Gaussian Limits for Generalized Spacings

Yu. Baryshnikov and J. E. Yukich[1]

Bell Laboratories and Lehigh University

December 12, 2004

Abstract

Nearest neighbor cells in $\mathbb{R}^d$, $d \geq 1$, are used to define coefficients of divergence (ϕ-divergences) between continuous multivariate samples. For large sample sizes, such distances are shown to converge to a normal distribution 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 upon and generalizes 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 Introduction

Let $X_i, 1 \leq i \leq n$, denote 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 $\mathbb{R}$ [35] take the form of an *empirical ϕ-divergence*

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

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

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

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

[1] Research supported in part by NSF Grant DMS-0203720

where $\phi : \mathbb{R}^+ \to \mathbb{R}$ is a specified function. When F and G have densities f and g, respectively, the functionals (1.1) form a discrete version of the ϕ-divergence between f and g, namely $\int f(x)\phi(\frac{g(x)}{f(x)})dx$. ϕ-divergences, introduced independently by Csiszár [10, 11], and Ali and Silvey [1, 2, 3], are a measure of the distance between the distributions F and G. They are widely used in non-parametric estimation and are well suited for goodness-of-fit tests [7, 9, 15, 22, 34, 37, 44].

This paper has two main goals. The first is to use nearest neighbor cells to establish high dimensional analogs of the one dimensional ϕ-divergences (1.1). The nearest neighbor cells are employed as an adaptive scheme to define statistical distances of continuous samples in $\mathbb{R}^d$. We establish a general central limit theorem (CLT) showing that the resulting distance functionals converge to a normal random variable whenever the distributions F and G have densities bounded away from zero and infinity. Using one and two point correlation functions, together with the fact that nearest neighbor cells have local interactions, we show that the large scale behavior of generalized ϕ-divergences, as defined by limiting variances and covariances in the CLT, is determined by the local behavior of the underlying density of points. The high dimensional analogs of the ϕ-divergences (1.1) generate canonical random point measures, denoted $\mu^g_{n,\phi}$, $n \geq 1$, whose finite dimensional distributions, after re-normalization, converge to those of a generalized Gaussian field.

The second goal of this paper is to establish a general CLT providing asymptotic normality of (1.1) in $d = 1$ under general conditions on F, G, and ϕ. This umbrella CLT generalizes and extends upon classical CLTs for sum functions of spacings [7, 9, 12, 13, 17, 41, 44] and establishes asymptotic normality for information gain, log-likelihood ratios, and sums of logarithmic k-spacings whenever the densities of F and G are bounded away from zero and infinity. The latter resolves an open question of Darling (p. 249 of [13]). The methods extend to yield a CLT for the number of pairs of sample points within a fixed distance.

Our approach relies upon stabilization methods to quantify the dependency between functionals defined in terms of nearest neighbor cells. Such methods have been used [5, 31, 32] to establish the limit theory for random measures and functionals in geometric probability. Unfortunately, our functionals depend on both the local density f of points and also on an auxiliary weight g and thus the methods of [5, 31, 32] are not directly applicable and need to be substantially modified.

Notation. Let $\mathcal{X} \subset \mathbb{R}^d$ be locally finite. Given $a \in \mathbb{R}^+$ and $y \in \mathbb{R}^d$, let $a\mathcal{X} := \{ax : x \in \mathcal{X}\}$ and $y + \mathcal{X} := \{y + x : x \in \mathcal{X}\}$. For $x \in \mathbb{R}^d$, let $|x|$ be its Euclidean modulus and for $r > 0$, let $B_r(x)$ denote the Euclidean ball $\{y \in \mathbb{R}^d : |y - x| \leq r\}$. Let ω_d denote the volume of the unit radius ball in $\mathbb{R}^d$ and let $\mathbf{0}$ denote the origin of $\mathbb{R}^d$.

If $B \subset \mathcal{B}(\mathbb{R}^d)$, where $\mathcal{B}(\mathbb{R}^d)$ denotes the Borel subsets of $\mathbb{R}^d$, then $|B|$ denotes its Lebesgue measure. Unless otherwise specified, $A \subset \mathcal{B}(\mathbb{R}^d)$ denotes a compact convex subset of $\mathbb{R}^d$ with non-empty interior.

Given $f \in C(A)$, where $C(A)$ denote the continuous functions on A, and μ a Borel measure on $\mathbb{R}^d$, $\langle f, \mu \rangle$ denotes the integral of f with respect to μ. Let $\mathcal{F} := \{\phi : \mathbb{R}^+ \to \mathbb{R}^+ : \phi \text{ is continuous and } \mathbb{E}\left[\phi^p(\alpha\Gamma_1)\right] < \infty \text{ for all } \alpha > 0 \text{ and some } p > 3\}$, where Γ_1 is a gamma random variable. Let $\mathcal{P}_\tau$ denote a homogeneous Poisson point process on $\mathbb{R}^d$ with intensity τ. When $\tau = 1$, we write $\mathcal{P}$ for $\mathcal{P}_1$. Finally, K denotes a generic positive constant whose value may change from line to line.

2 Main Results

Here we establish two general central limit theorems. The first (Theorem 2.1) provides asymptotic normality for a high dimensional analog of (1.1) while the second (Theorem 2.2) establishes asymptotic normality of (1.1) under general conditions on F, G, and ϕ.

2.1 High dimensional ϕ-divergence based on nearest neighbor cells

Throughout $g : A \to \mathbb{R}^+$ is continuous and $X_1, X_2, ...$ are independent random variables in $\mathbb{R}^d$, $d \geq 1$, with common continuous probability density $f : A \to \mathbb{R}^+$. We assume once and for all that f and g are bounded away from zero and infinity.

For each $X_i, 1 \leq i \leq n$, consider the ball $C_i := C_i(X_1, ..., X_n)$ centered at X_i with radius equal to the distance to the nearest neighbor in the sample $\{X_1, ..., X_n\}$. We use the cells C_i to define high dimensional spacing functionals analogous to the classical one-dimensional functionals (1.1). Define for $1 \leq i \leq n$ the sample spacings

$$D_{i,n} := D_i(X_1, ..., X_n) := \int_{C_i(X_1,...,X_n)} dx$$

and the transformed spacings

$$D^g_{i,n} := D^g_i(X_1, ..., X_n) := \int_{C_i(X_1,...,X_n)} g(x)dx.$$

We will measure the discrepancy between g and the sample density f by comparing the transformed spacings $\{D^g_{i,n},\ 1 \leq i \leq n\}$ with $\{D^f_{i,n},\ 1 \leq i \leq n\}$. Given $\phi \in \mathcal{F}$ consider the 'nearest neighbors ϕ-divergence' functionals:

$$I_\phi(\{D^g_{i,n}\}, \{D^f_{i,n}\}) := \sum_{i=1}^{n} D^f_{i,n} \phi\left(\frac{D^g_{i,n}}{D^f_{i,n}}\right). \tag{2.1}$$

I_ϕ is a measure of the 'distance' between g and f and is a discrete version induced by the balls of the nearest neighbors graph of the ϕ-divergence between g and f, namely $\int_A f(x)\phi\left(\frac{g(x)}{f(x)}\right) dx$.

The ϕ-divergences (2.1), or 'coefficients of divergence', are used heavily in goodness-of-fit tests [37] and are useful in characterizing the amount of information of one distribution contained in another [37, 38]. If P and Q are measures with densities f and g, respectively, then the ϕ-divergence of f and g is a measure of the statistical distinguishability of P and Q. 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) 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, $r > 0$.

If f is unknown, we can replace $D^f_{i,n}$ in (2.1) by its empirical estimate $\hat{D}^f_{i,n} := n^{-1}$, thus providing a high-dimensional analog of the statistic (1.1):

$$N^g_{n,\phi}(X_1, ..., X_n) := \sum_{i=1}^{n} \phi(nD^g_{i,n}). \tag{2.2}$$

We will henceforth call $N^g_{n,\phi}$ the 'nearest neighbors spacing statistic', or 'empirical nearest neighbor ϕ-divergence'. We write $N^\beta_{n,\phi}(X_1, ..., X_n)$ when $g \equiv \beta$, β a constant.

