

Theorem 2.2 *(Gaussian limit for sum functions of spacings) Let* $A := [c_1, c_2]$, $\mathcal{K} = (0, \infty)$, *and* $\phi \in \mathcal{F}$. *Then the conclusion of Theorem 2.1 holds with* $\nu^g_{n,\phi,k}$ *replaced by* $\nu^*_{n,\phi,k}$. *Moreover, in this case* $V_{\phi,k}(\beta)$ *is given by* (2.13).

Applications of Theorem 2.2 are given in Section 3 and the proof is in Section 7. This result, like our main result, shows that sum functions of spacings are asymptotically distribution free under the null hypothesis $f = g$.

Remarks. (i) Darling [15] undertook the first systematic study of the functionals $S_{n,\phi,k}$ when $k = 1$, but restricted attention to uniform samples. Theorem 2.2

generalizes Holst [24], as well as earlier work of Cressie [10], who proves asymptotic normality (but not convergence of finite-dimensional distributions) for sum functions of k-spacings over *uniform* points. Holst uses a generalization of LeCam's method and a CLT for k-dependent random variables. In $d = 1$, Holst and Rao [25] prove asymptotic normality of $S^g_{n,\phi,k}$ under 'somewhat stringent conditions' on f and g. For non-uniform samples the asymptotics of $S^g_{n,\phi,k}$ have been widely studied under the assumption that G runs through a sequence of alternatives G_n approaching the uniform distribution; see Hall [22] and del Pino [40]. Khashimov [29] establishes asymptotic normality of $S^1_{n,\phi,k}$ under rather technical differentiability conditions on ϕ and f.

(ii) The approach used here also yields a weak law of large numbers, namely convergence in mean-square of $n^{-1}S^g_{n,\phi,k}$ to $I_{M_\phi}(g, f)$. This extends the corresponding weak laws in [28]. Analogous results hold for non-overlapping k-spacings [46].

2.5 Number of pairs of sample points within a fixed distance

Instead of considering point measures based on spacings, we now consider a functional using cells of *fixed* radius depending on a continuous $g : A \to \mathbb{R}^+$ and a parameter t. Thus, for all $t > 0$, all point sets $\mathcal{Y}_n \subset \mathbb{R}^d$ having cardinality n, and all $\phi \in \mathcal{F}$, we define the functional

$$H^{g,t}_{n,\phi}(\mathcal{Y}_n) := \frac{1}{2} \sum_{x \in \mathcal{X}_n} \phi\left(\mathrm{card}\{\mathcal{Y}_n \cap B_{t(ng(x))^{-1/d}}(x)\}\right).$$

When $\phi(x) \equiv x$ and $g \equiv 1$, then $H^{g,t}_{n,\phi}(\mathcal{Y}_n)$ counts the total number of pairs of points in $\mathcal{Y}_n$ distant t from each other. Recalling that $X_1, ..., X_n$ are i.i.d. with density f, we seek the asymptotic distribution of $H^{g,t}_{n,\phi}(\mathcal{X}_n)$ as well as that of the point measure

$$\nu^{g,t}_{n,\phi} := \sum_{i=1}^{n} \phi\left(\mathrm{card}\{\mathcal{X}_n \cap B_{t(ng(x))^{-1/d}}(x)\}\right) \delta_{X_i}.$$

The following CLT is obtained by modifying the proof of Theorem 2.1; see Section 7 for details.

Theorem 2.3 *(Gaussian limit for the number of pairs of points within distance t) For all continuous $g : A \to \mathbb{R}^+$, $t > 0$, and $\phi \in \mathcal{F}$, we have as $n \to \infty$*

$$n^{-1}\mathrm{Var}[H^{g,t}_{n,\phi}(\mathcal{X}_n)] \to \sigma^2_{t,\phi,g}(f)$$

for some constant $\sigma^2_{t,\phi,g}(f)$ *and*

$$n^{-1/2}(H^{g,t}_{n,\phi}(\mathcal{X}_n) - \mathbb{E}\, H^{g,t}_{n,\phi}(\mathcal{X}_n)) \xrightarrow{\mathcal{D}} N(0, \sigma^2_{t,\phi,g}(f))$$

while $n^{-1/2}\overline{\nu}^{g,t}_{n,\phi}$ *converges in law to a mean zero Gaussian field.*

Remarks. The limiting variance $\sigma^2_{t,\phi,g}(f)$ takes the same form as in the right hand side of (2.10) with the functions $V_{\phi,k}$ and $\delta_{\phi,k}$ suitably modified; see Section 7.2 for details.

Various authors have studied the functional $H^{g,t}_{n,\phi}$ when $\phi(x) \equiv x$ and $g \equiv 1$. L'Ecuyer et al. [18] considers $H^{g,t}_{n,\phi}$ from the point of view of multidimensional goodness-of-fit tests, but restricts attention to uniform samples. Penrose [32] (Ch.4) proves that the finite-dimensional distributions of the process $H^{g,t}_{n,\phi}(\mathcal{X}_n),\ t > 0$, converge to those of a Gaussian process.

3 Applications

3.1 Classical spacing statistics

For many tests involving goodness-of-fit (Dudewicz et al. [17], Blumenthal [8], Cressie [10], Holst and Rao [25], delPino [40], Weiss [50]) and parametric estimation (Ghosh and Jammalamadaka [21]) it is important to know the asymptotic distribution of $S^g_{n,\phi,k}$ (defined at (2.17)) for arbitrary g and f and for various choices of ϕ. The following provides some illustrative examples. Throughout Section 3.1 we write $V^S_{\phi,k}$ for the value of $V_{\phi,k}$ given by (2.13).

3.1.1 Limit theory for logarithms of spacings

Let

$$S^g_{n,\log,k} := \sum_{i=1}^{n-k} \log(n[G(X_{(i+k)}) - G(X_{(i)})])$$

denote the sum of the logarithmic k-spacings. Setting $\phi(x) = \log x$ in Theorem 2.2 and appealing to (2.8) and (2.13), we find a CLT for logarithms of k-spacings as follows.

Let ψ be the di-gamma function with $\psi(k) := \sum_{i=1}^{k-1} i^{-1} - \gamma$, where γ is Euler's constant, and let $\psi'(k) := -\sum_{i=1}^{k-1} i^{-2} + \pi^2/6$.

By Cressie [10] and Holst [24],

$$\sum_{j=1}^{k-1} \mathrm{Cov}(\log \Gamma_k, \log(\Gamma_{k+j} - \Gamma_j)) = k(k-1)\psi'(k) - (k-1).$$

Also, $\mathbb{E}[\log \Gamma_k] = \psi(k)$, so we have $2k\mathbb{E}[\log \Gamma_k](\mathbb{E} \log \Gamma_k - \mathbb{E} \log \Gamma_{k+1}) = -2\psi(k)$. Also, $\mathbb{E}[\log^2 \Gamma_k] = \psi'(k) + (\psi(k))^2$. So, combining terms and using (2.13) for $\phi(x) = \log x$ gives

$$V^S_{\log,k}(1) = \psi'(k) + (\psi(k))^2 - 2\psi(k) + 2[k(k-1)\psi'(k) - (k-1)]. \qquad (3.1)$$

By (2.8) we have $\Delta_{\log,k}(1) = (k+1)\psi(k) - k\psi(k+1) = \psi(k) - 1$.

Using simple relations such as $\mathrm{Cov}(\log \beta X, \log \beta Y) = \mathrm{Cov}(\log X, \log Y)$, it is straightforward to deduce that $V^S_{\log,k}(\beta) = V^S_{\log,k}(1) + \log^2 \beta + 2\log\beta(\psi(k) - 1)$ and $\Delta_{\log,k}(\beta) = \Delta_{\log,k}(1) + \log\beta$. Substituting this into Theorem 2.2, putting $\tau_k := (2k^2 - 2k + 1)\psi'(k) - 2k + 1$, and re-arranging terms yields:

Corollary 3.1 *(CLT for logarithmic k-spacings) Let $X, X_1, X_2, \ldots$ be i.i.d. with density f on $[0,1]$. As $n \to \infty$, $n^{-1}\mathrm{Var}[S^g_{n,\log,k}] \to \tau_k + \mathrm{Var}\left[\log(\frac{f(X)}{g(X)})\right]$ and*

$$n^{-1/2}(S^g_{n,\log,k} - \mathbb{E}\, S^g_{n,\log,k}) \xrightarrow{\mathcal{D}} N\left(0,\ \tau_k + \mathrm{Var}\left[\log(\frac{f(X)}{g(X)})\right]\right).$$

Remarks. When $A = [0,1]$ and $f \equiv g \equiv 1$, then the CLT for $S^g_{n,\log,k}$ was established by Darling (sect. 7 of [15]) for $k = 1$ and later by Holst [24] and Cressie [10] for general k. When the X_i have a step density then Cressie shows asymptotic normality of $S^g_{n,\log,k}$ including cases when $k \to \infty$. Czekala (Thm. 1 of [14]) apparently re-discovered Cressie's result. Shao and Hahn [47] treat general densities for $k = 1$, although their proof depends upon interchanging limits in order to pass from step densities to arbitrary densities. When $k = 1$, Blumenthal (Thm. 2 of [8]), proves Corollary 3.1 for densities f satisfying special conditions. Corollary 3.1 extends all of these results to f and g bounded away from zero and infinity, resolving a conjecture of Darling ([15], p. 249) affirmatively.

3.1.2 Information gain of order r

Let $\phi(x) = x^r$, $r > 0$. We write $S^g_{n,r,1}$ to denote $S^g_{n,\phi,1}$, also known as Rényi's information gain (I-divergence) of order r in $d = 1$, that is

$$S^g_{n,r,1} := \sum_{i=1}^{n-1} \left(n[G(X_{(i+1)}) - G(X_{(i)})] \right)^r .$$

Denote the associated point measures by

$$\nu^*_{n,r,1} := \sum_{i=1}^{n-1} (n[G(X_{(i+1)}) - G(X_{(i)})])^r \delta_{X_i}.$$

Let $w_r := -2r\Gamma^2(r+1) + \Gamma(2r+1)$ and $t_r := \Gamma(r+1)(1-r)$. It is a simple matter to verify via (2.13) and (2.8), respectively, that for all $\beta > 0$

$$V^S_{\phi,1}(\beta) := w_r \beta^{2r} \quad \text{and} \quad \Delta_{\phi,1}(\beta) := 2\mathbb{E}\left[\phi(\beta\Gamma_1)\right] - \mathbb{E}\left[\phi(\beta\Gamma_2)\right] = t_r \beta^r .$$

Put

$$\sigma_r^2(f,g) := w_r \int_A \left(\frac{g(x)}{f(x)} \right)^{2r} f(x)dx - t_r^2 \left(\int_A \left(\frac{g(x)}{f(x)} \right)^r f(x)dx \right)^2 .$$

Theorem 2.2 yields:

Corollary 3.2 *(Gaussian limits for information gain) Let $X_1, X_2, ...$ be i.i.d. with density f on $A := [c_1, c_2]$. As $n \to \infty$,*

$$n^{-1}\mathrm{Var}[S^g_{n,r,1}] \to \sigma_r^2(f,g)$$

and $n^{-1/2}(S^g_{n,r,1} - \mathbb{E}\, S^g_{n,r,1}) \xrightarrow{\mathcal{D}} N(0, \sigma_r^2(f,g))$, while $n^{-1/2}\overline{\nu}^*_{n,r,1}$ converges in law to a mean zero Gaussian field with covariance kernel

$$\begin{aligned}(h_1,h_2) \mapsto \quad & w_r \int_A h_1(x)h_2(x) \left(\tfrac{g(x)}{f(x)} \right)^{2r} f(x)dx \\ & -t_r^2 \int_A h_1(x) \left(\tfrac{g(x)}{f(x)} \right)^r f(x)dx \int_A h_2(x) \left(\tfrac{g(x)}{f(x)} \right)^r f(x)dx.\end{aligned}$$

Remarks. It is easy to verify using [5] that $\sigma_r^2(f,g) > 0$ except when $r = 1$. Corollary 3.2 extends upon the CLTs of Darling [15] (uniform case) and Weiss [50]. Moran [30] proved a CLT for the functional S_2 over uniform random variables.

3.1.3 Limit theory for log-likelihood ratio

Let $\phi(x) = x \log x$. Consider the log-likelihood point measure

$$\nu^g_{n,\phi,1} := \sum_{i=1}^{n-1} \phi(n[G(X_{(i+1)}) - G(X_{(i)})])\delta_{X_i}$$

and let S^g_ϕ denote the total mass of this measure, also called the log-likelihood statistic. Again denoting Euler's constant by γ, we have for $\beta > 0$ that

$$\begin{aligned} \mathbb{E}\,[\beta\Gamma_1 \log(\beta\Gamma_1)] &= \beta \log \beta + \beta(1-\gamma); \\ \mathbb{E}\,[\beta\Gamma_2 \log(\beta\Gamma_2)] &= 2\beta \log \beta + \beta(3-2\gamma); \\ \mathbb{E}\,[(\beta\Gamma_1 \log(\beta\Gamma_1))^2] &= \beta^2[2(\log\beta)^2 + (6-4\gamma)\log\beta + 2 + \pi^2/3 - 6\gamma + 2\gamma^2]. \end{aligned}$$

Using these in (2.13) and (2.8) respectively, it is easily verified that

$$V^S_{\phi,1}(\beta) := (\frac{\pi^2}{3} - 2)\beta^2 \text{ and } \Delta_{\phi,1}(\beta) := 2\mathbb{E}\,\phi(\beta\Gamma_1) - \mathbb{E}\,\phi(\beta\Gamma_2) = -\beta.$$

Put

$$\sigma^2_\phi(f,g) := (\frac{\pi^2}{3} - 2)\int_A \frac{g^2(x)}{f(x)}dx - \left(\int_A g(x)dx\right)^2.$$

Let X have density f and note that since g is a density we have

$$\sigma^2_\phi(f,g) = (\frac{\pi^2}{3} - 2)\mathrm{Var}\left[\frac{g(X)}{f(X)}\right] + \frac{\pi^2}{3} - 3.$$

Using the above values for $V_{\phi,1}$, $\Delta_{\phi,1}$, $\sigma^2_\phi(f,g)$, and applying Theorem 2.2 for $\phi(x) = x \log x$ yields:

Corollary 3.3 *(Gaussian limit for log-likelihood) Let $X_1, X_2, ...$ be i.i.d. with density f on $A := [c_1, c_2]$. As $n \to \infty$, $n^{-1}\mathrm{Var}[S^g_{n,\phi,1}] \to \sigma^2_\phi(f,g)$ and*

$$n^{-1/2}(S^g_{n,\phi,1} - \mathbb{E}\,S^g_{n,\phi,1}) \xrightarrow{\mathcal{D}} N(0, \sigma^2_\phi(f,g)),$$

while $n^{-1/2}\overline{\nu}^*_{n,\phi,1}$ converges in law to a mean zero Gaussian field with covariance kernel

$$(h_1,h_2) \mapsto (\frac{\pi^2}{3} - 2)\int_A h_1(x)h_2(x)\frac{g^2(x)}{f(x)}dx - \int_A h_1(x)g(x)dx \int_A h_2(x)g(x)dx.$$

Remarks. Corollary 3.3 extends the results of Gebert and Kale [20], who assume uniformity of X_i and Czekala (Thm. 2 of [14]), who assumes that X_i have a step density. Van Es [49] establishes asymptotic normality for S_ϕ^g whenever $k,\ n \to \infty$, $k = o(n^{1/2})$, and $f : A \to [0, \infty)$ is Lipschitz.