Before stating our main results we need two definitions. For any locally finite $\mathcal{X} := \{x_1, x_2, ...\} \subset \mathbb{R}^d$ and for all $i = 1, 2, ...$, let $C(x_i, \mathcal{X}) := C_i(\mathcal{X})$ be the ball centered at x_i with radius equal to the distance to the nearest neighbor in $\mathcal{X}$. Let $D_i(\mathcal{X}) := \int_{C_i(\mathcal{X})} dx$ and $D^g_i(\mathcal{X}) := \int_{C_i(\mathcal{X})} g(x)dx$ be the generalizations of $D_i(X_1, ..., X_n)$ and $D^g_i(X_1, ..., X_n)$, respectively. For all $t > 0$ and all finite point sets $\mathcal{X}_n := \{x_1, ..., x_n\} \subset \mathbb{R}^d$, let

$$N^g_{t,\phi}(\mathcal{X}_n) := \sum_{i=1}^{n} \phi(tD^g_i(\mathcal{X}_n)). \tag{2.3}$$

Definition 2.1 *(Variance function) For all $\phi \in \mathcal{F}$ and all $\beta > 0$ put*

$$V_\phi(\beta) := \lim_{\lambda \to \infty} \frac{\mathrm{Var}[N^\beta_{\lambda,\phi}(\mathcal{P}_\lambda \cap [0,1]^d)]}{\lambda}. \tag{2.4}$$

The existence of the limit (2.4) is not immediate but is a consequence of the proof of Theorem 2.1 which shows that $y \to \left[\mathbb{E}\left[\phi(\beta|C(\mathbf{0}, \mathcal{P} \cup y)|)\phi(\beta|C(y, \mathcal{P} \cup \mathbf{0})|)\right] - \mathbb{E}\,\phi^2(\beta\Gamma_1)\right]$ is integrable on $\mathbb{R}^d$ for all $\beta > 0$ and that

$$V_\phi(\beta) = \mathbb{E}\left[\phi^2(\beta\Gamma_1)\right] + \int_{\mathbb{R}^d} \left[\mathbb{E}\left[\phi(\beta|C(\mathbf{0}, \mathcal{P} \cup y)|)\phi(\beta|C(y, \mathcal{P} \cup \mathbf{0})|)\right] - \mathbb{E}\,\phi^2(\beta\Gamma_1)\right] dy. \tag{2.5}$$

The existence of the limit in the next definition is a consequence of Lemma 6.1 of [31] and follows because inserting an extra point into the Poisson point set $\mathcal{P}$ induces only local changes in nearest neighbors distances [31].

Definition 2.2 *('Average add-one cost' function) For all $\phi \in \mathcal{F}$ and all $\beta > 0$, let*

$$\Delta_\phi(\beta) := \mathbb{E}\left[\lim_{r \to \infty} [N^\beta_{1,\phi}(\mathcal{P} \cap B_r(\mathbf{0}) \cup \mathbf{0}) - N^\beta_{1,\phi}(\mathcal{P} \cap B_r(\mathbf{0}))]\right].$$

2.2 A general CLT for ϕ-divergences

The volume of the nearest neighbors cell around the origin with respect to $\mathcal{P}$ is a Γ_1 random variable. Given $X_i = x$, $nD^g_{i,n}$ is roughly equal to $g(x)n|C_i|$, which is approximated in distribution by $\frac{g(x)}{f(x)}\Gamma_1$, since for large n, the points around x resemble a Poisson point process with intensity $f(x)$. Thus the summands in (2.2) behave asymptotically like $\phi(\frac{g(X)}{f(X)}\Gamma_1)$. Conditional on $X_i = x$ and neglecting cell dependence, one expects in view of Definition 2.1 and scaling arguments that the limiting variance of $\phi(nD^g_{i,n})$ should equal $V_\phi(\frac{g(x)}{f(x)})$. Furthermore, in the absence of conditioning, one expects that the limiting variance of $n^{-1/2}\sum_{i=1}^n \phi(nD^g_{i,n})$ should behave like a weighted average of $V_\phi(\frac{g(x)}{f(x)})$. This is indeed the case and is the content of the first part of Theorem 2.1 below.

Consider the point measures induced by the empirical nearest neighbor ϕ-divergence:

$$\mu^g_{n,\phi} := \sum_{i=1}^n \phi(nD^g_{i,n})\delta_{X_i}, \tag{2.6}$$

and let $\overline{\mu}^g_{n,\phi} := \mu^g_{n,\phi} - \mathbb{E}\,\mu^g_{n,\phi}$ be their centered versions. The following general CLT, our first main result, establishes convergence of the finite-dimensional distributions of $n^{-1/2}\overline{\mu}^g_{n,\phi}$ as $n \to \infty$ (i.e., the convergence of the m-vector $\langle\langle h_1, n^{-1/2}\overline{\mu}^g_{n,\phi}\rangle, ..., \langle h_m, n^{-1/2}\overline{\mu}^g_{n,\phi}\rangle\rangle$ for all $h_1, ..., h_m$ in $C(A)$) to the finite dimensional distributions of a mean zero finitely additive Gaussian field whose covariance kernel is a weighted average of the variance and add-one cost functions. Henceforth, by 'convergence in law' we shall mean convergence of finite dimensional distributions. Applications, corollaries, and examples follow in section three.

Theorem 2.1 *Let $\phi \in \mathcal{F}$ and $h \in C(A)$. Then as $n \to \infty$*

$$\frac{\mathrm{Var}[\langle h, \mu^g_{n,\phi}\rangle]}{n} \to \int_A h^2(x) V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx - \left(\int_A h(x)\Delta_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx\right)^2 \tag{2.7}$$

and $n^{-1/2}\overline{\mu}^g_{n,\phi}$ converges in law as $n \to \infty$ to a Gaussian field with covariance kernel

$$\int_A h_1(x)h_2(x)V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx - \int_A h_1(x)\Delta_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx \int_A h_2(x)\Delta_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx. \tag{2.8}$$

For certain choices of ϕ we may identify V_ϕ and Δ_ϕ, thus providing limiting distributions of some statistical distances of interest, including information gain and log-likelihood in high dimensions (section 3.5). Notice that if we put

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

then Theorem 2.1 yields the following CLT for the nearest neighbors spacing statistic $N^g_{n,\phi}$:

$$\frac{N^g_{n,\phi}(X_1, ..., X_n) - \mathbb{E}\, N^g_{n,\phi}(X_1, ..., X_n)}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0, \sigma^2_\phi(f,g)).$$

Remarks.

(i) (*Related work*) Bickel and Breiman [6], and subsequently Schilling [39], consider the functionals $N^g_{n,\phi}(X_1, ..., X_n)$ when $\phi(x) = \exp(-x)$. Using the approximation $D^g_{i,n} \approx g(X_i)|C_i(X_1, ..., X_n)|$, they establish a CLT for the empirical process of nearest neighbor distances, but do not consider convergence of the associated random measures $\mu^g_{n,\phi}$, $n \geq 1$. Strong limit theorems for multivariate spacings using general 'shapes' are given by Deheuvels et al. [14]. Penrose [29] finds a CLT for k-nearest neighbor distances and a strong law [28] for the largest nearest-neighbor link. Henze [20] establishes the limit distribution for the maxima of weighted nearest neighbor distances.

(ii)(*Poisson CLT*) For all $\lambda > 0$ let $\mathcal{P}_{\lambda f}$ be a Poisson point process on A with intensity measure $\lambda f : A \to \mathbb{R}^+$. Recalling the general definition (2.3), the Poisson version of the ϕ-divergence (2.2) is

$$N^g_{\lambda,\phi}(\mathcal{P}_{\lambda f}) := \sum_{x \in \mathcal{P}_{\lambda f}} \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f})} g(u)du\right) \tag{2.9}$$

and the Poisson analog of the point measures (2.6) is

$$\mu^g_{\lambda,\phi} := \sum_{x\in\mathcal{P}_{\lambda f}} \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f})} g(u)du\right)\delta_x. \tag{2.10}$$

An integral part of the proof of Theorem 2.1 involves first showing the following Poisson version: If $\phi \in \mathcal{F}$ and $h \in C(A)$, then as $\lambda \to \infty$

$$\frac{\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi}\rangle]}{\lambda} \to \int_A h^2(x) V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx \tag{2.11}$$

and $\lambda^{-1/2}\overline{\mu}^g_{\lambda,\phi}$ converges in law to a Gaussian field with covariance kernel

$$\int_A h_1(x)h_2(x)V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx. \tag{2.12}$$

(iii) (*Limits are distribution free*) [25] The empirical nearest neighbors estimator is a.s. consistent [25, 32] whenever ϕ there is some $p > 4$ such that $\mathbb{E}\left[\phi^p(\alpha\Gamma_1)\right] < \infty$ for all $\alpha > 0$, i.e.,

$$\lim_{n\to\infty} \frac{N^g_{n,\phi}(X_1, ..., X_n)}{n} = \int_A \mathbb{E}\left[\phi\left(\frac{g(x)}{f(x)}\Gamma_1\right)\right] f(x)dx \quad a.s.$$

Thus, combining the above with Theorem 2.1, when $g = f$ the limiting mean, variance, and distribution for $n^{-1}N^g_{n,\phi}(X_1, ..., X_n)$ and $n^{-1/2}\overline{\mu}^g_{n,\phi}$ do not depend on f. In particular $\sigma^2_\phi(f, g) = V_\phi(1)-(\Delta_\phi(1))^2$. Therefore the nearest neighbor functionals are *asymptotically distribution free under the null hypothesis* $g = f$ and have asymptotic variance $V_\phi(1) - (\Delta_\phi(1))^2$. A possible goodness-of-fit test would be to take the density g to be tested, compute the functional $N^g_{n,\phi}(X_1, ..., X_n)$, and see whether the cumulative distribution function is close to the cumulative distribution function of a mean normal random variable with variance $V_\phi(1) - (\Delta_\phi(1))^2$.