3.2 Information gain and log-likelihood in high dimensions

In this section we put $k = 1$ and $\mathcal{K} = \mathbb{R}^d$.

3.2.1 Information gain of order r

Let $\phi(x) = x^r,\ r \in \mathbb{R}^+$, so that $N_{n,\phi,1}^g$ defined by (2.3) yields Rényi's information gain (I-divergence) of order r. For all $r \in \mathbb{R}^+$, define the constant

$$K_r := r^2 \int_0^\infty \int_0^\infty s^{r-1} t^{r-1} e^{-(s+t)} [J_d(s,t) - \max(s,t)] ds dt,$$

with $J_d(s,t)$ given by (2.14). Since ϕ satisfies the conditions of Proposition 2.2 and since $\mathbb{E}\left[\phi^2(\Gamma_1)\right] = \Gamma(2r+1)$, the following is immediate.

Lemma 3.1 *For all $\beta > 0$ and for $\phi(x) = x^r,\ r > 0$, we have $V_{\phi,1}(\beta) = \beta^{2r}\left[\Gamma(2r+1) + K_r\right]$.*

Note that $\beta^{2r} = \phi^2(\beta)$. Combining Lemma 3.1 with Theorem 2.1 yields the following CLT for $N_{n,\phi,1}^g$ and the associated measures $\nu_{n,\phi,1}^g$ defined by (2.3).

Corollary 3.4 *Let $\phi(x) = x^r,\ r > 0$. Then as $n \to \infty$ $n^{-1}\mathrm{Var}[N_{n,\phi,1}^g]$ converges to $\left[\Gamma(2r+1) + K_r\right] I_{\phi^2,1}(g,f) - (I_{\Delta_{\phi,1}}(f,g))^2$, and*

$$n^{-1/2}(N_{n,\phi,1}^g - \mathbb{E}\, N_{n,\phi,1}^g) \xrightarrow{\mathcal{D}} N(0, \left[\Gamma(2r+1) + K_r\right] I_{\phi^2,1}(g,f) - (I_{\Delta_{\phi,1}}(g,f))^2)$$

while $n^{-1/2}\overline{\nu}_{n,\phi,1}^g$ converges in law to a mean zero Gaussian field with covariance kernel

$$(h_1, h_2) \mapsto [\Gamma(2r+1) + K_r] I_{\phi^2,1}(g, f, h_1 h_2) - I_{\Delta_{\phi,1}}(g, f, h_1) I_{\Delta_{\phi,1}}(g, f, h_2).$$

3.2.2 Log-likelihood

When $\phi(x) = x \log x$, $N^g_{n,\phi}$ defined by (2.3) yields the log-likelihood statistic. To apply Theorem 2.1 , we define the constants

$$I_1 := \int_0^\infty \int_0^\infty (\log s + 1)(\log t + 1)e^{-(s+t)}[J_d(s,t) - \max(s,t)]dsdt,$$

$$I_2 := \int_0^\infty \int_0^\infty (\log s + 1)e^{-(s+t)}[J_d(s,t) - \max(s,t)]dsdt,$$

and

$$I_3 := \int_0^\infty \int_0^\infty e^{-(s+t)}[J_d(s,t) - \max(s,t)]dsdt.$$

Also, set $K_1 := 2 + \frac{\pi}{3} - 6\gamma + 2\gamma^2 + I_1$, $K_2 := 6 - 4\gamma + 2I_2$, and $K_3 := 2 + I_3$. The following is an easy consequence of Proposition 2.2.

Lemma 3.2 *For $\phi(x) = x \log x$, $V_{\phi,1}(\beta) = \beta^2(K_1 + K_2 \log \beta + K_3(\log \beta)^2)$, $\beta > 0$.*

Theorem 2.1 yields asymptotic normality for the log-likelihood functional $N^g_{n,\phi,1}$ and the associated measures $\nu^g_{n,\phi,1}$ given by (2.3). Put

$$\sigma^2_\phi(f,g) := \int_A \left(\frac{g(x)}{f(x)}\right)^2 \left[K_1 + K_2 \log\left(\frac{g(x)}{f(x)}\right) + K_3\left(\log\frac{g(x)}{f(x)}\right)^2\right] f(x)dx.$$

Corollary 3.5 *Let $\phi(x) = x \log x$. Then as $n \to \infty$, $n^{-1}\mathrm{Var}[N^g_{n,\phi,1}] \to \sigma^2_\phi(f,g) - (I_{\Delta_{\phi,1}}(g,f))^2$ and*

$$n^{-1/2}(N^g_{n,\phi,1} - \mathbb{E}\, N^g_{n,\phi,1}) \xrightarrow{\mathcal{D}} N(0, \sigma^2_\phi(f,g) - (I_{\Delta_{\phi,1}}(g,f))^2)$$

while $n^{-1/2}\overline{\nu}^g_{\lambda,\phi,1}$ converges in law to a mean zero Gaussian field with covariance kernel

$$(h_1,h_2) \mapsto I_{V_{\phi,1}}(g,f,h_1h_2) - I_{\Delta_{\phi,1}}(g,f,h_1)I_{\Delta_{\phi,1}}(g,f,h_2).$$

with $V_{\phi,1}$ given by Lemma 3.2.

4 Proof of Theorem 2.1 for ϕ bounded on $(0,1]$

The proof of Theorem 2.1 involves expressing the measure $\nu^g_{n,\phi,k}$ as a sum of weakly spatially dependent terms having the property that the spatial dependence can be quantified, allowing us to show convergence of the associated two point correlation function (over Poisson samples) and thus crucially establish convergence of the variance of the measure $\nu^g_{n,\phi,k}$ (Proposition 4.1) in the context of Poisson samples. Even though the measures in question share neither the same representation nor the same scaling properties as those considered in previous work ([5], [35]-[39]), once we have shown the crucial variance convergence for measures defined in terms of Poisson samples we can draw upon some well established dependency graph techniques ([35], [36], [39]) to deduce a Poissonized version of Theorem 2.1. Further, as a direct application of the arguments in Penrose [36], we may easily de-Poissonize and deduce Theorem 2.1 when ϕ is bounded on (0,1]. Deducing Theorem 2.1 for general ϕ requires considerable extra technical effort (see Section 5).

Recall that for all $a > 0$, $\mathcal{H}_a$ is a homogeneous Poisson point process of intensity a on $\mathbb{R}^d$. Suppose we fix the set $A \subset \mathbb{R}^d$, the densities f and g and their corresponding distributions F and G on $\mathbb{R}^d$, as described in Section 2.1.

For all $\lambda > 0$, $x \in \mathbb{R}^d$, and all finite $\mathcal{X} \subset \mathbb{R}^d$, we lighten the notation and *write* $C(x, \mathcal{X})$ *for* $C_k^{\mathcal{K}}(x, \mathcal{X})$ given by (2.2). For all Borel $B \subset \mathbb{R}^d$, define the numbers $\Phi_\lambda(x, \mathcal{X}) = \Phi^g_\lambda(x, \mathcal{X})$ and $\xi_\lambda(x, \mathcal{X}, B) := \xi^g_\lambda(x, \mathcal{X}, B)$ by

$$\Phi_\lambda(x, \mathcal{X}) := \phi(\lambda G(C(x, \mathcal{X})); \quad \xi_\lambda(x, \mathcal{X}, B) := \Phi_\lambda(x, \mathcal{X})\delta_x(B).$$

Clearly we have

$$\nu^g_{n,\phi,k}(B) = \sum_{i=1}^n \xi^g_n(X_i, \mathcal{X}_n, B)$$

where we recall $\mathcal{X}_n := \{X_1, ..., X_n\}$. Then $\xi_\lambda(x, \mathcal{X}, \cdot)$ is a (signed) point measure on $\mathbb{R}^d$ determined by x and $\mathcal{X}$, in a similar manner to the general case considered in [36]. Unfortunately, unlike in [36] it is not the case here that for each λ the measure $\xi_\lambda(x, \mathcal{X}, \cdot)$ is obtained by scaling the measure ξ_1 (because the function g enters in a more complicated way into the definition of ξ_λ here) so we cannot directly apply results from [36]. We write $\langle h, \xi_\lambda(x, \mathcal{X})\rangle$ for $\int_{\mathbb{R}^d} h(y)\xi_\lambda(x, \mathcal{X}, dy)$.

We note some immediate consequences of the assumptions set out in Section 2.1. The assumptions on f and g imply that there is a finite constant K_4 such that for any Borel $B \subseteq \mathbb{R}^d$,

$$F(B) \leq K_4 G(B); \quad G(B) \leq K_4 F(B). \tag{4.1}$$

Also, if $\phi \in \mathcal{F}$ then it is not hard to show that the dominating function ϕ^* defined by (2.1) satisfies $\log(\max(\phi^*(t), 1)) = o(t)$ as $t \to \infty$, and thence to deduce that for all $\beta > 0$ and all $\delta > 0$ we have

$$\mathbb{E}\,[\phi^*(\beta \Gamma_k)^2] < \infty; \tag{4.2}$$

$$\mathbb{E}\,[\phi^*(\beta \Gamma_k)^4 \mathbf{1}\{\Gamma_k \geq \delta\}] < \infty. \tag{4.3}$$

Recall that $\mathcal{F}_0$ is the class of functions in $\mathcal{F}$ which are bounded on $(0, 1]$. As indicated, we shall first prove Theorem 2.1 when $\phi \in \mathcal{F}_0$. For $\phi \in \mathcal{F}_0$ define $\phi^{**} : \mathbb{R}^+ \to [0, \infty)$ to be the minimal nondecreasing function on $\mathbb{R}^+$ which dominates $|\phi(\cdot)|$ pointwise, i.e. set

$$\phi^{**}(t) := \sup\{|\phi(s)| : 0 < s \leq t\}.$$

Then ϕ^{**} satisfies $\log(\max(\phi^*(t), 1)) = o(t)$ as $t \to \infty$, so that for all $\beta > 0$,

$$\mathbb{E}\,[\phi^{**}(\beta \Gamma_k)^4] < \infty. \tag{4.4}$$

Let $\mathcal{S}_3$ denote the class of all finite subsets of A having at most three elements. For $\lambda > 0$, recall that $\mathcal{P}_\lambda$ denotes a Poisson process in $\mathbb{R}^d$ with intensity function λf.

Lemma 4.1 *Under our stated conditions on f and g, if $\phi \in \mathcal{F}_0$ then*

$$\sup_{\lambda \geq 1, x \in A} \mathbb{E}\,[\Phi_\lambda(x, \mathcal{P}_\lambda)^4] < \infty; \tag{4.5}$$

$$\sup_{\lambda \geq 1, x \in A, y \in A} \mathbb{E}\,[\Phi_\lambda(x, \mathcal{P}_\lambda \cup \{y\})^4] < \infty \tag{4.6}$$

and

$$\sup_{\lambda \geq 1, x \in A, \mathcal{A} \in \mathcal{S}_3} \; \sup_{(\lambda/2) \leq m \leq (3\lambda/2)} \mathbb{E}\,[\Phi_\lambda(x, \mathcal{X}_m \cup \mathcal{A})^4] < \infty. \tag{4.7}$$

Remark. For general $\phi \in \mathcal{F}$, condition (4.7) is unfortunately not satisfied. This is the main reason we need to work initially with the class $\mathcal{F}_0$. We will show Theorem 2.1 holds for $\phi \in \mathcal{F}_0$ and then use truncation in Section 5 to show that Theorem 2.1 also holds for $\phi \in \mathcal{F}$ provided either $\mathcal{K} = \mathbb{R}^d$ or $d = 1$.

Proof. Let $x \in A$. By (4.1), for $t > 0$ we have

$$P[\lambda G(C(x, \mathcal{P}_\lambda)) > t] \leq P[\lambda F(C(x, \mathcal{P}_\lambda)) > t/K_4]. \tag{4.8}$$

Put $r(x,u) := \inf\{r : F(B_r^{\mathcal{K}}(x)) > u\}$, with the infimum of the empty set taken to be ∞. If $r(x, t/(K_4\lambda))$ is infinite, then the left-hand side of (4.8) is zero. If not, then the event in the right-hand side of (4.8) occurs iff $\mathrm{card}(\mathcal{P}_\lambda \cap B^{\mathcal{K}}_{r(x,t/(K_4\lambda))}(x)) < k$.

We thus have

$$P[\lambda G(C(x, \mathcal{P}_\lambda)) > t] \leq \exp\left(-\frac{t}{K_4}\right) \sum_{j=0}^{k-1} \frac{1}{j!} \left(\frac{t}{K_4}\right)^j$$

whence we obtain the stochastic domination relations

$$\lambda G(C(x, \mathcal{P}_\lambda \cup \{y\})) \prec \lambda G(C(x, \mathcal{P}_\lambda)) \prec K_4 \Gamma_k.$$

Hence by (4.4)

$$\mathbb{E}\,[\Phi_\lambda(x, \mathcal{P}_\lambda)^4] \leq \mathbb{E}\,[\phi^{**}(\lambda G(C(x, \mathcal{P}_\lambda))^4] \leq \mathbb{E}\,[\phi^{**}(K_4 \Gamma_k)^4] < \infty$$

yielding (4.5); (4.6) follows similarly. Also, by an argument similar to that above, for any $t > 0$, $m \geq k$, $G(C(x, \mathcal{X}_m \cup \mathcal{A})) \geq t/\lambda$ implies that there are at most $k-1$ points of $\mathcal{X}_m$ in $B^{\mathcal{K}}_{r(x,t/(\lambda K_4))}(x)$. So for $x \in A$ and $\mathcal{A} \in \mathcal{S}_3$,

$$P[\lambda G(C(x, \mathcal{X}_m \cup \mathcal{A})) \geq t] \leq \sum_{j=0}^{k-1} \binom{m}{j} \left(1 - \frac{t}{K_4 \lambda}\right)^{m-j} \left(\frac{t}{K_4\lambda}\right)^j.$$

For $2k \leq \lambda/2 \leq m \leq 3\lambda/2$ and $m/2 \leq m - j$ we have $\lambda^{-j}\binom{m}{j} \leq \frac{m^j}{j!\lambda^j} \leq \frac{1}{j!}(\frac{3}{2})^j$, and so

$$P[\lambda G(C(x, \mathcal{X}_m \cup \mathcal{A})) \geq t] \leq \sum_{j=0}^{k-1} \frac{1}{j!} \exp\left(-\frac{t}{4K_4}\right) \left(\frac{3t}{2K_4}\right)^j \leq 6^k P[4K_4\Gamma_k > t].$$

Thus

$$\mathbb{E}\left[\Phi_\lambda(x, \mathcal{X}_m \cup \mathcal{A})^4\right] \le \mathbb{E}\left[\phi^{**}(\lambda G(C(x, \mathcal{X}_m \cup \mathcal{A}))^4\right] \le 6^k \mathbb{E}\left[\phi^{**}(4K_4\Gamma_k)^4\right] < \infty.$$

This yields (4.7) for $\lambda \in [4k, \infty)$. For $\lambda \in [1, 4k)$, we have $\mathbb{E}\left[\Phi_\lambda(x, \mathcal{X}_m \cup \mathcal{A})^4\right] \le \mathbb{E}\left[\phi^{**}(\lambda G(C(x, \mathcal{X}_m \cup \mathcal{A})))^4\right] \le (\phi^{**}(4k))^4$ since ϕ^{**} is non-decreasing and since $G(B) \le 1$ for all sets B. Thus (4.7) holds for all $\lambda \ge 1$. □

For locally finite $\mathcal{X} \subset \mathbb{R}^d$ and $x \in A$ we define

$$\xi_\infty^{g,x}(\mathcal{X}) := \xi_\infty^{g,x,k}(\mathcal{X}) := \phi\left(g(x)|C(\mathbf{0}, \mathcal{X})|\right). \tag{4.9}$$

Our next result illustrates one of the central ideas of this paper, namely that *the local behavior of Φ_λ near $x \in A$ is approximated by the behavior of the limit functional $\xi_\infty^{g,x}$ on homogeneous Poisson point sets on the (infinite) space $\mathbb{R}^d$.* This result, combined with the localization (stabilization) property of the upcoming Lemma 4.3, lies at the heart of our approach, facilitating convergence of the pair correlation function and thus the variance convergence of Proposition 4.1.

Recall that $x \in \mathbb{R}^d$ is a *Lebesgue point* of f if $r^{-d}\int_{B_r(x)} |f(y) - f(x)|dy \to 0$ as $r \downarrow 0$, and that almost all $x \in \mathbb{R}^d$ are Lebesgue points of f.

Lemma 4.2 *Suppose $x \in A$ is a Lebesgue point both of f and of g. Then for all $z \in \mathbb{R}^d$, as $n \to \infty$ we have*

• $n \to \lambda$. *JY*

$$\Phi_\lambda(x + \lambda^{-1/d}z, \mathcal{P}_\lambda) \xrightarrow{\mathcal{D}} \xi_\infty^{g,x}(\mathcal{H}_{f(x)}) \tag{4.10}$$

and

$$\begin{aligned}\Phi_\lambda(x, \mathcal{P}_\lambda \cup \{x + \lambda^{-1/d}z\})\Phi_\lambda(x + \lambda^{-1/d}z, \mathcal{P}_\lambda \cup \{x\}) \\ \xrightarrow{\mathcal{D}} \xi_\infty^{g,x}(\mathcal{H}_{f(x)} \cup z)\xi_\infty^{g,x}(-z + (\mathcal{H}_{f(x)} \cup \mathbf{0})).\end{aligned} \tag{4.11}$$

Proof. Let $v_\lambda := x + \lambda^{-1/d}z$. By Lemma 3.2 of [35], it is the case that using the metric on point sets introduced in [35], as $\lambda \to \infty$ we have the weak convergence

$$\lambda^{1/d}(-v_\lambda + \mathcal{P}_\lambda) \xrightarrow{\mathcal{D}} \mathcal{H}_{f(x)}; \tag{4.12}$$

$$\lambda^{1/d}(-x + \mathcal{P}_\lambda) \xrightarrow{\mathcal{D}} \mathcal{H}_{f(x)}. \tag{4.13}$$

By (4.12) and the Continuous Mapping Theorem ([6], Chapter 1, Theorem 5.1),

$$\lambda|C(v_\lambda,\mathcal{P}_\lambda)| = |C(\mathbf{0},\lambda^{1/d}(-v_\lambda+\mathcal{P}_\lambda))| \xrightarrow{\mathcal{D}} |C(\mathbf{0},\mathcal{H}_{f(x)})|. \tag{4.14}$$

Also, using (4.12) and the assumption that x is a Lebesgue point of g, we have

$$\lambda \int_{C(v_\lambda,\mathcal{P}_\lambda)} (g(y)-g(x))dy \xrightarrow{P} 0. \tag{4.15}$$

Combining (4.14) with (4.15) and using Slutsky's theorem we obtain

$$\lambda \int_{C(v_\lambda,\mathcal{P}_\lambda)} g(y)dy \xrightarrow{\mathcal{D}} g(x)|C(\mathbf{0},\mathcal{H}_{f(x)})|$$

so that the Continuous Mapping Theorem gives us (4.10).

For (4.11), note that by (4.13) and the Continuous Mapping Theorem we have

$$\begin{aligned} &(\lambda|C(x,\mathcal{P}_\lambda\cup v_\lambda)|,\lambda|C(v_\lambda,\mathcal{P}_\lambda\cup x)|) \\ &= (|C(\mathbf{0},(\lambda^{1/d}(-x+\mathcal{P}_\lambda))\cup z)|,|C(z,(\lambda^{1/d}(-x+\mathcal{P}_\lambda))\cup\mathbf{0})|) \\ &\xrightarrow{\mathcal{D}} (|C(\mathbf{0},\mathcal{H}_{f(x)}\cup z)|,|C(z,\mathcal{H}_{f(x)}\cup\mathbf{0})|). \end{aligned} \tag{4.16}$$

Also, using (4.12), (4.13) and the assumption that x is a Lebesgue point of g, we have

$$\lambda \int_{C(x,\mathcal{P}_\lambda\cup v_\lambda)} (g(y)-g(x))dy \xrightarrow{P} 0; \qquad \lambda \int_{C(v_\lambda,\mathcal{P}_\lambda\cup x)} (g(y)-g(x))dy \xrightarrow{P} 0,$$

and combined with (4.16) and a 2-dimensional version of Slutsky's theorem, this gives us

$$\begin{aligned} &\left(\lambda \int_{C(x,\mathcal{P}_\lambda\cup v_\lambda)} g(y)dy, \lambda \int_{C(v_\lambda,\mathcal{P}_\lambda\cup x)} g(y)dy\right) \\ &\xrightarrow{\mathcal{D}} (g(x)|C(\mathbf{0},\mathcal{H}_{f(x)}\cup z)|, g(x)|C(z,\mathcal{H}_{f(x)}\cup\mathbf{0})|). \end{aligned}$$

The Continuous Mapping Theorem then gives (4.11). □

For all $x \in A$ and given $k \in \mathbb{N}$, let $t_0(x)$ denote the infimum of all t with the property that $B_u^{\mathcal{K}}(x)\cap A$ is the same for all $u \geq t$. Given also a locally finite set $\mathcal{X} \subset A$ and $\mathcal{K}$, and writing $\#(\cdot)$ for card$(\cdot)$ and $\mathcal{X}\setminus x$ for $\mathcal{X}\setminus\{x\}$ here, define

$$R_1(x,\mathcal{X}) := \begin{cases} \inf\{t\in\mathbb{R}^+ : \#(B_t^{\mathcal{K}}(x)\cap\mathcal{X}\setminus x)\geq k\} & \text{if } \#((x+\mathcal{K})\cap\mathcal{X}\setminus x)\geq k \\ t_0(x) & \text{otherwise} \end{cases}$$

Thus $R_1(x, \mathcal{X})$ is the distance between x and its kth nearest neighbor in $\mathcal{X}$ in the direction of the cone $\mathcal{K}$ or if no such neighbor exists, the furthest one has to look from x to ascertain that this is the case. For $\lambda > 0$ let $R_\lambda(x, \mathcal{X}) := \lambda^{1/d} R_1(x, \mathcal{X})$. The following lemma establishes the equivalent of the 'exponential stabilization' conditions discussed in [36].

Lemma 4.3 *It is the case that*

$$\limsup_{t\to\infty} \sup_{x\in A, \lambda\geq 1} t^{-1} \log P[R_\lambda(x, \mathcal{P}_\lambda) > t] < 0 \tag{4.17}$$

and

$$\limsup_{t\to\infty} \sup_{x\in A, \lambda\geq 1, (\lambda/2)\leq n\leq (3\lambda/2), \mathcal{A}\in\mathcal{S}_3} t^{-1} \log P[R_\lambda(x, \mathcal{X}_n \cup \mathcal{A}) > t] < 0. \tag{4.18}$$

Proof. By the assumptions that f is bounded away from zero on A, and that the sets $B_r^{\mathcal{K}}(x) \cap A$ have bounded aspect ratio uniformly over $x \in A$ and $r > 0$,

$$\inf\{r^{-d} F(B_r^{\mathcal{K}}(x)) : x \in A, r \in (0, t_0(x))\} > 0. \tag{4.19}$$

Hence, there is a finite constant $K \geq 1$ such that if $\lambda^{-1/d} t < t_0(x)$, then

$$\begin{aligned} P[R_\lambda(x, \mathcal{P}_\lambda) > t] = P[R_1(x, \mathcal{P}_\lambda) > \lambda^{-1/d} t] &= P[\mathrm{card}(\mathcal{P}_\lambda \cap B^{\mathcal{K}}_{\lambda^{-1/d} t}(x)) < k] \\ &\leq \exp\left(-\frac{t^d}{K}\right) \sum_{j=0}^{k-1} \frac{1}{j!} \left(\frac{t^d}{K}\right)^j, \end{aligned} \tag{4.20}$$

and for $\mathcal{A} \in \mathcal{S}_3$,

$$\begin{aligned} P[R_\lambda(x, \mathcal{X}_n \cup \mathcal{A}) > t] &\leq P[\mathrm{card}(\mathcal{X}_n \cap B^{\mathcal{K}}_{t/\lambda^{1/d}}(x)) < k] \\ &\leq \sum_{j=0}^{k-1} \binom{n}{j} \left(1 - \frac{t^d}{K\lambda}\right)^{n-j} \left(\frac{t^d}{K\lambda}\right)^j. \end{aligned}$$

For $\lambda/2 \leq n \leq 3\lambda/2$ we have $\lambda^{-j}\binom{n}{j} \leq \frac{1}{j!}(\frac{3}{2})^j$ and so for $1 < t < \lambda^{-1/d} t_0(x)$,

$$P[R_\lambda(x, \mathcal{X}_n \cup \mathcal{A}) > t] \leq \sum_{j=0}^{k-1} \frac{1}{j!} \left(\frac{3t^d}{2}\right)^j \left(1 - \frac{t^d}{K\lambda}\right)^{n-j} \leq k \left(\frac{3t^d}{2}\right)^k \left(1 - \frac{t^d}{K\lambda}\right)^{n-k} \tag{4.21}$$

If $\lambda^{-1/d} t \geq t_0(x)$ then the bounds (4.20) and (4.21) still hold since then the probabilities in question are zero. Equations (4.20) and (4.21) imply (4.17) and (4.18) respectively. □

Recall (2.15) that $\mu^g_{\lambda,\phi,k}$ denotes the Poissonized version of $\nu^g_{\lambda,\phi,k}$. The following is a Poissonized version of Theorem 2.1 for $\phi \in \mathcal{F}_0$ and is of independent interest.

Proposition 4.1 *Let $\phi \in \mathcal{F}_0$ and $h \in \mathcal{B}(A)$. Then as $\lambda \to \infty$*

$$\lambda^{-1}\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi,k}\rangle] \to \int_A h^2(x) V_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx = I_{V_{\phi,k}}(g,f,h^2) \tag{4.22}$$

and $\lambda^{-1/2}\overline{\mu}^g_{\lambda,\phi,k}$ converges in law as $\lambda \to \infty$ to a mean zero Gaussian field with covariance kernel $(h_1,h_2) \mapsto I_{V_{\phi,k}}(g,f,h_1) I_{V_{\phi,k}}(g,f,h_2)$.

Proof. For simplicity we first assume that h is a.e. continuous. It is the case that

$$\lambda^{-1}\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi,k}\rangle] = \lambda \int_A \int_A h(x)h(y)\{\mathbb{E}\left[\Phi_\lambda(x, \mathcal{P}_\lambda \cup y)\Phi_\lambda(y, \mathcal{P}_\lambda \cup x)\right]$$
$$-\mathbb{E}\left[\Phi_\lambda(x,\mathcal{P}_\lambda)\right]\mathbb{E}\left[\Phi_\lambda(y,\mathcal{P}_\lambda)\right]\} f(x)f(y)dxdy + \int_A h^2(x)\mathbb{E}\left[\Phi^2_\lambda(x,\mathcal{P}_\lambda)\right]f(x)dx. \tag{4.23}$$

We will use Lemmas 4.1 -4.3 to show that $\lambda^{-1}\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi,k}\rangle]$ converges to

$$\int_A \int_{\mathbb{R}^d} h^2(x)[\mathbb{E}\, \xi^{g;x}_\infty(\mathcal{H}_{f(x)} \cup z)\xi^{g;x}_\infty(-z + (\mathcal{H}_{f(x)} \cup \mathbf{0})) - (\mathbb{E}\, \xi^{g;x}_\infty(\mathcal{H}_{f(x)}))^2] f^2(x)dzdx$$
$$+ \int_A h^2(x)\mathbb{E}\left[(\xi^{g;x}_\infty(\mathcal{H}_{f(x)})^2\right] f(x)dx. \tag{4.24}$$

Putting $y = x + \lambda^{-1/d}z$ in the right-hand side in (4.23) reduces the double integral to

$$= \int_A \int_{-\lambda^{1/d}x + \lambda^{1/d}A} h(x)h(x+\lambda^{-1/d}z)\{...\}f(x)f(x+\lambda^{-1/d}z)dzdx \tag{4.25}$$

where

$$\{...\} := \{\mathbb{E}\left[\Phi_\lambda(x, \mathcal{P}_\lambda \cup \{x+\lambda^{-1/d}z\})\Phi_\lambda(x+\lambda^{-1/d}z, \mathcal{P}_\lambda \cup x)\right]$$
$$-\mathbb{E}\left[\Phi_\lambda(x,\mathcal{P}_\lambda)\right] E[\Phi_\lambda(x+\lambda^{-1/d}z, \mathcal{P}_\lambda)]\}$$

is the two point correlation function for Φ_λ. By Lemmas 4.1 and 4.2, it follows for all $x \in A$ and all $z \in \mathbb{R}^d$, that the two point correlation function for Φ_λ converges to the two point correlation function for $\xi^{g;x}_\infty$, i.e., the bracketed expression in the first term of (4.24). Moreover, the integrand in (4.25) is dominated by an integrable function of

z over $\mathbb{R}^d$ (see Lemma 4.2 of [36]; here we are using Lemma 4.3). The convergence of the double integral in (4.23) to that in (4.24) now follows by dominated convergence, the continuity of h, and the moment bounds of Lemma 4.1. To show convergence of general $h \in \mathcal{B}(A)$ we refer to [36].