(iv) (*Voronoi cells*) Volumes of nearest neighbor cells are computationally attractive and have exponentially decaying correlations. Defining analogous functionals 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, Δ_ϕ is strictly positive, showing that Poissonization contributes extra randomness which manifests itself in a larger limiting variance. When V_ϕ and Δ_ϕ are convex, which is the case when $\phi(x) = x^r$, $r \geq 1$, or when $\phi(x) = x \log x$ (see sections 3.2, 3.3), then $\int_A V_\phi(\frac{g(x)}{f(x)})f(x)dx$ and $\int_A \Delta_\phi(\frac{g(x)}{f(x)})f(x)dx$ are themselves divergences. Thus the limiting variance (2.11) serves as a natural measure of distance. Basic properties of ϕ-divergences [3, 10] imply that the limiting variance over Poisson samples is minimized when $g = f$.

2.3 Asymptotic normality of sum functions of spacings

Classical k-spacings on the line may be represented via analogs of the transformed spacings $D^g_{i,n}$. Viewing spacings on the line in this fashion and making minor modifications to the proof of Theorem 2.1 gives a general CLT for the classical statistics (1.1). This is our second general result and goes as follows.

Given a point set $\mathcal{X} \subset \mathbb{R}^d$ and a convex set $\mathcal{C} \subset \mathbb{R}^d$ containing the origin, define for all $x \in \mathbb{R}^d$ and all $k = 1, 2, ...$

$$d^k_{\mathcal{C}}(x, \mathcal{X}) := \inf_t \{\text{card}(x + t\mathcal{C}) \cap \mathcal{X} \geq k + 1\}.$$

Thus $d^k_{\mathcal{C}}(x, \mathcal{X})$ is the set-induced (asymmetric) distance between x and its kth nearest neighbor in $\mathcal{X}$. For every $x \in \mathcal{X}$, consider the cell $C^k(x, \mathcal{X}) := x + d^k_{\mathcal{C}}(x, \mathcal{X}) \cdot \mathcal{C}$.

Given $\mathcal{X} := \{x_1, x_2, ...\} \subset \mathbb{R}$ and $\mathcal{C} := [0, 1]$, put for all $i = 1, 2, ...$

$$C^k_i := C^k(x_i, \mathcal{X}) \quad \text{and} \quad D^g_i := \int_{C^k_i} g(x)dx$$

and note that C^k_i is the difference of the order statistics $x_{(i+k)} - x_{(i)}$, i.e., corresponds to k-spacings on the line. If g is a probability density with distribution function G then

$$\phi\left(\int_{C^k_i} g(x)dx\right) = \phi(G(x_{(i+k)}) - G(x_{(i)})).$$

If $\mathcal{X}_n := \{x_1, ..., x_n\}$ has order statistics $x_{(1)} < x_{(2)} < ... < x_{(n)}$, then put for all $t > 0$

$$S^g_{t,\phi,k}(\mathcal{X}_n) := \sum_{i=1}^{n-k} \phi(t[G(x_{(i+k)}) - G(x_{(i)})])$$

and when $g \equiv \beta$, write $S^\beta_{t,\phi,k}(\mathcal{X}_n)$. When $\mathcal{X}_n = \{X_1, ..., X_n\}$, where $X_i,\ 1 \leq i \leq n$, are i.i.d.,

$$S^g_{\phi,k}(X_1, ..., X_n) := S^g_{n,\phi,k}(X_1, ..., X_n) =: \sum_{i=1}^{n-k} \phi(n[G(X_{(i+k)})) - G(X_{(i)})])$$

is the classical *sum function of spacings.* Developing the limit theory for $S^g_{n,\phi,k}(X_1, ..., X_n)$ over continuous multivariate samples is important in goodness of fit tests. The passage from uniform to non-uniform densities requires interchanging limits, which has proved to be an obstacle to rigorous analysis.

For all $k = 1, 2, ...$ let $\mathcal{F}_k := \{\phi : \mathbb{R}^+ \to \mathbb{R}^+ : \phi \text{ is continuous and } \mathbb{E}\,[\phi^p(\alpha\Gamma_k)] < \infty \text{ for all } \alpha > 0 \text{ and some } p > 3\}$, where Γ_k is a gamma random variable with parameters k and 1. For all $\phi \in \mathcal{F}_k$, put

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

The convergence in law of the measures $\overline{\mu}^g_{n,\phi,k} := \mu^g_{n,\phi,k} - \mathbb{E}\,\mu^g_{n,\phi,k}$ follows *exactly along the lines of Theorem 2.1.* In this setting we may explicitly identify the variance and add-one functions (V_ϕ and Δ_ϕ) of Theorem 2.1 and therefore the limiting covariance structure as well. This goes as follows.

For all $\beta > 0$ we put

$$V^S_{\phi,k}(\beta) := 2\sum_{l=1}^{k-1} \mathrm{Cov}[\phi(\beta\Gamma_k), \phi(\beta(\Gamma_{k+l} - \Gamma_l))] + \tag{2.13}$$

$$+\ 2k[\mathbb{E}\,\phi(\beta\Gamma_k) - \mathbb{E}\,\phi(\beta\Gamma_{k+1})]\mathbb{E}\,\phi(\beta\Gamma_k) + \mathbb{E}\,\phi^2(\beta\Gamma_k).$$

As will be shown in section five, the analog of the combined identities (2.4) and (2.5) for the spacings functional $S^\beta_{\phi,k}$ is

$$\lim_{\lambda\to\infty} \frac{\mathrm{Var}[S^\beta_{\lambda,\phi,k}(\mathcal{P}_\lambda \cap [0,1])]}{\lambda} = V^S_{\phi,k}(\beta). \tag{2.14}$$

Moreover, putting for all $\beta > 0$

$$\Delta^S_{\phi,k}(\beta) := (k+1)\mathbb{E}\,[\phi(\beta\Gamma_k)] - k\mathbb{E}\,[\phi(\beta\Gamma_{k+1})], \tag{2.15}$$

and letting $\mathcal{P}$ be a homogeneous Poisson point process on $\mathbb{R}$ with intensity 1, it is straightforward to show for all $k = 1, 2, ...$ and $\beta > 0$ that

$$\mathbb{E}\,[\lim_{r\to\infty} (S^\beta_{1,\phi,k}(\mathcal{P} \cap B_r(\mathbf{0}) \cup \mathbf{0}) - S^\beta_{1,\phi,k}(\mathcal{P} \cap B_r(\mathbf{0}))] = \Delta^S_{\phi,k}(\beta).$$

Using the functions $V^S_{\phi,k}$ and $\Delta^S_{\phi,k}$ instead of V_ϕ and Δ_ϕ in Theorem 2.1, we state our second main result, a general CLT for sum functions of spacings. Applications are given in section three.

Theorem 2.2 *(Gaussian limit for sum functions of spacings) Let* $\phi \in \mathcal{F}_k$ *and* $h \in C(A)$, $A := [a,b]$. *Then as* $n \to \infty$

$$\frac{\mathrm{Var}[\langle h, \mu^g_{n,\phi,k}\rangle]}{n} \to \int_A h^2(x) V^S_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx - \left(\int_A \Delta^S_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx\right)^2$$

while $n^{-1/2}\overline{\mu}^g_{n,\phi,k}$ *converges in law to a Gaussian field with covariance kernel*

$$\int_A h_1(x)h_2(x)V^S_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx \ - \int_A h_1(x)\Delta^S_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx \int_A h_2(x)\Delta^S_{\phi,k}\left(\frac{g(x)}{f(x)}\right) f(x)dx.$$

Remarks. (i) Darling [13] 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 [21], as well as earlier work of Cressie [9], who proves asymptotic normality (but not convergence in law) 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 [22] prove asymptotic normality of $S^g_{n,\phi,k}(X_1, ..., X_n)$ under 'somewhat stringent conditions' on f and g. For non-uniform samples the asymptotics of $S^g_{n,\phi,k}(X_1, ..., X_n)$ have been widely studied under the assumption that G runs through a sequence of alternatives G_n approaching the uniform distribution; see Hall [19] and del Pino [34]. Theorem 2.2 does not impose this restriction. Khashimov [26] establishes asymptotic normality of $S^1_{n,\phi,k}(X_1, ..., X_n)$ under rather technical differentiability conditions on ϕ and f.

(ii) When $\phi \in \mathcal{F}_k$, $S^g_{n,\phi,k}$ is a consistent estimator of the ϕ-divergence [25, 32] in the sense that

$$\lim_{n\to\infty} \frac{S^g_{n,\phi,k}(X_1, ..., X_n)}{n} = \int_A \mathbb{E}\left[\phi\left(\frac{g(x)}{f(x)}\Gamma_k\right)\right] f(x)dx \quad a.s.$$

Analogous consistency results hold for non-overlapping k-spacings [40].

3 Applications

For many tests involving goodness-of-fit (Dudewicz et al. [15], Blumenthal [7], Cressie [9], Holst and Rao [22], delPino [34], Weiss [44]) and parametric estimation (Ghosh and Jammalamadaka [18]) it is important to know the asymptotic distribution of $S^g_{n,\phi,k}(X_1, ..., X_n)$ for arbitrary g and f and for various choices of ϕ. The following provides some illustrative examples.

3.1 Limit theory for logarithms of spacings

If $X_1, ..., X_n$ are i.i.d. with density f, let

$$S^g_{\log,k}(X_1, ..., X_n) := \sum_{i=1}^{n-k} \log(n[G(X_{(i+k)})) - G(X_{(i)})])$$

denote the sum of the logarithmic k-spacings, where $\log x$ denotes the natural logarithm of x. Setting $\phi(x) = \log x$ in Theorem 2.2 yields a CLT for logarithms of k-spacings as follows.

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

By Cressie [9] and Holst [21],

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

Since $\mathbb{E}\,[\log\Gamma_k] = \psi(k)$, we also have

$$2k(\mathbb{E}\,\log\Gamma_k - \mathbb{E}\,\log\Gamma_{k+1})\mathbb{E}\,\log\Gamma_k = -2\psi(k).$$

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

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

By(2.15) we have $\Delta^S_{\log,k}(1) = (k+1)\psi(k) - k\psi(k+1) = \psi(k) - 1$. Letting $U_1, U_2, ...$ be i.i.d. with the uniform density on $[0,1]$ and putting $\tau_k := (2k^2 - 2k + 1)\psi'(k) - 2k + 1$ and $g \equiv 1$, Theorem 2.2 yields:

Corollary 3.1 *(Gaussian limits for logarithmic k-spacings, uniform densities) As $n \to \infty$,*

$$\frac{\mathrm{Var}[S^1_{\log,k}(U_1, ..., U_n)]}{n} \to \tau_k,$$

and

$$\frac{S^1_{\log,k}(U_1, ..., U_n) - \mathbb{E}\,S^1_{\log,k}(U_1, ..., U_n)}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0, \tau_k)$$

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

$$V^S_{\log,k}(1)\int_0^1 h_1(x)h_2(x)dx - (\psi(k) - 1)^2\int_0^1 h_1(x)dx\int_0^1 h_2(x)dx,$$

where $V^S_{\log,k}(1)$ is as in (3.1).

By using simple relations such as

$$\mathrm{Cov}(\log\beta\Gamma_k, \log\beta(\Gamma_{k+l} - \Gamma_l)) = \mathrm{Cov}(\log\Gamma_k, \log(\Gamma_{k+l} - \Gamma_l)),\ \beta > 0,$$

it follows that $V^S_{\log,k}(\beta) = V^S_{\log,k}(1) + \log^2\beta + 2\log\beta(\psi(k) - 1)$ and $\Delta^S_{\log,k}(\beta) = \Delta^S_{\log,k}(1) + \log\beta$. Substituting this into Theorem 2.2 and re-arranging terms yields the following generalization of Corollary 3.1:

Corollary 3.2 *(CLT limits for logarithmic k-spacings, general densities) Let $X, X_1, X_2, ...$ be i.i.d. with density f on $[0,1]$. As $n \to \infty$*

$$\frac{\mathrm{Var}[S^g_{\log,k}(X_1, ..., X_n)]}{n} \to \tau_k + \mathrm{Var}\left[\log(\frac{f(X)}{g(X)})\right]$$

and

$$\frac{S^g_{\log,k}(X_1, ..., X_n) - \mathbb{E}\,S^g_{\log,k}(X_1, ..., X_n)}{n^{1/2}} \xrightarrow{\mathcal{D}} N\left(0,\ \tau_k + \mathrm{Var}\left[\log(\frac{f(X)}{g(X)})\right]\right).$$

Remarks. When X_i are i.i.d. uniform on $[0,1]$, and when $g \equiv 1$, then the CLT for $S^g_{\log,k}(X_1,...,X_n)$ was established by Darling (sect. 7 of [13]) for $k = 1$ and later by Holst [21] and Cressie [9] for general k. When the X_i have a step density then Cressie shows asymptotic normality of $S^g_{\log,k}(X_1,...,X_n)$ including cases when $k \to \infty$. Czekala (Thm. 1 of [12]) apparently re-discovered Cressie's result. Shao and Hahn [41] 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 [7]), proves Corollary 3.2 for densities f satisfying special conditions. Corollary 3.2 extends all of these results to general f and g, resolving a conjecture of Darling ([13], p. 249) affirmatively.

3.2 Information gain of order r

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

$$S^g_r(X_1,...,X_n) := \sum_{i=1}^{n-1} \left(n[G(X_{(i+1)})) - G(X_{(i)})]\right)^r .$$

Denote the associated point measures by

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

Let $V_r := -2r\Gamma^2(r+1) + \Gamma(2r+1)$ and $D_r := \Gamma(r+1)(1-r)$. It is a simple matter to verify via (2.13) and (2.15), respectively, that for all $\beta > 0$

$$V^S_r(\beta) := \lim_{\lambda\to\infty} \frac{\mathrm{Var}[S^\beta_r(\mathcal{P}_\lambda \cap [0,1])]}{\lambda} = V_r \beta^{2r}$$

and

$$\Delta^S_r(\beta) := 2\mathbb{E}\left[\phi(\beta\Gamma_1)\right] - \mathbb{E}\left[\phi(\beta\Gamma_2)\right] = D_r\beta^r.$$

Put

$$\sigma^2_r(f,g) := V_r \int_A \left(\frac{g(x)}{f(x)}\right)^{2r} f(x)dx - D^2_r \left(\int_A \left(\frac{g(x)}{f(x)}\right)^r f(x)dx\right)^2 .$$

Theorem 2.2 yields:

Corollary 3.3 *(Gaussian limits for information gain, general densities) Let $X_1, X_2, ...$ be i.i.d. with density f on $A := [a,b]$. As $n \to \infty$*

$$\frac{\mathrm{Var}[S^g_r(X_1,...,X_n)]}{n} \to \sigma^2_r(f,g)$$

and

$$\frac{S_r^g(X_1, ..., X_n) - \mathbb{E}\, S_r^g(X_1, ..., X_n)}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0, \sigma_r^2(f, g)),$$

while $n^{-1/2}\overline{\mu}_{n,r}^g$ *converges in law to a Gaussian field with covariance kernel*

$$V_r \int_A h_1(x) h_2(x) \left(\frac{g(x)}{f(x)}\right)^{2r} f(x)dx - D_r^2 \int_A h_1(x) \left(\frac{g(x)}{f(x)}\right)^r f(x)dx \int_A h_2(x) \left(\frac{g(x)}{f(x)}\right)^r f(x)dx.$$

Remarks. It is easy to verify using [5] that $\sigma_r^2(f, g) > 0$ except when $r = 1$. Corollary 3.3 extends upon the CLTs of Darling [13] (uniform case) and Weiss [44]. The latter assumes that GF^{-1} has a step function derivative and incorrectly asserts that his results hold in more general situations (cf. p. 417 of Pyke [35]). Moran [27] proved a CLT for the functional S_2 over uniform random variables.

3.3 Limit theory for log-likelihood ratio

Let $\phi(x) = x \log x$. We will use Theorem 2.2 to develop the limit theory for the log-likelihood statistic

$$S_\phi^g(X_1, ..., X_n) := \sum_{i=1}^{n-1} \phi(n[G(X_{(i+1)})) - G(X_{(i)})])$$

as well as for the log-likelihood point measure

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

Using the relations

$$\mathbb{E}\,[(\beta\Gamma_1 \log(\beta\Gamma_1))] = \beta \log \beta + \beta(1 - \gamma) \text{ and } \mathbb{E}\,[(\beta\Gamma_2 \log(\beta\Gamma_2))] = 2\beta \log \beta + \beta(3 - 2\gamma),$$

as well as

$$\mathbb{E}\,[(\beta\Gamma_1 \log(\beta\Gamma_1))^2] = \beta^2[2(\log \beta)^2 + (6 - 4\gamma) \log \beta + 2 + \pi^2/3 - 6\gamma + 2\gamma^2],$$

it is a simple matter to verify via (2.13) and (2.15), respectively, that for all $\beta > 0$

$$V_\phi^S(\beta) := \lim_{\lambda \to \infty} \frac{\mathrm{Var}[S_{\lambda,\phi}^\beta(\mathcal{P}_\lambda \cap [0, 1])]}{\lambda} = (\frac{\pi^2}{3} - 2)\beta^2$$

and that

$$\Delta_\phi^S(\beta) := 2\mathbb{E}\, \phi(\beta\Gamma_1) - \mathbb{E}\, \phi(\beta\Gamma_2) = -\beta.$$

Put

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

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

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

Using the above values for $V_\phi^S(\beta)$, $\Delta_\phi^S(\beta)$, $\sigma_\phi^2(f,g)$, and applying Theorem 2.2 for $\phi(x) = x \log x$ yields:

Corollary 3.4 *(Gaussian limit for log likelihood, general densities) Let $X_1, X_2, ...$ be i.i.d. with density f on A, $A := [a,b]$. As $n \to \infty$*

$$\frac{\mathrm{Var}[S_\phi^g(X_1,...,X_n)]}{n} \to \sigma_\phi^2(f,g)$$

and

$$\frac{S_\phi^g(X_1,...,X_n) - \mathbb{E}\, S_\phi^g(X_1,...,X_n)}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0, \sigma_\phi^2(f,g))$$

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

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

Remarks. Corollary 3.4 extends the results of Gebert and Kale [17], who assume uniformity of $X_i,\ i \geq 1$, and Czekala (Thm. 2 of [12]), who assumes that $X_i,\ i \geq 1$, have a step density. van Es [43] establishes asymptotic normality for S_ϕ^g whenever $k,\ n \to \infty$, $k = o(n^{1/2})$, and $f : A \to \mathbb{R}^+$ is Lipschitz.