To complete the proof that (4.23) converges to (4.24) we need only to show convergence of $\int_A h^2(x)\mathbb{E}\,[\Phi_\lambda^2(x, \mathcal{P}_\lambda)] f(x)dx$. This is a simple consequence of the convergence (4.10), the moment bounds (4.5), and dominated convergence.

For all $x \in A$ we define $V_{\phi,k}^{\xi}(x, 0) := 0$ and for all $a > 0$ we put

$$V_{\phi,k}^{\xi}(x,a) := \mathbb{E}\,[\xi_\infty^{g,x}(\mathcal{H}_a)^2] + a\int_{\mathbb{R}^d} \left[\mathbb{E}\,\xi_\infty^{g,x}(\mathcal{H}_a \cup z)\xi_\infty^{g,x}(-z + (\mathcal{H}_a \cup \mathbf{0})) - (\mathbb{E}\,\xi_\infty^{g,x}(\mathcal{H}_a))^2\right] dz.$$

Using (4.9), it is easy to see that

$$V_{\phi,k}^{\xi}(x,a) = \mathbb{E}\,[\phi(\frac{g(x)}{a}\Gamma_k)^2]$$

$$+ \int_{\mathbb{R}^d} \left[\mathbb{E}\,\left[\phi(\frac{g(x)}{a}|C(\mathbf{0}, \mathcal{H} \cup y)|)\phi(\frac{g(x)}{a}|C(y, \mathcal{H} \cup \mathbf{0})|)\right] - (\mathbb{E}\,[\phi(\frac{g(x)}{a}\Gamma_k)])^2\right] dy \quad (4.26)$$

and in particular, by definition of $V_{\phi,k}$ (recall (2.9)) we have

$$V_{\phi,k}^{\xi}(x, f(x)) = V_{\phi,k}\left(\frac{g(x)}{f(x)}\right).$$

By combining this with (4.24) we thus obtain the desired limiting variance (4.22).

The proof of the second part of Proposition 4.1 (i.e., convergence to the normal) follows from Lemmas 4.1-4.3 and that of Theorem 2.2 of [36], which itself follows dependency graph arguments in [39]. By Lemma 4.2, $\xi_\infty^{g,x}(\mathcal{H}_{f(x)})$ corresponds in our setting to the limiting expression from Lemma 3.4 of [36], and consequently appears in expressions for limiting variances arising from following the proofs in [36], where all expressions for limits are obtained through Lemmas 3.4 and 3.5 of [36]. □

To obtain Theorem 2.1 we will need to de-Poissonize Proposition 4.1 and then extend the result to all $\phi \in \mathcal{F}$. The former is achieved by suitably adapting the proofs of Theorems 2.1, 2.2 and 2.3 of [36] whereas the latter is achieved via truncation arguments.

The next result plays the role of Lemma 3.6 of [36], which is used in de-Poissonizing the limit theorems. We write $\tilde{\mathcal{H}}_\lambda$ for an independent copy of the homogeneous Poisson process $\mathcal{H}_\lambda$.

• *need to define $\mathcal{X}_\ell^y$, remove ξ_λ, etc. JY*

Lemma 4.4 *Let $(x, y) \in A^2$ with $x \neq y$ and x, y both Lebesgue points of both f and g. Let $(z, w) \in (\mathbb{R}^d)^2$. Given integer-valued functions $(\ell(\lambda), \lambda \geq 1)$ and $(m(\lambda), \lambda \geq 1)$ with $\ell(\lambda) \sim \lambda$ and $m(\lambda) \sim \lambda$ as $\lambda \to \infty$, we have convergence in joint distribution, as $\lambda \to \infty$, of the 11-dimensional random vector*

$$\Big(\Phi_\lambda(x; \mathcal{X}_\ell), \Phi_\lambda(x; \mathcal{X}_\ell^y), \Phi_\lambda(\xi_\lambda(x; \mathcal{X}_\ell^{x+\lambda^{-1/d}z}), \Phi_\lambda(\xi_\lambda(x; \mathcal{X}_\ell^{x+\lambda^{-1/d}z} \cup \{y\}),$$
$$\Phi_\lambda(x; \mathcal{X}_m), \Phi_\lambda(x; \mathcal{X}_m^y), \Phi_\lambda(x; \mathcal{X}_m^y \cup \{x + \lambda^{-1/d}z\}), \Phi_\lambda(y; \mathcal{X}_m),$$
$$\Phi_\lambda(y; \mathcal{X}_m^x), \Phi_\lambda(y; \mathcal{X}_m^x \cup \{x + \lambda^{-1/d}z\}),$$
$$\Phi_\lambda(y; \mathcal{X}_m^x \cup \{x + \lambda^{-1/d}z, y + \lambda^{-1/d}w\})\Big)$$

to

$$\Big(g(x)|C(\mathbf{0}, \mathcal{H}_{f(x)})|, g(x)|C(\mathbf{0}, \mathcal{H}_{f(x)})|, g(x)|C(\mathbf{0}, \mathcal{H}^z_{f(x)})|, g(x)|C(\mathbf{0}, \mathcal{H}^z_{f(x)})|,$$
$$g(x)|C(\mathbf{0}, \mathcal{H}_{f(x)})|, g(x)|C(\mathbf{0}, \mathcal{H}_{f(x)})|, f(x)|C(\mathbf{0}, \mathcal{H}^z_{f(x)})|, g(y)|C(\mathbf{0}, \tilde{\mathcal{H}}_{f(y)})|,$$
$$g(y)|C(\mathbf{0}, \tilde{\mathcal{H}}_{f(y)})|, g(y)|C(\mathbf{0}, \tilde{\mathcal{H}}_{f(y)})|, g(y)|C(\mathbf{0}, \tilde{\mathcal{H}}^w_{f(y)})\Big).$$

Proof. We may follow the proof of Lemma 3.6 of [36], since Lemma 3.2 of [36] remains valid in our setting. The 'Binomial exponential stabilization' condition of [36] holds by (4.18). □

• $f(x) \to g(x)$

Proposition 4.2 *Suppose $\phi \in \mathcal{F}_0$. Then the conclusions of Theorem 2.1 hold.*

Proof. Taking Proposition 4.1 as our starting point, we can follow the de-Poissonization argument of Section 5 of [36] which is used there to prove Theorem 2.3 of [36]. The argument carries through verbatim; the condition A5′ in [36] follows from Lemma 4.1 and (4.18).

To obtain the limiting variance in the present setting, again we compare the limiting expressions in Lemma 4.2 with the corresponding limits obtained in Lemmas 3.4 and 3.5 of [36]. That is, for all $x \in A$ and all $a > 0$ we define

$$\Delta^\xi_{\phi,k}(x, a) := \mathbb{E}\,[\xi^{g,x}_\infty(\mathcal{H}_a)] + a \int_{\mathbb{R}^d} [\mathbb{E}\,\xi^{g,x}_\infty(\mathcal{H}_a \cup y) - \xi^{g,x}_\infty(\mathcal{H}_a)]\, dy,$$

analogously to the definition of $\delta(x, a)$ in [36]. It is easy to see that

$$\Delta^\xi_{\phi,k}(x, a) = \mathbb{E}\,[\phi(\frac{g(x)}{a}\Gamma_k)] + \ a \int_{\mathbb{R}^d} \mathbb{E}\,[\phi(g(x)|C(\mathbf{0}, \mathcal{H}_a \cup y)|) - \phi(g(x)|C(\mathbf{0}, \mathcal{H}_a)|)] dy.$$

By a change of variables $y \to a^{1/d}y$, the equivalence $a^{1/d}\mathcal{H}_a \stackrel{\mathcal{D}}{=} \mathcal{H}$, and (2.8) we obtain

$$\Delta^{\xi}_{\phi,k}(x,a) = M_{\phi,k}\left(\frac{g(x)}{a}\right) + \int_{\mathbb{R}^d} \mathbb{E}\,[\phi(\frac{g(x)}{a}|C(\mathbf{0},\mathcal{H}\cup y)|) - \phi(\frac{g(x)}{a}|C(\mathbf{0},\mathcal{H})|)]dy. \tag{4.27}$$

We now show that $\Delta^{\xi}_{\phi,k}(x,a)$ reduces to $\Delta_{\phi,k}(g(x)/a)$ defined by (2.8). Put $\beta := g(x)/a$ and $b_d := |B^{\mathcal{K}}_1(\mathbf{0})|$. Since $|C(\mathbf{0},\mathcal{H})|$ has the same distribution as Γ_k, (4.27) yields

$$\begin{aligned}\Delta^{\xi}_{\phi,k}(x,a) - M_{\phi,k}(\beta) &= \int_{\mathcal{K}} \mathbb{E}\,\left[(\phi(\beta|C(\mathbf{0},\mathcal{H}\cup y)|) - \phi(\beta|C(\mathbf{0},\mathcal{H})|))\mathbf{1}\{b_d|y|^d \le \Gamma_k\}\right] dy \\ &= \int_{\mathcal{K}} \mathbb{E}\,\left[(\phi(\beta\max(b_d|y|^d,\Gamma_{k-1})) - \phi(\beta\Gamma_k))\mathbf{1}\{b_d|y|^d \le \Gamma_k\}\right] dy.\end{aligned}$$

Putting $s := |B^{\mathcal{K}}_{|y|}(\mathbf{0})|$ shows that

$$\begin{aligned}\Delta^{\xi}_{\phi,k}(x,a) - M_{\phi,k}(\beta) = \mathbb{E}\int_0^{\Gamma_{k-1}} \phi(\beta\Gamma_{k-1})ds \\ +\mathbb{E}\int_{\Gamma_{k-1}}^{\Gamma_k} \phi(\beta s)ds - \mathbb{E}\,[\Gamma_k\phi(\beta\Gamma_k)].\end{aligned} \tag{4.28}$$

The third term in the right hand side of (4.28) is

$$\mathbb{E}\,[\Gamma_k\phi(\beta\Gamma_k)] = \int_0^\infty s\phi(\beta s)\frac{s^{k-1}}{(k-1)!}e^{-s}ds = k\mathbb{E}\,[\phi(\beta\Gamma_{k+1})] \tag{4.29}$$

and likewise, the first term is $(k-1)\mathbb{E}\,\phi(\beta\Gamma_k)$. Recalling that $\Gamma_k = \sum_{i=1}^k \Gamma_{1,i}$ and setting $t = s - \Gamma_{k-1}$, we find that the middle term in the right hand side of (4.28) is

$$\begin{aligned}\mathbb{E}\int_0^\infty \phi(\beta(\Gamma_{k-1}+t))\mathbf{1}_{\{t\le\Gamma_{1,k}\}}dt &= \mathbb{E}\int_0^\infty \phi(\beta(\Gamma_{k-1}+t))e^{-t}dt \\ &= \mathbb{E}\,[\phi(\beta\Gamma_k))] = M_{\phi,k}(\beta).\end{aligned}$$

Combining these expressions for terms in the right side of (4.28) yields

$$\Delta^{\xi}_{\phi,k}(x,a) = (k+1)M_{\phi,k}(\beta) - kM_{\phi,k+1}(\beta) := \Delta_{\phi,k}(\beta).$$

It follows from the proof of Theorem 2.3 of [36] that the finite-dimensional distributions of the random field $\left(n^{-1/2}\langle h, \overline{\nu}^g_{n,\phi}\rangle,\ h \in \mathcal{B}(A)\right)$, converge weakly as $n \to \infty$ to those of a mean zero finitely additive Gaussian field with covariance given by (2.12). This concludes the proof of Proposition 4.2. □

5 Proof of Theorem 2.1 for general ϕ

Section 4 establishes Theorem 2.1 for $\phi \in \mathcal{F}_0$. To prove Theorem 2.1 for $\phi \in \mathcal{F}$ when either $\mathcal{K} = \mathbb{R}^d$ or $d = 1$, we will use the Efron-Stein inequality [19], which we now recall. Suppose $S(x_1, \ldots, x_n)$ is a symmetric function of n vectors $x_i \in \mathbb{R}^d$. Define the random variables

$$S_i := S(X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_{n+1}), \quad 1 \leq i \leq n+1.$$

The Efron-Stein inequality says that for any random variable Z we have

$$\mathrm{Var}[S(X_1, \ldots, X_n)] \leq \mathbb{E} \sum_{i=1}^{n+1} (Z - S_i)^2.$$

We also recall the notation $\|X\|_p := \mathbb{E}\,[|X|^p]^{1/p}$ for any random variable X and any $p \geq 1$.

We shall extend Theorem 2.1 to cases with $\phi \in \mathcal{F} \setminus \mathcal{F}_0$ (i.e., where ϕ 'blows up' at 0) via a truncation argument. Given $\varepsilon > 0$, define the functions $\phi^\varepsilon : \mathbb{R}^+ \to \mathbb{R}$ and $\phi_\varepsilon : \mathbb{R}^+ \to \mathbb{R}$ by

$$\phi^\varepsilon(x) := \begin{cases} \phi(x) & \text{if } x \geq \varepsilon \\ 0 & \text{otherwise,} \end{cases} \qquad \phi_\varepsilon(x) := \begin{cases} \phi(x) & \text{if } x < \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 5.1 *Given $h \in \mathcal{B}(\mathbb{R}^d)$ and $\delta > 0$, there exists $\varepsilon_0 > 0$ and $n_0 > 0$ such that for $\varepsilon \in (0, \varepsilon_0)$ and $n \geq n_0$ we have $n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi_\varepsilon,k} \rangle] \leq \delta$.*

Proof. We first assume $\mathcal{K} = \mathbb{R}^d$. Recall that we write $C(x, \mathcal{X})$ for $C_k^{\mathcal{K}}(x, \mathcal{X})$. Applying the Efron-Stein inequality with

$$S(x_1, \ldots, x_n) := \sum_{i=1}^n h(x_i)\phi_\varepsilon(nG(C(x_i, \{x_1, \ldots, x_n\})))$$

and $Z := \sum_{i=1}^{n+1} h(X_i)\phi_\varepsilon(nD^g_{i,n+1,k})$, we obtain the bound

$$\begin{aligned} \mathrm{Var}[\langle h, \nu^g_{n,\phi_\varepsilon,k} \rangle] &= \mathrm{Var}\left[\sum_{i=1}^n h(X_i)\phi_\varepsilon(nD^g_{i,n,k})\right] \\ &\leq (n+1)\mathbb{E}\left[\left(\sum_{i=1}^{n+1} h(X_i)\phi(nD^g_{i,n+1,k}) - \sum_{i=1}^n h(X_i)\phi_\varepsilon(nD^g_{i,n,k})\right)^2\right] \\ &= (n+1)\mathbb{E}\left[\left(h(X_{n+1})\phi_\varepsilon(nD^g_{n+1,n+1,k}) + \sum_{i=1}^n h(X_i)\left(\phi_\varepsilon(nD^g_{i,n+1,k}) - \phi_\varepsilon(nD^g_{i,n,k})\right)\right)^2\right]. \end{aligned} \tag{5.1}$$

Write X for X_{n+1}. By (4.1), for $t \in [0,1]$, and $n \geq 2K_4$, we have

$$P[nD^g_{n+1,n+1,k} \geq t] = P[nG(C(X,\mathcal{X}_n)) \geq t] \geq P[nF(C(X,\mathcal{X}_n)) \geq tK_4]$$
$$= \left(1 - \frac{tK_4}{n}\right)^n \geq e^{-tK_5} = P[K_5^{-1}\Gamma_1 \geq t] \tag{5.2}$$

for some constant $K_5 > 0$. Hence

$$\min(nD^g_{n+1,n+1,k}, 1) \succ \min(K_5^{-1}\Gamma_1, 1). \tag{5.3}$$

Let the dominating function ϕ^* be defined by (2.1). Let $\varepsilon \in (0,1]$ and define the function ϕ^*_ε by

$$\phi^*_\varepsilon(x) := \begin{cases} \phi^*(x) & \text{if } x \leq \varepsilon \\ 0 & \text{otherwise.} \end{cases}$$

Then $\phi^*_\varepsilon(x)$ is a non-increasing function of $\min(x,1)$, so by (5.3),

$$\mathbb{E}\,[\phi_\varepsilon(nD^g_{n+1,n+1,k})^2] \leq \mathbb{E}\,[\phi^*_\varepsilon(nD^g_{n+1,n+1,k})^2] \leq \mathbb{E}\,[\phi^*_\varepsilon(K_5^{-1}\Gamma_1)^2]. \tag{5.4}$$

Next, let $W_1^0, \ldots, W_J^0$ be a finite collection of open cones in $\mathbb{R}^d$, each with vertex at the origin and angular radius $\pi/6$, such that $\mathbb{R}^d \setminus \{\mathbf{0}\} \subseteq \bigcup_{j=1}^J W_j^0$. For $1 \leq j \leq J$, let W_j be the translate of W_j^0 with its vertex at X.