3.4 Number of pairs of sample points within a fixed distance

Instead of considering spacing functionals defined with respect to the sample points, we now define a functional using cells of *fixed* radius t. Thus, for all $t > 0$ and all locally finite point sets $\mathcal{X} \subset \mathbb{R}^d$ and $x \in \mathcal{X}$ we set $\Phi^t(x, \mathcal{X}) := \mathrm{card}(B_t(x) \cap \mathcal{X})$ and for all $\lambda > 0$ we define $\Phi_\lambda^t(x, \mathcal{X}) := \Phi^t(\lambda^{1/d}x, \lambda^{1/d}\mathcal{X})$. The functional

$$H^t(\mathcal{X}) := \frac{1}{2}\sum_{x \in \mathcal{X}} \Phi^t(x, \mathcal{X})$$

counts the total number of pairs of points in $\mathcal{X}$ distant t from each other. Define the scaled version $H_\lambda^t(\mathcal{X}) := H^t(\lambda^{1/d}\mathcal{X})$. Given i.i.d. random variables $X_1, ..., X_n$ with continuous density $f : A \to \mathbb{R}^+$, where A is a compact convex subset of $\mathbb{R}^d$, we seek the asymptotic distribution of $H_n^t(X_1, ..., X_n)$ as well as that of the point measure

$$\mu_n^t := \sum_{i=1}^n \Phi_n^t(X_i, \{X_i\}_{i=1}^n)\delta_{X_i}.$$

For all $h \in C(A)$ and $t > 0$ put

$$\sigma_t^2(h) := v_t \int_A h^2(x)dx + v_t^2 \left(2 \int_A h^3(x)dx - (\int_A h^2(x)dx)^2 \right),$$

where $v_t := t^d \omega_d$ is the volume of a d-dimensional ball of radius t. The following CLT is obtained by modifying the proof of Theorem 2.1; details are in section six.

Corollary 3.5 *(Gaussian limit for the number of pairs of points within distance t) Let* $X_1, X_2, ...$ *be i.i.d. with density* f *on* A*. As* $n \to \infty$

$$\frac{\mathrm{Var}[H_n^t(X_1, ..., X_n)]}{n} \to \sigma_t^2(f) \tag{3.2}$$

and

$$\frac{H_n^t(X_1, ..., X_n) - \mathbb{E}\, H_n^t(X_1, ..., X_n)}{n^{1/2}} \xrightarrow{\mathcal{D}} N(0, \sigma_t^2(f))$$

while $n^{-1/2}\overline{\mu}_{n,t}^g$ *converges in law to a Gaussian field with covariance kernel*

$$\int_A h_1(x)h_2(x)[v_t f^2(x) + 2v_t^2 f^3(x)]dx - v_t^2 \int_A h_1(x)f^2(x)dx \int_A h_2(x)f^2(x)dx.$$

Remarks. The functional H^t has been considered by various authors. L'Ecuyer et al. [16] considers H^t from the point of view of multidimensional goodness-of-fit tests, but restricts attention to uniform samples. Penrose [30] (Ch.4) proves that the finite dimensional distributions of the process $H_n^t(X_1, ..., X_n),\ t > 0$, converge to those of a Gaussian process.

3.5 Information gain and log-likelihood in high dimensions

For differentiable ϕ with $\phi(0) = 0$, the limit $V_\phi(\beta),\ \beta > 0$, in Definition 2.1 can be easily computed. We will use this expression for $V_\phi(\beta)$ in conjunction with the Poisson CLT (2.11 - 2.12) to develop the limit theory for the information gain and log-likelihood functional on Poisson point sets in all dimensions.

For all $s, t, u \in \mathbb{R}^+$, 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

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

The following is proved in section seven.

Proposition 3.1 *(evaluation of the variance function* V_ϕ*) Assume* $\phi \in \mathcal{F}$ *is differentiable and* $\phi(0) = 0$*. Then for all* $\beta > 0$*,*

$$V_\phi(\beta) = \mathbb{E}\,[\phi^2(\beta\Gamma_1)] +$$

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

provided that the integral exists.

In what follows we recall that $\mathcal{P}_{\lambda f}$ is a Poisson point process with intensity λf on the compact convex subset A of $\mathbb{R}^d$.

3.5.1 Information gain of order r

When $\phi(x) = x^r, \ r > 0$, $N^g_{\lambda,\phi}(\mathcal{P}_{\lambda f})$ defined by (2.9) yields Rényi's information gain (I-divergence) of order r. Since $\phi(x) = x^r$ satisfies the conditions of Proposition 3.1, the following is immediate. For all $r > 0$, define the constant

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

Lemma 3.1 *For all $\beta > 0$ and for $\phi(x) = x^r, \ r > 0$, we have*

$$V_\phi(\beta) = \beta^{2r} \left[\mathbb{E}\left[(\Gamma_1)^{2r}\right] + K_r \right].$$

Combining (2.11 - 2.12) and Lemma 3.1 we easily deduce asymptotic normality for the information gain $N^g_{\lambda,\phi}(\mathcal{P}_{\lambda f})$ and the associated point measures $\mu^g_{\lambda,\phi}$ defined by (2.10). Let

$$\sigma^2_\phi(f,g) := \left[\mathbb{E}\left[(\Gamma_1)^{2r}\right] + K_r\right] \int_A \left(\frac{f(x)}{g(x)}\right)^{2r} f(x) dx.$$

Corollary 3.6 *Let $\phi(x) = x^r, \ r > 0$. Then as $\lambda \to \infty$*

$$\frac{\mathrm{Var}[N^g_{\lambda,\phi}(\mathcal{P}_{\lambda f})]}{\lambda} \to \sigma^2_\phi(f,g)$$

and

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

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

$$[\mathbb{E}\left[(\Gamma_1)^{2r}\right] + K_r] \int_A h_1(x) h_2(x) \left(\frac{f(x)}{g(x)}\right)^{2r} f(x) dx.$$

3.5.2 Log likelihood

When $\phi(x) = x \log x$, $N^g_{\lambda,\phi}(\mathcal{P}_{\lambda f})$ defined by (2.9) yields the log-likelihood statistic. To apply Theorem 2.1 , we define the constants

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

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

and

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

The following is an easy consequence of Proposition 3.1.

Lemma 3.2 *For all $\beta > 0$ and for $\phi(x) = x \log x$,*

$$V_\phi(\beta) = \beta^2 \left[2 + \pi/3 - 6\gamma + 2\gamma^2 + I_1\right] + \beta^2 \log \beta [6 - 4\gamma + 2I_2] + (\beta \log \beta)^2 [2 + I_3].$$

Using (2.11 - 2.12) we may deduce asymptotic normality for the log-likelihood functional $N_\phi^g(\mathcal{P}_{\lambda f})$ and for the associated point measures $\mu_{\lambda,\phi}^g$ defined by (2.10). Let

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

where $K_1 := 2 + \frac{\pi}{3} - 6\gamma + 2\gamma^2 + I_1$, $K_2 := 6 - 4\gamma + 2I_2$, and $K_3 := 2 + I_3$.

Corollary 3.7 *Let $\phi(x) = x \log x$. Then as $\lambda \to \infty$*

$$\frac{\mathrm{Var}[N_{\lambda,\phi}^g(\mathcal{P}_{\lambda f})]}{\lambda} \to \sigma_\phi^2(f,g)$$

and

$$\frac{N_{\lambda,\phi}^g(\mathcal{P}_{\lambda f}) - \mathbb{E}\, N_{\lambda,\phi}^g(\mathcal{P}_{\lambda f})}{\lambda^{1/2}} \xrightarrow{\mathcal{D}} N(0, \sigma_\phi^2(f,g))$$

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

$$\int_A h_1(x) h_2(x) \left(\frac{g(x)}{f(x)}\right)^2 \left[K_1 + K_2 \log\left(\frac{g(x)}{f(x)}\right) + K_3 \left(\log \frac{g(x)}{f(x)}\right)^2\right] f(x) dx.$$

4 Proof of Theorem 2.1

4.1 Stabilization

The proof of Theorem 2.1 relies heavily upon the fact that correlations between nearest neighbor cells decay exponentially fast with the distance between cells. This fact is neither surprising nor difficult to establish. However, by quantifying the dependency between cells through the notion of *stabilization* [5, 31, 32] we obtain explicit formulas for limiting variances and covariances. The proofs are involved but we believe that this cannot be avoided. We have tried to make the proofs as self-contained as possible, but when there is obvious overlap with existing results we appeal to those instead.