For $1 \leq j \leq J$, and $1 \leq l \leq k$, let $Y_{j,l,n}$ denote the lth nearest neighbor of X in the point set $\mathcal{X}_n \cap W_j$, with $Y_{j,l,n} := X$ if $\mathrm{card}(\mathcal{X}_n \cap W_j) < l$. Let $I_{j,l,n}$ be the indicator function of the event that $\mathrm{card}(\mathcal{X}_n \cap W_j) \geq l$.

For $1 \leq i \leq n$, and $1 \leq j \leq J$, if $X_i \in W_j$ and X_i is *not* among the k nearest neighbors of X in $\mathcal{X}_n \cap W_j$, then there are at least k points of $\mathcal{X}_n \cap W_j$ lying closer to X than X_i does, in which case by elementary geometry these points lie closer to X_i than X does, so that $C(X_i, \mathcal{X}_{n+1}) = C(X_i, \mathcal{X}_n)$ and $D^g_{i,n+1} = D^g_{i,n}$. Therefore

$$\left|\sum_{i=1}^n h(X_i)(\phi_\varepsilon(nD^g_{i,n+1}) - \phi_\varepsilon(nD^g_{i,n}))\right|$$
$$\leq \sum_{j=1}^J \sum_{l=1}^k I_{j,l,n}|h(Y_{j,l,n})[\phi_\varepsilon(nG(C(Y_{j,l,n},\mathcal{X}_{n+1}))) - \phi_\varepsilon(nG(C(Y_{j,l,n},\mathcal{X}_n)))]|.$$

If X is not among the k nearest neighbors of $Y_{j,l,n}$ in $\mathcal{X}_{n+1}$ then $C(Y_{j,l,n},\mathcal{X}_{n+1}) = C(Y_{j,l,n},\mathcal{X}_n)$ so that the corresponding term in the above sum is zero. On the other hand, if X is among the k nearest neighbors of $Y_{j,l,n}$ in $\mathcal{X}_{n+1}$, then

$$|\phi_\varepsilon(nG(C(Y_{j,l,n},\mathcal{X}_{n+1}))) - \phi_\varepsilon(nG(C(Y_{j,l,n},\mathcal{X}_n)))|$$
$$\leq \phi^*_\varepsilon(nG(C(Y_{j,l,n},\mathcal{X}_{n+1}))) + \phi^*_\varepsilon(nG(C(Y_{j,l,n},\mathcal{X}_n))) \leq 2\phi^*_\varepsilon(nG(C_1^{\mathcal{K}}(Y_{j,l,n},\{X\})))$$

since ϕ_ε^* is a non-increasing function. Hence,

$$\left|\sum_{i=1}^n h(X_i)(\phi_\varepsilon(nD^g_{i,n+1}) - \phi_\varepsilon(nD^g_{i,n}))\right| \leq 2\|h\|_\infty \sum_{j=1}^J \sum_{l=1}^k I_{j,l,n} \phi_\varepsilon^*(nG(C_1^{\mathcal{K}}(Y_{j,l,n}, \{X\})))$$
$$\leq 2\|h\|_\infty \sum_{j=1}^J \sum_{l=1}^k \phi_\varepsilon^*(nG(C_1^{\mathcal{K}}(Y_{j,l,n}, \{X\}))) \quad (5.5)$$

Define the variable $R_{j,l,n}$ by

$$R_{j,l,n} := \begin{cases} |X - Y_{j,l,n}| & \text{if } \mathrm{card}(\mathcal{X}_n \cap W_j) \geq l \\ +\infty \text{ otherwise.} & \end{cases}$$

Since f is assumed bounded, there is a constant K_6 such that for any $r > 0$, $x \in A$ we have $F(B_r(x)) \leq K_6 r^d$. Hence,

$$P[R_{j,l,n} > r] = P[\mathrm{card}(\mathcal{X}_n \cap B_r^{\mathcal{K}}(X) \cap W_j) < l]$$
$$\geq P[\mathrm{card}(\mathcal{X}_n \cap B_r^{\mathcal{K}}(X)) = 0] \geq (1 - K_6 r^d)^n. \quad (5.6)$$

Since we assume $\mathcal{K} = \mathbb{R}^d$ here, by our assumptions on g and A there is a constant $K_7 > 1$ such that for all $x \in A$ and $0 < r \leq 1$ we have $G(B_r^{\mathcal{K}}(x)) \geq r^d / K_7$. Then for $t \in [0, 1]$ and all $n \geq K_7$ we have

$$P[nG(C_1^{\mathcal{K}}(Y_{j,l,n}, \{X\})) > t] \geq P[nR^d_{j,l,n}/K_7 > t] = P[R_{j,l,n} > (tK_7/n)^{1/d}]$$

so that by (5.6) we have

$$P[nG(C_1^{\mathcal{K}}(Y_{j,l,n}, \{X\}) > t] \geq \exp(-tK_8) \quad (5.7)$$

for some constant K_8. Thus, for large enough n we have $\min(nG(C_1^{\mathcal{K}}(Y_{j,l,n}, \{X\})), 1) \succ \min(K_8^{-1}\Gamma_1, 1)$ and since $\phi_\varepsilon^*(x)$ is a non-increasing function of $\min(x, 1)$ we obtain for all $1 \leq j \leq J, 1 \leq l \leq k$,

$$\mathbb{E}\,[(\phi_\varepsilon^*(nG(C_1(Y_{j,l,n}, \{X\}))))^2] \leq \mathbb{E}\,[\phi_\varepsilon^*(K_8^{-1}\Gamma_1)^2]. \quad (5.8)$$

By (5.5), (5.8) and Minkowski's inequality,

$$\left\|\sum_{i=1}^n h(X_i)(\phi_\varepsilon(nD^g_{i,n+1,k}) - \phi_\varepsilon(nD^g_{i,n,k}))\right\|_2 \leq 2Jk\|h\|_\infty \|\phi_\varepsilon^*(K_8^{-1}\Gamma_1)\|_2. \quad (5.9)$$

By (5.1), (5.4) and (5.9), and Minkowski's inequality, for all large enough n we have

$$(n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi_\varepsilon,k}\rangle])^{1/2} \leq 4\|h\|_\infty (\|\phi_\varepsilon^*(K_5^{-1}\Gamma_1)\|_2 + Jk\|\phi_\varepsilon^*(K_8^{-1}\Gamma_1)\|_2). \quad (5.10)$$

Since $\phi \in \mathcal{F}$, both $\mathbb{E}\,[\phi_1^*(K_5^{-1}\Gamma_1)^2]$ and $\mathbb{E}\,[\phi_1^*(K_8^{-1}\Gamma_1)^2]$ are finite by (4.2), so the bound in (5.10) can be made small by taking ε small. This gives Lemma 5.1 when $\mathcal{K} = \mathbb{R}^d$.

When $d = 1$ and $\mathcal{K} = (0, \infty)$, then we modify the above arguments as follows. Suppose $A := [c_1, c_2]$. Then (5.1) remains unchanged. Since $C_1^{\mathcal{K}}(X, \mathcal{X}_n)$ could be exceptionally small, as would be the case if X were close to the right endpoint c_2, we alter (5.2) as follows. Let E be the event $E := \{nF([X, c_2]) \geq tK_4\}$. Then for all $t \in [0, 1]$ and all large enough $n \geq 2K_4$ we have

$$P[nF(C_1^{\mathcal{K}}(X, \mathcal{X}_n)) \geq tK_4] \geq P[nF(C_1^{\mathcal{K}}(X, \mathcal{X}_n)) \geq tK_4 | E]P[E] = \left(1 - \frac{tK_4}{n}\right)^{n+1},$$

showing that (5.4) holds for a possibly different value of K_5.

Next, for all $1 \leq l \leq k$, let $Y_{l,n}$ denote the lth nearest neighbor of X in $\mathcal{X}_n$ to the *left* of X with $Y_{l,n} := X$ if no such point exists. Then (5.5) becomes

$$\left| \sum_{i=1}^{n} h(X_i)(\phi_\varepsilon(nD^g_{i,n+1}) - \phi_\varepsilon(nD^g_{i,n})) \right| \leq 2\|h\|_\infty \sum_{l=1}^{k} \phi_\varepsilon^*(nG(C_1^{\mathcal{K}}(Y_{l,n}, \{X\}))).$$

Define the event $E' := \{nF([c_1, X]) \geq tK_4\}$. If $E \cap E'$ occurs, and if the interval with F-measure tK_4/n with X as its right endpoint contains no point of $\mathcal{X}_n$, then $nG(C_1^{\mathcal{K}}(Y_{l,n}, \{X\}) > t$. Hence, $P[nG(C_1^{\mathcal{K}}(Y_{l,n}, \{X\})) > t | E \cap E'] \geq (1 - tK_4/n)^n$, and so we obtain $P[nG(C_1^{\mathcal{K}}(Y_{l,n}, \{X\})) > t] \geq \exp(-tK_8')$, the analog of (5.7). The argument after (5.7) then carries through verbatim, so Lemma 5.1 holds when $A = [c_1, c_2]$ and $\mathcal{K} = (0, \infty)$. □

Before stating the next lemma we define for all $\beta > 0$, $y \in \mathbb{R}^d$, $\phi \in \mathcal{F}$, and $\varepsilon > 0$

$$\psi(\beta, y) := \mathbb{E}\,[\phi(\beta|C(\mathbf{0}, \mathcal{H} \cup y)|)\phi(\beta|C(y, \mathcal{H} \cup \mathbf{0})|)] - (\mathbb{E}\,\phi(\beta\Gamma_k))^2 \tag{5.11}$$

and

$$\psi_\varepsilon(\beta, y) := \mathbb{E}\,[\phi^\varepsilon(\beta|C(\mathbf{0}, \mathcal{H} \cup y)|)\phi^\varepsilon(\beta|C(y, \mathcal{H} \cup \mathbf{0})|)] - (\mathbb{E}\,\phi^\varepsilon(\beta\Gamma_k))^2. \tag{5.12}$$

We also define $a_K := \mathbb{E}\,[\phi^*(\Gamma_k/K)^2] + \mathbb{E}\,[\phi^*(K\Gamma_k)^2]$ for $K > 0$, and observe for any $K > 1$ that $a_K < \infty$. Also, if $K^{-1} \leq \beta \leq K$, then since ϕ^* is decreasing on $(0, 1)$ and increasing on $(1, \infty)$,

$$\begin{aligned} \mathbb{E}\,[\phi^*(\beta\Gamma_k)^2] &= \mathbb{E}\,[\phi^*(\beta\Gamma_k)^2\mathbf{1}\{\Gamma_k \leq 1/\beta\}] + \mathbb{E}\,[\phi^*(\beta\Gamma_k)^2\mathbf{1}\{\Gamma_k > 1/\beta\}] \\ &\leq \mathbb{E}\,[\phi^*(\Gamma_k/K)^2\mathbf{1}\{\Gamma_k \leq 1/\beta\}] + \mathbb{E}\,[\phi^*(K\Gamma_k)^2\mathbf{1}\{\Gamma_k > 1/\beta\}] \leq a_K. \end{aligned} \tag{5.13}$$

Lemma 5.2 *Let $K > 1$. Then there exists a Lebesgue integrable function $\psi_K^* : \mathbb{R}^d \to [0,\infty)$, such that*

$$|\psi_\varepsilon(\beta, y)| \leq \psi_K^*(y) \quad \forall y \in \mathbb{R}^d \setminus \{\mathbf{0}\},\ \varepsilon \in (0,1],\ \beta \in [1/K, K]. \tag{5.14}$$

Proof. Set

$$X_{\beta,\varepsilon,y} := \phi^\varepsilon(\beta|C(\mathbf{0}, \mathcal{H} \cup y)|); \quad Y_{\beta,\varepsilon,y} := \phi^\varepsilon(\beta|C(y, \mathcal{H} \cup \mathbf{0})|),$$
$$X_{\beta,\varepsilon} := \phi^\varepsilon(\beta|C(\mathbf{0}, \mathcal{H})|); \quad Y_{\beta,\varepsilon} := \phi^\varepsilon(\beta|C(y, \mathcal{H})|),$$

so that $\psi_\varepsilon(\beta, y) = \mathbb{E}\,[X_{\beta,\varepsilon,y} Y_{\beta,\varepsilon,y}] - \mathbb{E}\,[X_{\beta,\varepsilon}]\mathbb{E}\,[Y_{\beta,\varepsilon}]$. Define the events

$$E_y := \{\mathrm{card}(\mathcal{H} \cap B^{\mathcal{K}}_{|y|/2}(\mathbf{0})) \geq k\} \text{ and } \ F_y := \{\mathrm{card}(\mathcal{H} \cap B^{\mathcal{K}}_{|y|/2}(y)) \geq k\}.$$

Then $X_{\beta,\varepsilon,y}\mathbf{1}_{E_y}$ and $Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y}$ are independent and $X_{\beta,\varepsilon,y}\mathbf{1}_{E_y} = X_{\beta,\varepsilon}\mathbf{1}_{E_y}$ and $Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y} = Y_{\beta,\varepsilon}\mathbf{1}_{F_y}$, so that

$$\begin{aligned}\mathbb{E}\,[X_{\beta,\varepsilon,y} Y_{\beta,\varepsilon,y}] &= \mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y}]\mathbb{E}\,[Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y}] + \mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y} Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y^c}] + \mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y^c} Y_{\beta,\varepsilon,y}] \\ &= \mathbb{E}\,[X_{\beta,\varepsilon}\mathbf{1}_{E_y}]\mathbb{E}\,[Y_{\beta,\varepsilon}\mathbf{1}_{F_y}] + \mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y} Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y^c}] + \mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y^c} Y_{\beta,\varepsilon,y}] \end{aligned} \tag{5.15}$$