We now recall the notion of stabilization. Recall that for all $\lambda > 0$, $\mathcal{P}_{\lambda f}$ is a Poisson point process on A with intensity measure $\lambda f : A \to \mathbb{R}^+$. Define for all locally finite point sets $\mathcal{X} \subset \mathbb{R}^d, x \in \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}$ the 'g content of the cell around x with respect to $\mathcal{X}$' by

$$I^g(x, \mathcal{X}) := \int_{C(x,\mathcal{X})} g(u)du.$$

I^g stabilizes since the following property is satisfied: Given $f, g : A \to \mathbb{R}^+$ bounded away from zero and infinity, for all $x \in A$ and all $\lambda > 1$ there exists a random variable $R := R_x := R^{f,g}_{x,\lambda}$, called a *radius of stabilization for K^g with respect to $\mathcal{P}_{\lambda f}$ at x* such that

$$I^g(\lambda^{1/d}x, \lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_R(\lambda^{1/d}x) \cup \mathcal{A}) = I^g(\lambda^{1/d}x, \lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_R(\lambda^{1/d}x)) \tag{4.1}$$

for all finite $\mathcal{A} \subset \lambda^{1/d}A \setminus B_R(\lambda^{1/d}x)$. I^g is said to stabilize since changes in point configurations distant greater than R do not affect the value of I^g.

The proof of Theorem 2.1 consists of the following three steps. We fix $\phi \in \mathcal{F}, g : A \to \mathbb{R}^+,\ g$ continuous, and $f : A \to \mathbb{R}^+$ a continuous density.

Step 1. (Poissonization of variance) Let

$$\mu^g_{\lambda,\phi} := \sum_{x \in \mathcal{P}_{\lambda f}} \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f})} g(u)du\right)\delta_x := \sum_{x \in \mathcal{P}_{\lambda f}} \phi\left(\lambda K^g(x, \mathcal{P}_{\lambda f})\right)\delta_x. \tag{4.2}$$

We use stabilization (4.1) together with properties of the one and two point correlation functions for I^g to show for all $h \in C(A)$,

$$\lim_{\lambda\to\infty} \frac{\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi}\rangle]}{\lambda} = \int_A h^2(x) V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx. \tag{4.3}$$

Step 2. (de-Poissonization) We de-Poissonize the limit (4.3) and show the limit (2.7), namely

$$\lim_{n\to\infty} \frac{\mathrm{Var}[\langle h, \mu^g_{n,\phi}\rangle]}{n} = \int_A h^2(x) V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx - \left(\int_A h(x)\Delta_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx\right)^2. \tag{4.4}$$

Step 3. (Completion of proof of Theorem 2.1) By appealing to general results and methods of [33] and by using the stabilization of I^g, we obtain for all $h \in C(A)$ as $\lambda \to \infty$:

$$\lambda^{-1/2}\langle h, \overline{\mu}^g_{\lambda,\phi}\rangle \xrightarrow{\mathcal{D}} N\left(0, \int_A h^2(x) V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx\right). \tag{4.5}$$

We then establish Theorem 2.1 by de-Poissonizing the limit (4.5).

Variations of the above approach have been used in [5, 31] in the context of proving CLTs for random measures in geometric probability. Unfortunately the functional I^g depends on the auxiliary function g and thus the methods of [5, 31] are not directly applicable and need to be substantially modified.

4.2 Step 1: Variance convergence over Poisson samples

To simplify the notation, for all $\lambda > 0$, $x \in A$, $g : A \to \mathbb{R}^+$, and locally finite $\mathcal{X} \subset A$, define

$$\Phi_\lambda^g(x, \mathcal{X}) := \phi\left(\lambda \int_{C(x,\mathcal{X})} g(u)du\right) := \phi\left(\lambda I^g(x, \mathcal{X})\right)$$

so that (4.2) becomes

$$\mu_{\lambda,\phi}^g = \sum_{x \in \mathcal{P}_{\lambda f}} \Phi_\lambda^g(x, \mathcal{P}_{\lambda f})\delta_x.$$

Since I^g stabilizes it is clear that Φ^g *stabilizes as well* and has the same radius of stabilization.

Define also

$$\Phi_\lambda^{g(x)}(y, \mathcal{X}) := \phi\left(\lambda \int_{C(y,\mathcal{X})} g(x)du\right).$$

We refer to $\Phi_\lambda^{g(x)}(y, \mathcal{P}_{\lambda f(x)})$ as a first order approximation to $\Phi_\lambda^g(y, \mathcal{P}_{\lambda f})$ in that f and g are respectively approximated by the scalars $f(x)$ and $g(x)$. When x is close to y, the continuity of ϕ and g ensures that this is a good approximation. We approximate $\mu_{\lambda,\phi}^g$ by the first order approximation $\sum_{x \in \mathcal{P}_{\lambda f}} \Phi_\lambda^{g(x)}(x, \mathcal{P}_{\lambda f(x)})\delta_x$ as follows.

For all $\lambda > 0$, let $\mathcal{P}'_{\lambda f}$ be a Poisson point process equidistributed with and independent of $\mathcal{P}_{\lambda f}$, i.e., $\mathcal{P}'_{\lambda f}$ is a copy of $\mathcal{P}_{\lambda f}$. For all $\lambda > 0$ and $x \in A$ we introduce two auxiliary homogeneous independent Poisson point processes $\mathcal{P}_{\lambda f(x)}$ and $\mathcal{P}'_{\lambda f(x)}$ such that:

- $\mathcal{P}_{\lambda f(x)}, \mathcal{P}'_{\lambda f(x)}$ have *constant* intensity $\lambda f(x)$ on A,
- $\mathcal{P}_{\lambda f}$ and $\mathcal{P}_{\lambda f(x)}$ are coupled in the sense that for any $B \in \mathcal{B}(\mathbb{R}^d)$,

  $$P\left[\mathcal{P}_{\lambda f}(B) \neq \mathcal{P}_{\lambda f(x)}(B)\right] \leq \lambda \int_B |f(y) - f(x)|\, dy, \tag{4.6}$$

  and the same is true for $\mathcal{P}'_{\lambda f}$ and $\mathcal{P}'_{\lambda f(x)}$.

To show (4.3), we will show for all $h \in C(A)$ that

$$\lambda^{-1}\mathrm{Var}[\langle h, \mu_{\lambda,\phi}^g\rangle] = \int_A h^2(x)E[(\Phi_\lambda^g(x, \mathcal{P}_{\lambda f}))^2]f(x)dx \tag{4.7}$$

$$+\lambda \int\int_{A\times A} h(x)h(y)\mathbb{E}\left[\Phi_\lambda^g(x,\mathcal{P}_{\lambda f}\cup y)\Phi_\lambda^g(y,\mathcal{P}_{\lambda f}\cup x) - \Phi_\lambda^g(x,\mathcal{P}_{\lambda f})\Phi_\lambda^g(y,\mathcal{P'}_{\lambda f})\right] f(x)f(y)dxdy$$

is approximated by its 'first order version'

$$\int_A h^2(x)E[(\Phi_\lambda^{g(x)}(x,\mathcal{P}_{\lambda f(x)}))^2]f(x)dx + \lambda\int\int_{A\times\mathbb{R}^d} h(x)h(y)\mathbb{E}\;[\;...\;]\;f(x)f(y)dxdy \tag{4.8}$$

where $\mathbb{E}\,[...]$ is shorthand for

$$\mathbb{E}\left[\Phi_\lambda^{g(x)}(x,\mathcal{P}_{\lambda f(x)}\cup y)\Phi_\lambda^{g(x)}(y,\mathcal{P}_{\lambda f(x)}\cup x) - \Phi_\lambda^{g(x)}(x,\mathcal{P}_{\lambda f(x)})\Phi_\lambda^{g(x)}(y,\mathcal{P'}_{\lambda f(x)})\right].$$

Then we show in the large λ limit that (4.8) reduces to $\int_A h^2(x)V_\phi(\frac{g(x)}{f(x)})f(x)dx$.

The approximation of (4.7) by (4.8) would be easy if all weights $\Phi_\lambda(\cdot,\cdot)$ were independent. If x and y are far enough apart in $A\times A$, then the corresponding weights are independent and the integrand vanishes in this case. However for x and y close, the situation is much different and requires some careful analysis which is inspired by (but does not follow from) the stabilization methods of [5].

We will measure the closeness of (4.7) and (4.8) in terms of the one and two point correlation functions for Φ_λ^g, defined for all $x,y\in\mathbb{R}^d$ by

$$q_\lambda^g(x) := \mathbb{E}\,[(\Phi_\lambda^g(x,\mathcal{P}_{\lambda f}))^2]$$

and

$$c_\lambda^g(x,y) := \mathbb{E}\,[\Phi_\lambda^g(x,\mathcal{P}_{\lambda f}\cup y)\Phi_\lambda^g(y,\mathcal{P}_{\lambda f}\cup x) - \Phi_\lambda^g(x,\mathcal{P}_{\lambda f})\Phi_\lambda^g(y,\mathcal{P'}_{\lambda f})]\,.$$

The corresponding correlation functions for the first order approximations are:

$$\tilde{q}_\lambda^g(x) := \mathbb{E}[(\Phi_\lambda^{g(x)}(x,\mathcal{P}_{\lambda f(x)}))^2]$$

and

$$\tilde{c}_\lambda^g(x,y) := \mathbb{E}\left[\Phi_\lambda^{g(x)}(x,\mathcal{P}_{\lambda f(x)}\cup y)\Phi_\lambda^{g(x)}(y,\mathcal{P}_{\lambda f(x)}\cup x) - \Phi_\lambda^{g(x)}(x,\mathcal{P}_{\lambda f(x)})\Phi_\lambda^{g(x)}(y,\mathcal{P'}_{\lambda f(x)})\right].$$

The shorthand notation $\tilde{q}$ and $\tilde{c}$ signifies that we approximate $q_\lambda^g(x)$ and $c_\lambda^g(x,y)$ by using $f(x)$ and $g(x)$ in place of f and g, respectively.

Write simply $q^g(x)$ and $c^g(x,y)$ for $q_1^g(x)$ and $c_1^g(x,y)$, respectively, and similarly for $\tilde{q}^g$ and $\tilde{c}^g$. Since for all $\lambda>0$, $\lambda\int_{C(x,\mathcal{P}_{\lambda f(x)})} du \overset{\mathcal{D}}{=} \int_{C(x,\mathcal{P}_{f(x)})} du$, it follows for all $x,y\in\mathbb{R}^d$ and all $\lambda>0$, that

$$\tilde{q}_\lambda^g(x) = \tilde{q}^g(x) \;\text{ and }\; \tilde{c}_\lambda^g(x,x+y) = \tilde{c}^g(x,x+\lambda^{1/d}y).$$

Notice that for all $x, y \in \mathbb{R}^d$, $\Phi_\lambda^{g(x)}(y) \stackrel{\mathcal{D}}{=} \phi(\frac{g(x)}{f(x)}\Gamma_1)$. By the bounds on g and f, the assumption $\phi \in \mathcal{F}$, and Cauchy-Schwarz, we have

$$\sup_{x,y\in\mathbb{R}^d} \sup_{\lambda>0} [|c_\lambda^g(x,y)|, |\tilde{c}_\lambda^g(x,y)|] \leq K. \tag{4.9}$$

We next describe properties of the correlation functions resulting from stabilization of I^g. The first result describes the exponential decay of correlations.

Lemma 4.1 *For all $x, y \in \mathbb{R}^d$, the correlations $c_\lambda^g(x,y)$ and $\tilde{c}_\lambda^g(x,y)$ satisfy the bounds:*

$$|c_\lambda^g(x,y)| \leq K \exp(-K\lambda^{1/d}|x-y|) \tag{4.10}$$

and

$$|\tilde{c}_\lambda^g(x,y)| \leq K \exp(-K\lambda^{1/d}|x-y|). \tag{4.11}$$

Proof. We prove (4.10); the proof of (4.11) is identical. Given $x, y \in A$, let R_x and R_y be the radii of stabilization for Φ^g with respect to $\mathcal{P}_{\lambda f}$ at x and y respectively. R_x and R_y are both less than $\delta := \lambda^{1/d}|x-y|/2$ except on a set with probability at most $2\exp(-K\lambda^{1/d}|x-y|)$. Let E denote the event that R_x and R_y are both less than $\lambda^{1/d}|x-y|/2$. On E the sets $\lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_{R_x}(\lambda^{1/d}x)$ and $\lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_{R_y}(\lambda^{1/d}y)$ do not intersect and thus on E we have $\Phi_\lambda^g(x, \mathcal{P}_{\lambda f} \cup y) = \Phi_\lambda^g(x, \mathcal{P}_{\lambda f})$ and similarly $\Phi_\lambda^g(y, \mathcal{P}_{\lambda f} \cup x) = \Phi_\lambda^g(y, \mathcal{P}_{\lambda f})$.

It therefore follows that there is a constant K such that for all $x, y \in \mathbb{R}^d$,

$$\left|\mathbb{E}\left[\Phi_\lambda^g(x, \mathcal{P}_{\lambda f} \cup y)\Phi_\lambda^g(y, \mathcal{P}_{\lambda f} \cup x) - \Phi^g(\lambda^{1/d}x, \lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_\delta(\lambda^{1/d}x))\Phi^g(\lambda^{1/d}y, \lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_\delta(\lambda^{1/d}y))\right]\right|$$

$$\leq K(P[E^c])^{1/2},$$

by Hölder's inequality. Using independence in the second expectation, the bound

$$\sup_{x\in\mathbb{R}^d} \left|\mathbb{E}\left[\Phi^g(\lambda^{1/d}x, \lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_\delta(\lambda^{1/d}x)) - \Phi_\lambda^g(x, \mathcal{P}_{\lambda f})\right]\right| \leq K(P[E^c])^{1/2}$$

and

$$\sup_{x\in\mathbb{R}^d} \left|\mathbb{E}\left[\Phi^g(\lambda^{1/d}x, \lambda^{1/d}\mathcal{P}_{\lambda f} \cap B_\delta(\lambda^{1/d}x)\right]\right| < \infty$$

yields for all $x, y \in \mathbb{R}^d$

$$\left|\mathbb{E}\left[\Phi_\lambda^g(x, \mathcal{P}_{\lambda f} \cup y)\Phi_\lambda^g(y, \mathcal{P}_{\lambda f} \cup x)\right] - \mathbb{E}\left[\Phi^g(\lambda^{1/d}x, \lambda^{1/d}\mathcal{P}_{\lambda f})\right]\mathbb{E}\left[\Phi^g(\lambda^{1/d}y, \lambda^{1/d}\mathcal{P}_{\lambda f})\right]\right| \leq K(P[E^c])^{1/2}.$$

This yields the desired bound (4.10). □

Compactness of A and the continuity of f and h implies uniform continuity, so we fix *moduli of continuity* $t_f, t_h : \mathbb{R}^+ \to \mathbb{R}^+$ such that for any $x, y \in A : |x - y| \leq \delta$, $|f(x) - f(y)| \leq t_f(\delta)$ and $|h(x) - h(y)| \leq t_h(\delta)$.

The next coupling lemma shows that the correlation functions q_λ^g and c_λ^g are uniformly approximated by the first order approximations $c_\lambda^{g(x)}$ and $q_\lambda^{g(x)}$, respectively. This lemma, whose proof is deferred to the end of the section, depends heavily on the stabilization of I^g.

Lemma 4.2 *Let $\phi \in \mathcal{F}$. There exists a function $\delta : \mathbb{R}^+ \to \mathbb{R}^+$, increasing to ∞, such that as $\lambda \to \infty$, $\delta(\lambda)/\lambda \to 0$ and*

(i) For all $x \in A$ distant at least $(\delta/\lambda)^{1/d}$ from ∂A, $|q_\lambda^g(x) - \tilde{q}_\lambda^g(x)| \to 0$,

(ii) For all $x, y \in A$ distant at least $(\delta/\lambda)^{1/d}$ from ∂A,

$$\delta(\lambda)|c_\lambda^g(x,y) - \tilde{c}_\lambda^g(x,y)| \to 0,$$

(iii) $\delta(\lambda)t_f((\delta(\lambda)/\lambda)^{1/d}) \to 0$ and $\delta(\lambda)t_h((\delta(\lambda)/\lambda)^{1/d}) \to 0$.

The following lemma establishes the variance limit (4.3).

Lemma 4.3 *Let $\phi \in \mathcal{F}$ and $h \in C(A)$. Then*

$$\lim_{\lambda\to\infty} \frac{\mathrm{Var}[\langle h, \mu_{\lambda,\phi}^g \rangle]}{\lambda} = \int_A h^2(x) V_\phi\left(\frac{g(x)}{f(x)}\right) f(x)dx.$$

Proof. Using the definition of q_λ^g and c_λ^g, we rewrite the right hand side of (4.7) as

$$\int_A h^2(x) q_\lambda^g(x) f(x)dx + \lambda \int\int_{A\times A} h(y)h(x)c_\lambda^g(x,y)f(x)f(y)dxdy \quad =$$

$$\int_A h(x)f(x)\left[h(x)q_\lambda^g(x) + \lambda\int_A h(y)c_\lambda^g(x,y)f(y)dy\right]dx.$$

We want to compare the bracketed expression in the last integral, namely

$$h(x)q_\lambda^g(x) + \lambda\int_A h(y)c_\lambda^g(x,y)f(y)dy \tag{4.12}$$

with its corresponding uniform version

$$h(x)\tilde{q}_\lambda^g(x) + \lambda\int_{\mathbb{R}^d} h(x)\tilde{c}_\lambda^g(x,y)f(x)dy. \tag{4.13}$$

Let $\delta := \delta(\lambda)$ be as in Lemma 4.2. Since A is compact and convex, it satisfies the following boundary regularity condition: $\lim_{\lambda\to\infty} \frac{|\partial_r(\lambda^{1/d}A)|}{\lambda} = 0$ for all $r > 0$, where $\partial_r(\lambda^{1/d}A)$ denotes the r-neighborhood of the boundary of $\lambda^{1/d}A$. Therefore, convexity and compactness of A, together with $\delta/\lambda \to 0$, imply that the set of $x \in A$ distant at most $2(\delta/\lambda)^{1/d}$ from ∂A has measure zero as $\lambda \to \infty$. Thus it will be enough to show for all $x \in A$ distant at least $2(\delta/\lambda)^{1/d}$ from ∂A, that the difference of the expressions (4.12) and (4.13) converges to zero as $\lambda \to \infty$.