Also,

$$\mathbb{E}\,[X_{\beta,\varepsilon}]\mathbb{E}\,[Y_{\beta,\varepsilon}] = \mathbb{E}\,[X_{\beta,\varepsilon}\mathbf{1}_{E_y}]\mathbb{E}\,[Y_{\beta,\varepsilon}\mathbf{1}_{F_y}] + \mathbb{E}\,[X_{\beta,\varepsilon}\mathbf{1}_{E_y}]\mathbb{E}\,[Y_{\beta,\varepsilon}\mathbf{1}_{F_y^c}] + \mathbb{E}\,[X_{\beta,\varepsilon}\mathbf{1}_{E_y^c}]\mathbb{E}\,[Y_{\beta,\varepsilon}]$$

so by (5.15), and Hölder's inequality,

$$\begin{aligned}|\psi_\varepsilon(\beta, y)| &= |\mathbb{E}\,[X_{\beta,\varepsilon,y} Y_{\beta,\varepsilon,y}] - \mathbb{E}\,[X_{\beta,\varepsilon}]\mathbb{E}\,[Y_{\beta,\varepsilon}]| \\ &= |\mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y} Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y^c}] + \mathbb{E}\,[X_{\beta,\varepsilon,y}\mathbf{1}_{E_y^c} Y_{\beta,\varepsilon,y}] \\ &\quad -\mathbb{E}\,[X_{\beta,\varepsilon}\mathbf{1}_{E_y}]\mathbb{E}\,[Y_{\beta,\varepsilon}\mathbf{1}_{F_y^c}] - \mathbb{E}\,[X_{\beta,\varepsilon}\mathbf{1}_{E_y^c}]\mathbb{E}\,[Y_{\beta,\varepsilon}]| \\ \leq \|X_{\beta,\varepsilon,y}\|_2 \|Y_{\beta,\varepsilon,y}\mathbf{1}_{F_y^c}\|_4 P[F_y^c]^{1/4} &+ \|X_{\beta,\varepsilon,y}\mathbf{1}_{E_y^c}\|_4 \|Y_{\beta,\varepsilon,y}\|_2 P[E_y^c]^{1/4} \\ &+\|X_{\beta,\varepsilon}\|_2 \|Y_{\beta,\varepsilon}\|_2 (P[F_y^c]^{1/2} + P[E_y^c]^{1/2}). \end{aligned} \tag{5.16}$$

Define the constant $C := C(d, \mathcal{K}) := |\mathcal{K} \cap B_1(\mathbf{0})|$. We have

$$P[E_y^c] = P[F_y^c] = \sum_{j=0}^{k-1} \exp(-C(|y|/2)^d) \frac{(C(|y|/2)^d)^j}{j!}$$

while by (5.13), for $1/K \leq \beta \leq K$,

$$\mathbb{E}\,[X_{\beta,\varepsilon}^2] = E[Y_{\beta,\varepsilon}^2] = \mathbb{E}\,[\phi^\varepsilon(\beta\Gamma_k)^2] \leq \mathbb{E}\,[\phi^*(\beta\Gamma_k)^2] \leq a_K < \infty. \tag{5.17}$$

Note that since ϕ^* dominates $|\phi^\varepsilon|$, if $y \in \mathcal{K}$ then

$$
\begin{aligned}
|X_{\beta,\varepsilon,y}| &= |\phi^\varepsilon(\beta \min(\Gamma_k, \max(\Gamma_{k-1}, C|y|^d)))| \\
&\le \phi^*(\beta \min(\Gamma_k, \max(\Gamma_{k-1}, C|y|^d))), \qquad y \in \mathcal{K}.
\end{aligned} \tag{5.18}
$$

On the other hand, if $y \notin \mathcal{K}$, then $|X_{\beta,\varepsilon,y}| = |\phi^\varepsilon(\beta\Gamma_k)| \le \phi^*(\beta\Gamma_k)$, and hence, using also (5.18) and the fact that ϕ^* is nondecreasing on $[1,\infty)$, we have

$$
|X_{\beta,\varepsilon,y}| \le \phi^*(\beta\Gamma_k) \text{ if } C\beta|y|^d \ge 1. \tag{5.19}
$$

Thus for y such that $C\beta|y|^d \ge 1$, we have

$$
\mathbb{E}\,[X^2_{\beta,\varepsilon,y}] \le \mathbb{E}\,[\phi^*(\beta\Gamma_k)^2] \le a_K < \infty. \tag{5.20}
$$

If $C\beta|y|^d \ge 1$, then by (5.19) and then (4.3),

$$
\begin{aligned}
\mathbb{E}\,[X^4_{\beta,\varepsilon,y}\mathbf{1}_{E^c_y}] &\le \mathbb{E}\,[\phi^*(\beta\Gamma_k)^4\mathbf{1}\{\Gamma_k \ge C(|y|/2)^d\}] \\
&\le \mathbb{E}\,[\phi^*(\beta\Gamma_k)^4\mathbf{1}\{\Gamma_k \ge 1/(2^d\beta)\}] \\
&\le \mathbb{E}\,[(\phi^*(K\Gamma_k)^4 + \phi^*(2^{-d})^4)\mathbf{1}\{\Gamma_k \ge 1/(2^dK)\}] < \infty
\end{aligned}
$$

and similarly,

$$
\begin{aligned}
\mathbb{E}\,[Y^2_{\beta,\varepsilon,y}] &\le a_K < \infty; \\
\mathbb{E}\,[Y^4_{\beta,\varepsilon,y}\mathbf{1}_{F^c_y}] &\le \mathbb{E}\,[(\phi^*(K\Gamma_k)^4 + \phi^*(2^{-d})^4)\mathbf{1}\{\Gamma_k \ge 1/K\}] < \infty.
\end{aligned}
$$

By combining (5.16) with the subsequent estimates on the terms in the right hand side of (5.16), we find that $(|\psi_\varepsilon(\beta,y)|,\ y \in \mathbb{R}^d \setminus B_{(C\beta)^{-1/d}}(\mathbf{0}))$ is bounded by an integrable function of y that is independent of ε and of $\beta \in [1/K, K]$.

Next we find an integrable bound on $|\psi_\varepsilon(\beta,y)|,\ y \in B_{(C\beta)^{-1/d}}(\mathbf{0})$. By (5.17), the second term $\mathbb{E}\,[\phi^\varepsilon(\beta\Gamma_k)]^2$ in the definition (5.12) of ψ_ε is bounded by a_K for $\beta \in [K^{-1}, K]$. So we need only to consider the first term, which is $\mathbb{E}\,[X_{\beta,\varepsilon,y}Y_{\beta,\varepsilon,y}]$.

If $y \in B_{(C\beta)^{-1/d}}(\mathbf{0})$ satisfies also $y \in \mathcal{K}$, then by (5.18), $X^2_{\beta,\varepsilon,y} \le \phi^*(\beta C|y|^d)^2 + \phi^*(\beta\Gamma_k)^2$. This bound also holds if $y \notin \mathcal{K}$. An identical bound holds for $Y_{\beta,\varepsilon,y}$. By the Cauchy-Schwarz inequality, when $C\beta|y|^d \le 1$ we thus have

$$
\mathbb{E}\,[X_{\beta,\varepsilon,y}Y_{\beta,\varepsilon,y}] \le \mathbb{E}\,[\phi^*(\beta\Gamma_k)^2] + \phi^*(C\beta|y|^d)^2 \le a_K + \phi^*(C|y|^d/K)^2, \tag{5.21}
$$

where the last inequality comes from (5.13). We have

$$\int_{y \in B_{(K/C)^{1/d}}(\mathbf{0})} \phi^*(C|y|^d/K)^2 dy = \omega_d \int_0^{K/C} \phi^*(Cv/K)^2 dv < \infty.$$

Hence $|\psi_\varepsilon(\beta, y)|$ is bounded on $y \in B_{(C\beta)^{-1/d}}(\mathbf{0})$ by an integrable function, not dependent on β or ε, and combined with the earlier argument for $y \in \mathbb{R}^d \backslash B_{(C\beta)^{-1/d}}(\mathbf{0})$ this shows that there is an integrable function $\psi_K^* : \mathbb{R}^d \to [0, \infty)$ such that (5.14) holds. □

Our next two lemmas show that $V_{\phi^\varepsilon,k}(\beta)$ and $\Delta_{\phi^\varepsilon,k}(\beta)$ defined by (2.9) and (2.8), respectively, converge to $V_{\phi,k}(\beta)$ and $\Delta_{\phi,k}(\beta)$ as $\varepsilon \downarrow 0$.

Lemma 5.3 *For all $\beta > 0$ and $k \in \mathbb{N}$, $V_{\phi,k}(\beta)$ satisfies*

$$\lim_{\varepsilon \downarrow 0} V_{\phi^\varepsilon,k}(\beta) = V_{\phi,k}(\beta). \tag{5.22}$$

Moreover, given $K \in [1, \infty)$, it is the case that

$$\sup\{|V_{\phi^\varepsilon,k}(\beta)| : 0 < \varepsilon \leq 1, 1/K \leq \beta \leq K\} < \infty. \tag{5.23}$$

Proof. Let ϕ^* be the dominating function given by (2.1). If $\beta > 0$, then by (4.2), $\phi^*(\beta\Gamma_k)^2$ is a nonnegative integrable random variable, which dominates $\phi^\varepsilon(\beta\Gamma_k)^2$, so by the dominated convergence theorem, as $\varepsilon \downarrow 0$ we have

$$\mathbb{E}\,[\phi^\varepsilon(\beta\Gamma_k)] \to \mathbb{E}\,[\phi(\beta\Gamma_k)]; \tag{5.24}$$

$$\mathbb{E}\,[(\phi^\varepsilon(\beta\Gamma_k))^2] \to \mathbb{E}\,[(\phi(\beta\Gamma_k))^2]. \tag{5.25}$$

Fix $K \in [1, \infty)$. Let $\phi^{\varepsilon,*}(t) := \phi^*(t)\mathbf{1}\{t \geq \varepsilon\}$. For $y \in \mathbb{R}^d \setminus \{\mathbf{0}\}$ and $\beta \in [1/K, K]$, by the Cauchy-Schwarz inequality, (5.20) and (5.21),

$$\mathbb{E}\,[\phi^{\varepsilon,*}(\beta|C(\mathbf{0}, \mathcal{H} \cup y)|)\phi^{\varepsilon,*}(\beta|C(y, \mathcal{H} \cup \mathbf{0})|)] \leq a_K + \phi^*(C|y|^d/K) < \infty$$

and by the monotone convergence theorem this still holds with $\phi^{\varepsilon,*}$ replaced by ϕ^* on the left so that by the dominated convergence theorem, as $\varepsilon \downarrow 0$ we have for fixed y and β

$$\mathbb{E}\,[\phi^\varepsilon(\beta|C(\mathbf{0}, \mathcal{H} \cup y)|)\phi^\varepsilon(\beta|C(y, \mathcal{H} \cup \mathbf{0})|)] \to \mathbb{E}\,[\phi(\beta|C(\mathbf{0}, \mathcal{H} \cup y)|)\phi(\beta|C(y, \mathcal{H} \cup \mathbf{0})|)].$$

Combined with (5.24) and recalling (5.11) and (5.12), this shows that

$$\psi_\varepsilon(\beta, y) \to \psi(\beta, y) \quad \text{as} \quad \varepsilon \downarrow 0. \tag{5.26}$$

By (5.25), (5.26), (5.14) and the dominated convergence theorem, the right hand side of the expression (2.9) for $V_{\phi^\varepsilon,k}(\beta)$ converges to the corresponding expression for $V_{\phi,k}(\beta)$, and (5.22) follows.

By (2.9), (5.13) and (5.14), it is also the case that for all $\varepsilon \in (0, 1]$ and $\beta \in [1/K, K]$ we have $|V_{\phi^\varepsilon,k}(\beta)| \le a_K + \int_{\mathbb{R}^d} \psi_K^*(y)dy$ and this bound is finite, so (5.23) follows. □

Lemma 5.4 *For any $\beta > 0$ and $k \in \mathbb{N}$ we have*

$$\lim_{\varepsilon \downarrow 0} \Delta_{\phi^\varepsilon,k}(\beta) = \Delta_{\phi,k}(\beta). \tag{5.27}$$

Also, given $K \in [1, \infty)$, it is the case that

$$\sup\{|\Delta_{\phi^\varepsilon,k}(\beta)| : 0 < \varepsilon \le 1, 1/K \le \beta \le K\} < \infty. \tag{5.28}$$

Proof. By (2.8) and (5.24) we obtain (5.27). Also (2.8) implies $|\Delta_{\phi^\varepsilon,k}(\beta)| \le (k+1)\mathbb{E}[\phi^*(\beta\Gamma_k)] + k\mathbb{E}[\phi^*(\beta\Gamma_{k+1})]$ and the bound (5.28) easily follows from this with (5.13). □

Given $h \in \mathcal{B}(A)$, let $L_h(\phi)$ be the limiting variance in the statement of Theorem 2.1, i.e. let

$$L_h(\phi) := \int_A h^2(x) V_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx - \left(\int_A h(x)\Delta_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx\right)^2. \tag{5.29}$$

Lemma 5.5 *Given $h \in \mathcal{B}(A)$, it is the case that*

$$\lim_{\varepsilon \downarrow 0} L_h(\phi^\varepsilon) = L_h(\phi). \tag{5.30}$$

Proof. By assumption, $(g(x)/f(x), x \in A)$ is bounded away from 0 and ∞, and f is bounded. Hence by (5.23), the integrand in the first integral in the expression (5.29) for $L_h(\phi^\varepsilon)$ is bounded by a constant, not depending on ε. Similarly, by (5.28), the integrand in the second integral in the expression (5.29) for $L_h(\phi^\varepsilon)$ is bounded by a constant, not depending on ε. By (5.22) and (5.27), for both integrals the integrand converges, as $\varepsilon \downarrow 0$, to the corresponding integrand for $L_h(\phi)$. So by the dominated convergence theorem, the integrals converge and (5.30) follows. $\square$

Proof of Theorem 2.1. Let $h \in \mathcal{B}(A)$. Given $\delta > 0$, by Lemmas 5.1 and 5.5 we can find $\varepsilon_0 > 0$ and $n_0 > 0$ such that for $\varepsilon < \varepsilon_0$ and $n \geq n_0$ we have

$$|L_h(\phi^\varepsilon) - L_h(\phi)| < \delta \tag{5.31}$$

and $n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi_\varepsilon,k}\rangle] \leq \delta$. The function ϕ^ε lies in the class $\mathcal{F}_0$, so by Proposition 4.2,

$$\lim_{n\to\infty} n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi^\varepsilon,k}\rangle] = L_h(\phi^\varepsilon) \tag{5.32}$$

and hence by the Cauchy-Schwarz inequality, for large enough n we have

$$\begin{aligned} &n^{-1}|\mathrm{Var}(\langle h, \nu^g_{n,\phi,k}\rangle) - \mathrm{Var}(\langle h, \nu^g_{n,\phi^\varepsilon,k}\rangle)| \\ &= |n^{-1}\mathrm{Var}(\langle h, \nu^g_{n,\phi_\varepsilon,k}\rangle) + 2\mathrm{Cov}(n^{-1/2}\langle h, \nu^g_{n,\phi^\varepsilon,k}\rangle, n^{-1/2}\langle h, \nu^g_{n,\phi_\varepsilon,k}\rangle)| \\ &\leq \delta + 2\delta^{1/2}(n^{-1}\mathrm{Var}\langle h, \nu^g_{n,\phi^\varepsilon,k}\rangle)^{1/2} \leq \delta + 2\delta^{1/2}(L_h(\phi) + \delta)^{1/2}. \end{aligned}$$

Using (5.31) and (5.32), for large enough n, we thus have

$$|n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi,k}\rangle] - L_h(\phi)| \leq 3\delta + 2\delta^{1/2}(L_h(\phi) + \delta)^{1/2}$$

and since $\delta > 0$ is arbitrary this shows that

$$n^{-1}\mathrm{Var}[\langle h, \nu^g_{n,\phi,k}\rangle] \to L_h(\phi) \quad \text{as} \quad n \to \infty,$$

which is the first part of the statement of Theorem 2.1.

To prove convergence to a Gaussian field with covariance (2.12), by standard arguments based on the Cramér-Wold device, it suffices to show that for any $h \in \mathcal{B}(\mathbb{R}^d)$,

$$n^{-1/2}\langle h, \overline{\nu}^g_{n,\phi,k}\rangle \xrightarrow{\mathcal{D}} N(0, L_h(\phi)). \tag{5.33}$$

Let $t \in \mathbb{R}$. Set $X_n := n^{-1/2}\langle h, \overline{\nu}^g_{n,\phi,k}\rangle$ and for $\varepsilon > 0$ set $X^\varepsilon_n := n^{-1/2}\langle h, \overline{\nu}^g_{n,\phi^\varepsilon,k}\rangle$. Since ϕ^ε is in $\mathcal{F}_0$, Proposition 4.2 shows that $X^\varepsilon_n \xrightarrow{\mathcal{D}} N(0, L_h(\phi^\varepsilon))$ as $n \to \infty$. Hence,

$$\mathbb{E}\left[\exp(itX^\varepsilon_n)\right] - \exp(-t^2 L_h(\phi^\varepsilon)/2) \to 0 \quad \text{as} \quad n \to \infty. \tag{5.34}$$

Given $\delta > 0$, by Lemmas 5.1 and 5.5 we can choose $\varepsilon > 0$ such that for large n,

$$\mathbb{E}\left[|\exp(itX_n) - \exp(itX_n^\varepsilon)|\right] \leq \mathbb{E}\left[|t(X_n - X_n^\varepsilon)|\right] \leq \delta$$

and also $|e^{-t^2 L_h(\phi)/2} - e^{-t^2 L_h(\phi^\varepsilon)/2}| \leq \delta$ so that combining with (5.34) we have for large n that

$$|\mathbb{E}\left[\exp(itX_n)\right] - e^{-t^2 L_h(\phi)/2}| \leq 3\delta$$

and since δ is arbitrary, this implies (5.33). □

6 Proofs of Propositions 2.1 and 2.2

6.1 Proof of Proposition 2.1

First we identify $V_{\phi,1}(\beta)$ when $\mathcal{K} \neq \mathbb{R}^d$, which implies $-y \notin \mathcal{K}$ for all $y \in \mathcal{K}$. The integral in (2.9) has contributions only from $y \in \mathcal{K}$ and from $\mathbf{0} \in (y + \mathcal{K})$, and these contributions are equal by a symmetry argument. Let $b_d := |B_1^{\mathcal{K}}(\mathbf{0})|$.

Consider $y \in \mathcal{K}$. Then $|C_1(\mathbf{0}, \mathcal{H} \cup y)|$ has the distribution of $\min(\Gamma_1, b_d|y|^d)$ and $|C_1(y, \mathcal{H} \cup \mathbf{0})|$ has the distribution of Γ_1, and they are independent. Hence, the integral in (2.9) is equal to

$$\begin{aligned} 2\int_{\mathcal{K}} &\mathbb{E}\left[\phi(\beta\Gamma_1)\right](\mathbb{E}\left[\phi(\beta\min(\Gamma_1, b_d|y|^d)) - \phi(\beta\Gamma_1)\right])dy \\ &= 2M_{\phi,1}(\beta)\int_0^\infty \mathbb{E}\left[\phi(\beta\min(\Gamma_1, s)) - \phi(\beta\Gamma_1)\right]ds \\ &= 2M_{\phi,1}(\beta)\mathbb{E}\int_0^{\Gamma_1}(\phi(\beta s) - \phi(\beta\Gamma_1))ds. \end{aligned}$$

In the last expectation, the first term is equal to $\int_0^\infty \phi(\beta s)P[\Gamma_1 \geq s]$ which comes to $M_{\phi,1}(\beta)$. The second term comes to $M_{\phi,2}(\beta)$ as in (4.29). Thus, the integral in (2.9) is equal to to $2M_{\phi,1}(\beta)(M_{\phi,1}(\beta) - M_{\phi,2}(\beta))$ and substituting in (2.9) we find that $V_{\phi,1}(\beta)$ is given by case $k = 1$ of formula (2.13) when $\mathcal{K} \neq \mathbb{R}^d$.

To complete the proof of Proposition 2.1, we need to show that $V_{\phi,k}$ is given by (2.13) in the case when $d = 1$ (for arbitrary k, but still assuming $\mathcal{K} \neq \mathbb{R}^d$). There are only two possibilities for $\mathcal{K}$ and by symmetry it suffices to consider the case with $\mathcal{K} = (0, \infty)$. In this case, the expression (2.9) becomes:

$$V_{\phi,k}(\beta) := M_{\phi^2,k}(\beta)] + \int_{-\infty}^\infty c^\beta(\mathbf{0}, y)dy \tag{6.35}$$

where

$$c^\beta(\mathbf{0},y) := \mathbb{E}\left[\phi(\beta C_{\mathbf{0}})\phi(\beta C_y)\right] - (\mathbb{E}\left[\phi(\beta\Gamma_k)\right])^2,$$

where $C_{\mathbf{0}}$ (respectively, C_y) denotes the length of the k-spacing starting at the origin (respectively, starting at y) with respect to the augmented point set $\mathcal{H} \cup \mathbf{0} \cup y$.

We proceed to evaluate the integral in (6.35). Write $e_k := \mathbb{E}\left[\phi(\beta\Gamma_k)\right]$. Then

$$\begin{aligned} c^\beta(\mathbf{0},y) &= \mathbb{E}\left[(\phi(\beta C_{\mathbf{0}})\phi(\beta C_y) - \phi(\beta\Gamma_k)e_k)(\mathbf{1}\{y \le \Gamma_k\} + \mathbf{1}\{y > \Gamma_k\})\right] \\ &= \mathbb{E}\left[(\phi(\beta C_{\mathbf{0}})\phi(\beta C_y) - \phi(\beta\Gamma_k)e_k)\mathbf{1}\{y \le \Gamma_k\}\right]. \end{aligned}$$

Integrating over y and setting $\Gamma_0 := 0$, we have that

$$\int_0^\infty c^\beta(\mathbf{0},y)dy = \left(\sum_{j=1}^k I_j\right) - \int_0^\infty \mathbb{E}\left[\phi(\beta\Gamma_k)e_k\mathbf{1}\{y \le \Gamma_k\}\right]dy, \tag{6.36}$$

where we set

$$I_j := \mathbb{E}\int_{\Gamma_{j-1}}^\infty (\phi(\beta C_{\mathbf{0}})\phi(\beta C_y)\cdot \mathbf{1}\{y \le \Gamma_j\})dy.$$

Recall that $\Gamma_j = \sum_{i=1}^j \Gamma_{1,i}$. We now compute I_j in the case with $1 \le j \le k-1$. For such j, if $\Gamma_{j-1} < y < \Gamma_j$ then $C_{\mathbf{0}} = \Gamma_{k-1}$ and $C_y = \Gamma_{j+k-1} - y$; setting $w = y - \Gamma_{j-1}$ we have for $1 \le j \le k-1$ that

$$\begin{aligned} I_j = \mathbb{E}\int_0^\infty &\phi(\beta(\Gamma_{j-1} + w + (\Gamma_{1,j} - w) + \sum_{i=j+1}^{k-1}\Gamma_{1,i})) \\ &\times\phi(\beta(\Gamma_{1,j} - w + \sum_{i=j+1}^{j+k-1}\Gamma_{1,i}))\mathbf{1}\{\Gamma_{1,j} \ge w\}dw. \end{aligned}$$

Now take the expectation inside the integral. Since $\Gamma_{1,j}$ is exponential we have $P[\Gamma_{1,j} \ge w] = e^{-w}$, and by conditioning on this event, using the memoryless property of the exponential distribution and independence of $\Gamma_{1,j}$ from the other random variables in the expression, we obtain

$$I_j = \int_0^\infty \mathbb{E}\,\phi(\beta(\Gamma_{j-1} + w + \Gamma_{1,j} + \sum_{i=j+1}^{k-1}\Gamma_{1,i}))\phi(\beta(\Gamma_{1,j} + \sum_{i=j+1}^{j+k-1}\Gamma_{1,i}))e^{-w}dw.$$

Now take the integral back inside the expectation. Letting $\Gamma_{1,0}$ be a further independent exponential random variable with density function $e^{-w}, w \ge 0$, we have that

$$I_j = \mathbb{E}\,[\phi(\beta(\Gamma_{j-1} + \Gamma_{1,0} + \Gamma_{1,j} + \sum_{i=j+1}^{k-1} \Gamma_{1,i}))\phi(\beta(\Gamma_{1,j} + \sum_{i=j+1}^{j+k-1} \Gamma_{1,i}))]$$

$$= \mathbb{E}\,[\phi(\beta(\sum_{i=0}^{k-1} \Gamma_{1,i}))\phi(\beta(\sum_{i=j}^{k+j-1} \Gamma_{1,i}))] = \mathbb{E}\,[\phi(\beta\Gamma_k)\phi(\beta(\Gamma_{k+j} - \Gamma_j)]. \qquad (6.37)$$

To deal with I_k we modify the preceding argument as follows. If $\Gamma_{k-1} < y < \Gamma_k$ then $C_{\mathbf{0}} = y$ and $C_y = \Gamma_{2k+1} - y$. Setting $w = y - \Gamma_{k-1}$ we have that

$$I_k = \mathbb{E} \int_0^\infty \phi(\beta(\Gamma_{k-1} + w))\phi(\beta(\Gamma_{1,k} - w + \sum_{i=k+1}^{2k-1} \Gamma_{1,i}))\mathbf{1}\{\Gamma_{1,k} \geq w\}dw.$$

Conditioning on the event that $\Gamma_{1,k} \geq w$ using the memoryless property of the exponential distribution and independence of $\Gamma_{1,k}$ from the other random variables in the expression, we obtain

$$I_k = \mathbb{E} \int_0^\infty \phi(\beta(\Gamma_{k-1} + w))\phi(\beta(\Gamma_{1,k} + \sum_{i=k+1}^{2k-1} \Gamma_{1,i}))e^{-w}dw.$$

Letting $\Gamma_{1,0}$ be a further independent exponential random variable we have that

$$I_k = \mathbb{E}\,[\phi(\beta(\Gamma_{k-1} + \Gamma_{1,0}))\phi(\beta(\Gamma_{1,k} + \sum_{i=k+1}^{2k-1} \Gamma_{1,i}))]$$

$$= \mathbb{E}\,[\phi(\beta(\sum_{i=0}^{k-1} \Gamma_{1,i}))\phi(\beta(\sum_{i=k}^{2k-1} \Gamma_{1,i}))] = e_k^2.$$

Now as in (4.29) the last term in (6.36) is

$$e_k \int_0^\infty \mathbb{E}\,[\phi(\beta\Gamma_k)\mathbf{1}\{y < \Gamma_k\}]dy = ke_k\mathbb{E}\,[\phi(\beta\Gamma_{k+1})].$$

Combining this with the preceding expressions for I_j $(j < k)$ and for I_k, we may rewrite (6.36) as

$$\int_0^\infty c^\beta(\mathbf{0}, y)dy = \left(\sum_{j=1}^{k-1} \mathbb{E}\,[\phi(\beta\Gamma_k)\phi(\beta\Gamma_{k+j} - \beta\Gamma_j)]\right) + e_k^2 - ke_ke_{k+1}$$

$$= \left(\sum_{j=1}^{k-1} \left(\mathbb{E}\,[\phi(\beta\Gamma_k)\phi(\beta\Gamma_{k+j} - \beta\Gamma_j)] - e_k^2\right)\right) + ke_k(e_k - e_{k+1})$$

$$= ke_k(e_k - e_{k+1}) + \sum_{j=1}^{k-1} \mathrm{Cov}(\phi(\beta\Gamma_k), \phi(\beta\Gamma_{k+j} - \beta\Gamma_j)).$$

By symmetry, for all β we have $\int_{-\infty}^0 c^\beta(\mathbf{0}, y)dy = \int_0^\infty c^\beta(\mathbf{0}, y)dy$ and thus from (6.35) we obtain for all $\beta > 0$ that $V_{\phi,k}(\beta)$ is given by (2.13). This completes the proof of Proposition 2.1. $\square$

6.2 Proof of Proposition 2.2

We deduce Proposition 2.2 as follows. From the definition (2.9) we obtain

$$V_{\phi,1}(\beta) = M_{\phi^2,1}(\beta) + \int_{\mathbb{R}^d} c(\mathbf{0}, y)dy,$$

where $c(\mathbf{0}, y) := \mathbb{E}\left[\phi(\beta|C_1(\mathbf{0}, \mathcal{H} \cup y)|)\phi(\beta|C_1(y, \mathcal{H} \cup \mathbf{0})|)\right] - (\mathbb{E}\left[\phi(\beta\Gamma_1)\right])^2$. For all $s, t \in \mathbb{R}^+$, let $p(s,t) := P[|C_1(\mathbf{0}, \mathcal{H} \cup y)| > s, |C_1(y, \mathcal{H} \cup \mathbf{0})| > t]$. Then for all $s, t \in [0, |y|^d \omega_d]$ we have

$$p(s,t) = e^{-(s+t)+I(s,t,|y|)}.$$

Otherwise $p(s,t) = 0$. Hence, for $y \in \mathbb{R}^d$, by the fundamental theorem of calculus, the assumption that ϕ is differentiable with $\lim_{t\downarrow 0} \phi(t) = 0$, and Fubini's theorem,

$$\begin{aligned} c(\mathbf{0}, y) &= \mathbb{E} \int_0^\infty \int_0^\infty \beta^2 \phi'(\beta s)\phi'(\beta t) \mathbf{1}_{\{|C_1(\mathbf{0},\mathcal{H}\cup y)|>s, |C_1(y,\mathcal{H}\cup\mathbf{0})|>t\}} dsdt \\ &\qquad - \left(\mathbb{E} \int_0^\infty \beta\phi'(\beta u)\mathbf{1}_{\{\Gamma_1 > u\}} du\right)^2 \\ &= \beta^2 \int_0^\infty \int_0^\infty \phi'(\beta s)\phi'(\beta t)[p(s,t) - e^{-(s+t)}]dsdt. \end{aligned}$$

Since $p(s,t)$ vanishes whenever $(s,t) \notin [0, |y|^d\omega_d]^2$ we obtain

$$\begin{aligned} c(\mathbf{0}, y) = \beta^2 &\int_0^{|y|^d\omega_d} \int_0^{|y|^d\omega_d} \phi'(\beta s)\phi'(\beta t)[e^{-(s+t)+I(s,t,|y|)} - e^{-(s+t)}]dsdt - \\ &-\beta^2 \int\int_{\max(s,t)\geq |y|^d\omega_d} \phi'(\beta s)\phi'(\beta t)e^{-(s+t)}dsdt. \end{aligned}$$

Therefore

$$\begin{aligned} \int_{\mathbb{R}^d} c(\mathbf{0}, y)dy = \beta^2 &\int_{\mathbb{R}^d} \int_0^{|y|^d\omega_d} \int_0^{|y|^d\omega_d} \phi'(\beta s)\phi'(\beta t)[e^{-(s+t)+I(s,t,|y|)} - e^{-(s+t)}]dsdtdy - \\ &-\beta^2 \int_{\mathbb{R}^d} \int\int_{\max(s,t)\geq |y|^d\omega_d} \phi'(\beta s)\phi'(\beta t)e^{-(s+t)}dsdtdy \end{aligned}$$

and letting $u := |y|^d\omega_d$, the above becomes

$$\begin{aligned} \int_0^\infty c(\mathbf{0}, y)dy = \beta^2 &\int_0^\infty \int_0^u \int_0^u \phi'(\beta s)\phi'(\beta t)e^{-(s+t)}[e^{I(s,t,(u/\omega_d)^{1/d})} - 1]dsdtdu - \\ &-\beta^2 \int_0^\infty \int\int_{\max(s,t)\geq u} \phi'(\beta s)\phi'(\beta t)e^{-(s+t)}dsdtdu. \end{aligned}$$

Finally, change the order of integration to obtain

$$\begin{aligned} \int_0^\infty c(\mathbf{0}, y)dy = \beta^2 &\int_0^\infty \int_0^\infty \phi'(\beta s)\phi'(\beta t)e^{-(s+t)} \int_{\max(s,t)}^\infty [e^{I(s,t,(u/\omega_d)^{1/d})} - 1]dudsdt - \\ &-\beta^2 \int_0^\infty \int_0^\infty \phi'(\beta s)\phi'(\beta t)e^{-(s+t)} \int_0^{\max(s,t)} dudsdt, \end{aligned}$$

which is exactly the desired limit. This proves Proposition 2.2. □

7 Proofs of Theorems 2.2 and 2.3

7.1 Proof of Theorem 2.2

We need to check that the proof of Theorem 2.1 in Sections 4 and 5 carries through when we modify $C_k^{\mathcal{K}}(x, \mathcal{X})$ to be the empty set whenever $\mathrm{card}(\mathcal{X} \cap (x + \mathcal{K}) \setminus x) < k$, and set $\phi(0) = 0$.

This change has the effect only of modifying the value of $\phi(C_k^{\mathcal{K}}(x, \mathcal{X}))$ to zero on certain outcomes. Therefore it is easy to see that all moments conditions carry through. This applies in particular to Lemmas 4.1 and 5.1.

The stated change also has no effect on the lemmas concerned with weak convergence (Lemmas 4.2 and 4.4) because for the limiting random object (a homogeneous Poisson process) the probability of seeing fewer than k points in the cone $\mathcal{K}$ is zero, and hence the value of $\phi(C_k^{\mathcal{K}}(\mathcal{H}_a))$ is unaffected by the modification being considered, for any $a > 0$.

Essentially as a result of these two observations, it can be checked that all steps in the proofs in Sections 4 and 5 carry through, and so the result holds. □

7.2 Proof of Theorem 2.3

We prove Theorem 2.3 by again following the proof of Theorem 2.1. The proof is a bit simpler since the analog of Lemma 4.1 holds for all $\phi \in \mathcal{F}$, thus avoiding the lengthy truncation arguments.

Put for all $t, \lambda > 0$, locally finite point sets $\mathcal{X}$, and $x \in \mathcal{X}$

$$\Phi_\lambda^{g,t}(x, \mathcal{X}) := \phi\left(\mathrm{card}\{\mathcal{X} \cap B_{t(\lambda g(x))^{-1/d}}(x)\}\right)$$

and

$$\xi_\infty^{g,x,t}(\mathcal{X}) := \phi\left(\mathrm{card}\{\mathcal{X} \cap B_{t(g(x))^{-1/d}}(x)\}\right).$$

Since $\Phi_\lambda^{g,t}(x, \mathcal{P}_\lambda) \stackrel{\mathcal{D}}{=} \phi(\mathrm{Po}(m))$ where $\mathrm{Po}(m)$ denotes a Poisson random variable with parameter $m := m(t, x)$, and since $\mathbb{E}\left[\phi^*(\mathrm{Po}(m))^4\right] < \infty$ for any m, it follows that Lemma 4.1 holds for any $\phi \in \mathcal{F}$. Moreover the analog of Lemma 4.2 goes through verbatim with this definition of $\Phi_\lambda^{g,t}(\cdot, \cdot)$. Since the functional $\Phi_\lambda^{g,t}(\cdot, \cdot)$ depends only on points at a fixed deterministic distance from x, the functional $\Phi_\lambda^{g,t}(\cdot, \cdot)$ clearly localizes and Lemma 4.3 holds as well.

It follows that Proposition 4.1 and Proposition 4.2 also hold verbatim with $V^{\xi}_{\phi,k}(x,a)$ replaced by

$$V^{\xi}(x,a) = \mathbb{E}\,[\xi_{\infty}^{g,x,t}(\mathcal{H}_a)^2] + a\int_{\mathbb{R}^d}\left[\mathbb{E}\,\xi_{\infty}^{g,x,t}(\mathcal{H}_a\cup y)\xi_{\infty}^{g,x,t}(-y+\mathcal{H}_a\cup\mathbf{0}) - (\mathbb{E}\,\xi_{\infty}^{g,x,t}(\mathcal{H}_a))^2\right]dy$$

and $\delta^{\xi}_{\phi,k}(x,a)$ replaced by

$$\delta^{\xi}(x,a) = \mathbb{E}\,[\xi_{\infty}^{g,x,t}(\mathcal{H}_a)] + a\int_{\mathbb{R}^d}\mathbb{E}\,\left[\xi_{\infty}^{g,x,t}(\mathcal{H}_a\cup y) - \xi_{\infty}^{g,x,t}(\mathcal{H}_a)\right]dy.$$

This proves Theorem 2.3. □

Acknowledgements: Yu. Baryshnikov thanks the Lehigh Mathematics department for hosting his research visit. J. Yukich gratefully thanks R. Jimenez for multiple inspiring discussions related to spacings.

References

[1] Ali, S. M. and S. D. Silvey (1964), Association between random variables and the dispersion of the Radon-Nikodym derivative. J. R. Stat. Soc. Ser. B, **26**, 100-107.

[2] Ali, S. M. and S. D. Silvey (1964), A further result on the relevance of the dispersion of a Radon-Nikodym derivative to the problem of measuring association. J. R. Statist. Soc. Ser. B, **26**, 108-110.

[3] Ali, S. M. and S. D. Silvey (1966), A general class of coefficients of divergence of one distribution from another. J. R. Statist. Soc. Ser. B, **28**, 131-142.

[4] Barbour, A. D., Holst, L. and S. Janson (1992), Poisson Approximation, Clarendon Press, Oxford.

[5] Baryshnikov, Yu. and J. E. Yukich (2005), Gaussian limits for random measures in geometric probability, *Annals Appl. Prob.*, **15**, 1A, 213-253.

[6] Billingsley, P. (1968). *Convergence of Probability Measures.* John Wiley, New York.

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

[8] Blumenthal, S. (1968), Logarithms of sample spacings. SIAM J. Appl. Math. **16**, 1184-1191.

[9] Byers, S., and A. Raftery (1998), Nearest-neighbor clutter removal for estimating features in spatial point processes. JASA, **93**, no. 442, pp. 577- 584.

[10] Cressie, N. (1976), On the logarithm of high-order spacings. Biometrika, **63**, 343-355.

[11] Csiszár, I. (1963), Eine informationstheoretische Ungleichung und ihre Anwendung auf den Beweis der Ergodizitaät von Markoffschen Ketten. Magyar Tud. Akad. Mat. Kutató Int. Közl. 8, 85-108.

[12] Csiszár, I. (1967), Information-type measures of difference of probability distributions and indirect observations. Studia Sci. Math. Hungarica., **4**, 2, 299-318.

[13] Csiszár, I. (1978), Information measures: A critical survey. In Transaction 7th Prague Conf. on Info. Th Statist., Decis. Funct., Random Process and 8th European Meeting of Statist. Academia, Prague, 73-86.

[14] Czekala, F. (1993), Asymptotic distributions of statistics based on logarithm of spacings. Zastos. Mat. **21**, 511-519.

[15] Darling, D. A. (1953), On a class of problems related to the random division of an interval. Ann. Math. Statist., **24**, 239-253.

[16] Deheuvels, P., J. H. J. Einmahl, D. Mason, and F. Ruymgaart (1988), The almost sure behavior of maximal and minimal multivariate k_n spacings. J. Mult. Anal., **24**, 155-176.

[17] Dudewicz, E. J. and E. C. Van der Meulen (1987), The empiric entropy, a new approach to non-parametric density estimation. In Eds. M. I. Puri, J. Vilaplana and M. Wertz. New perspectives in theoretical and applied statistics. Wiley, New York, 202-227.

[18] L'Ecuyer, P., J.F. Cordeau, and R. Simard (2000), Close-point spatial tests and their application to random number generators. Operations Res. **48**, 308-317.

[19] Efron, B. and Stein, C. (1981) The jackknife estimate of variance. *Ann. Statist.* **9**, 586–596.

[20] Gebert, J. R. and B. K. Kale (1969), Goodness of fit tests based on discriminatory information. Statist. Hefte. **3**, 192-200.

[21] Ghosh, K. and S. Rao Jammalamadaka (2001), A general estimation method using spacings. J. of Statistical Planning and Inference, **93**, 71-82.

[22] Hall, P. (1986), On powerful distributional tests based on sample spacings. J. Mult. Analysis, **19**, 201-224.

[23] Henze, N. (1982), The limit distribution for maxima of 'weighted' *rth* nearest-neighbor distances. J. Appl. Prob., **19**, 344-354.

[24] Holst, L. (1979), Asymptotic normality of sum-functions of spacings. Annals of Prob., **7**, 1066-1072.

[25] Holst, L. and J. S. Rao (1981), Asymptotic spacings theory with applications to the two-sample problem. Canadian J. Statist., **9**, 79-89.

[26] Jimenez, R. and J.E. Yukich (2002), Asymptotics for statistical distances based on Voronoi tessellations. Journal of Theoretical Probability, **15**, 2, 503-541.

[27] Jimenez, R. and J.E. Yukich (2002), Strong laws for Euclidean graphs with general edge weights. Statist. Probab. Lett., **56**, 251-259.

[28] Jimenez, R. and J. E. Yukich (2005), Statistical distances based on Euclidean graphs. Recent Advances in Probability and its Applications, Kluwer Acad. Pubs., ed. J. Glaz and H. Gyzl, 223-239.

[29] Khashimov, Sh. A. (1989), Asymptotic properties of functions of spacings, Theory Prob. Appl., **34**, 298-306.

[30] Moran, P. (1947), The random division of an interval. J. R. Statist. Soc. Ser. B, **9**, 92-98.

[31] Penrose, M. D. (1999), A strong law for the largest nearest-neighbor link between random points J. London Mathematical Society, Second Series, **60**, 951-960.

[32] Penrose, M. D. (2003), Random Geometric Graphs, Clarendon Press, Oxford.

[33] Penrose, M. D. (2000), Central limit theorems for k-nearest neighbor distances, Stochastic Processes and Their Applications, **85**, 295-320.

[34] Penrose, M. D. (2005), Multivariate spatial central limit theorems with applications to percolation and spatial graphs, *Ann. Probab.*, **33**, 1945- 1991.

[35] Penrose, M. D. (2006), Laws of large numbers in stochastic geometry with statistical applications. *Bernoulli*, to appear.

[36] Penrose, M. D. (2006), Gaussian limits for random geometric measures, *Electron. J. Probab.* , to appear.

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

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

[39] Penrose, M. D. and J. E. Yukich (2005), Normal approximation in geometric probability. In Stein's Method and Applications, Lecture Note Series, Institute for Mathematical Sciences, National University of Singapore, **5**, A. D. Barbour and Louis H. Y. Chen, Eds., 37-58, Electronically available via `http://www.lehigh.edu/∼jey0/publications.html`.

[40] del Pino, G. (1979), On the asymptotic distribution of k-spacings with applications to goodness-of-fit tests. Annals of Statistics, **7**, 1058-1065.

[41] Pyke, R. (1965), Spacings. J. R. Statist. Soc. Ser. B, **27**, 395-436.

[42] Ranneby, B. (1984), The maximum spacing method. An estimation method related to the maximum likelihood method, Scand. J. Statist., 93-112.

[43] Read, T. R. C. and N. Cressie (1988), Goodness-of-fit Statistics for Discrete Multivariate Data, Springer, New York

[44] Rényi, A. (1961), Measures of entropy and information. Proc. Fourth Berkeley Symp., **1**, 547-561.

[45] Schilling, M. (1983), Goodness of fit testing in $\mathbb{R}^m$ based on the weighted empirical distribution of certain nearest neighbor statistics, Ann. Stat., **11**, 1-12.

[46] Shao, Y. and R. Jimenez (1998), Entropy for random partitions and its applications. J. Theoretical Prob., 417-433.

[47] Shao, Y. and M. G. Hahn (1995), Limit theorems for logarithm of sample spacings. Statistics and Probability Letters, 121-132.

[48] Shao, Y. and M. G. Hahn (1999), Strong consistency of the maximum product of spacings estimates with applications in nonparametrics and in estimation of unimodal densities. Ann. Inst. Statist., **51**, Math. 31-49.

[49] van Es, B. (1992), Estimating functionals related to a density by a class of statistics based on spacings Scand. J. Statist., **19**, 61-72.

[50] Weiss, L. (1957), Asymptotic power of certain tests of fit based on sample spacings. Ann. Math. Statist., **28**, 783-786.