Lemma 4.2(i) implies for all x distant at least $(\delta/\lambda)^{1/d}$ from ∂A that the difference of the first terms in (4.12) and (4.13) goes to zero as $\lambda \to \infty$. For all $x \in A$, the difference of the integrals in (4.12) and (4.13) equals:

$$\lambda \int_{\mathbb{R}^d} [c^g_\lambda(x,y)h(y)f(y) - \tilde{c}^g_\lambda(x,y)h(x)f(x)]dy. \tag{4.14}$$

To evaluate (4.14), we integrate separately over $B_{(\delta/\lambda)^{1/d}}(x)$ and $\mathbb{R}^d \setminus B_{(\delta/\lambda)^{1/d}}(x)$. We split the integral over $B_{(\delta/\lambda)^{1/d}}(x)$ as

$$\begin{aligned} \lambda \int_{B_{(\delta/\lambda)^{1/d}}(x)} &(c^g_\lambda(x,y) - \tilde{c}^g_\lambda(x,y))h(y)f(y)dy \,+ \\ \lambda \int_{B_{(\delta/\lambda)^{1/d}}(x)} &\tilde{c}^g_\lambda(x,y)(h(y) - h(x))f(y)dy \,+ \\ \lambda \int_{B_{(\delta/\lambda)^{1/d}}(x)} &\tilde{c}^g_\lambda(x,y)h(x)(f(y) - f(x))dy. \end{aligned} \tag{4.15}$$

The first integral is bounded by the product of λ, the volume of $B_{(\delta/\lambda)^{1/d}}(x)$, and the maximum over y of $|(c^g_\lambda(x,y) - \tilde{c}^g_\lambda(x,y))h(y)f(y)|$. However, since y is distant at least $(\delta/\lambda)^{1/d}$ from ∂A, the product goes to zero by Lemma 4.2(ii). The second and third integrals also tend to zero as $\lambda \to \infty$ by the bound (4.9) and Lemma 4.2(iii).

Since f and h are bounded, the integral in (4.14) over $\mathbb{R}^d \setminus B_{(\delta/\lambda)^{1/d}}(x)$ is bounded by

$$K \int_{\mathbb{R}^d \setminus B_{(\delta/\lambda)^{1/d}}(x)} [c^g_\lambda(x,y) + \tilde{c}^g_\lambda(x,y)] d(\lambda^{1/d}y) \le K \int_{\mathbb{R}^d \setminus B_{\delta^{1/d}}(x)} \exp(-C|z-x|)dz$$

by Lemma 4.1. This last integral is bounded by $K\omega_d \int_{\delta^{1/d}}^{\infty} \exp(-Kt)t^{d-1}dt$, which tends to zero as $\lambda \to \infty$ since $\delta \to \infty$. We conclude that (4.14) converges to 0 uniformly in $x \in A$. Hence,

$$\lambda^{-1}\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi}\rangle] - \int_A f(x)h(x) \left[h(x)\tilde{q}^g_\lambda(x) + \lambda \int_{\mathbb{R}^d} h(x)\tilde{c}^g_\lambda(x, x+y)f(x)dy \right] dx \tag{4.16}$$

converges to 0 as $\lambda \to \infty$.

Using $\Phi^\alpha(x, \mathcal{P}_\beta) \overset{\mathcal{D}}{=} \Phi^{\frac{\alpha}{\beta}}(x, \mathcal{P}) \overset{\mathcal{D}}{=} \Phi^{\frac{\alpha}{\beta}}(\mathbf{0}, \mathcal{P})$ for all $\alpha, \beta \in \mathbb{R}^+$ and $x \in \mathbb{R}^d$, we get

$$\tilde{q}^g_\lambda(x) = \mathbb{E}\,(\Phi^{g(x)}(x, \mathcal{P}_{f(x)}))^2 = \mathbb{E}\,(\Phi^{\frac{g(x)}{f(x)}}(\mathbf{0}, \mathcal{P}))^2.$$

Also, by definition of $\tilde{c}^g_\lambda$ we have

$$\lambda \int_{\mathbb{R}^d} \tilde{c}^g_\lambda(x, x+y) f(x) dy = \int_{\mathbb{R}^d} \tilde{c}^g(x, x+y) f(x) dy$$

$$= \int_{\mathbb{R}^d} \left[\mathbb{E}\, \Phi^{g(x)}(\mathbf{0}, \mathcal{P}_{f(x)} \cup y) \Phi^{g(x)}(y, \mathcal{P}_{f(x)} \cup \mathbf{0}) - \mathbb{E}\, \Phi^{g(x)}(\mathbf{0}, \mathcal{P}_{f(x)}) \mathbb{E}\, \Phi^{g(x)}(y, \mathcal{P}_{f(x)}) \right] f(x) dy. \tag{4.17}$$

Since $\Phi^\alpha(\mathbf{0}, \mathcal{P}_\beta \cup y) \overset{\mathcal{D}}{=} \Phi^{\frac{\alpha}{\beta}}(\mathbf{0}, \mathcal{P} \cup \beta^{-1/d} y)$ holds for all $\alpha, \beta > 0$ and all $y \in \mathbb{R}^d$, (4.17) equals

$$= \int_{\mathbb{R}^d} \left[\mathbb{E}\, \Phi^{\frac{g(x)}{f(x)}}(\mathbf{0}, \mathcal{P} \cup z) \Phi^{\frac{g(x)}{f(x)}}(z, \mathcal{P} \cup \mathbf{0}) - \mathbb{E}\, \Phi^{\frac{g(x)}{f(x)}}(\mathbf{0}, \mathcal{P}) \mathbb{E}\, \Phi^{\frac{g(x)}{f(x)}}(z, \mathcal{P}) \right] dz.$$

Thus, combining (4.16)- (4.17) yields

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi} \rangle]}{\lambda} = \int_A h^2(x) \mathbb{E}\, (\Phi^{\frac{g(x)}{f(x)}}(\mathbf{0}, \mathcal{P}))^2 f(x) dx +$$

$$+ \int_A \int_{\mathbb{R}^d} h^2(x) \left[\mathbb{E}\, \Phi^{\frac{g(x)}{f(x)}}(\mathbf{0}, \mathcal{P} \cup z) \Phi^{\frac{g(x)}{f(x)}}(z, \mathcal{P} \cup \mathbf{0}) - \mathbb{E}\, \Phi^{\frac{g(x)}{f(x)}}(\mathbf{0}, \mathcal{P}) \mathbb{E}\, \Phi^{\frac{g(x)}{f(x)}}(z, \mathcal{P}) \right] dz f(x) dx.$$

Finally, since $\Phi^\beta(\mathbf{0}, \mathcal{P}) := \phi(\beta |C(\mathbf{0}, \mathcal{P})|)$ and using the definition of V_ϕ (recall (2.5)) we obtain

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[\langle h, \mu^g_{\lambda,\phi} \rangle]}{\lambda} = \int_A h^2(x) V_\phi \left(\frac{g(x)}{f(x)} \right) f(x) dx$$

as desired. Finally, taking $A = [0,1]^d$, $h \equiv f \equiv 1$, and $g \equiv \beta$, it follows that

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[N^\beta_{\lambda,\phi}(\mathcal{P}_\lambda \cap [0,1]^d)]}{\lambda} = \int_{[0,1]^d} V_\phi(\beta) dx = V_\phi(\beta),$$

furnishing the limit (2.4). □

Now we return to the

Proof of Lemma 4.2. To prove implication (i) we need to show for all $x \in A$ distant at least $(\delta/\lambda)^{1/d}$ from ∂A, that

$$\left| \mathbb{E}\, \phi^2 \left(\lambda \int_{C(x, \mathcal{P}_{\lambda f})} g(u) du \right) - \mathbb{E}\, \phi^2 \left(\lambda \int_{C(x, \mathcal{P}_{\lambda f(x)})} g(x) du \right) \right| \to 0 \tag{4.18}$$

as $\lambda \to \infty$. Let $x' := x'_\lambda \in C(x, \mathcal{P}_{\lambda f})$ satisfy

$$\int_{C(x, \mathcal{P}_{\lambda f})} g(u) du = g(x') \int_{C(x, \mathcal{P}_{\lambda f})} du. \tag{4.19}$$

Such a point exists because of the continuity of g and the mean value theorem for integrals.

Also, $g(x'_\lambda) - g(x) \xrightarrow{P} 0$ as $\lambda \to \infty$ and by the boundedness of g we have $\mathbb{E}\,|g(x'_\lambda) - g(x)|^2 \to 0$. Chebyshev's inequality and Cauchy-Schwarz imply that as $\lambda \to \infty$,

$$\lambda g(x'_\lambda) \int_{C(x,\mathcal{P}_{\lambda f})} du - \lambda g(x) \int_{C(x,\mathcal{P}_{\lambda f})} du \xrightarrow{P} 0, \tag{4.20}$$

since $\lambda \int_{C(x,\mathcal{P}_{\lambda f})} du$ has a finite second moment. Moreover (see e.g. Lemma 5.1 of [25]) for all $x \in A$ distant at least $(\delta/\lambda)^{1/d}$ from ∂A

$$\lambda g(x) \int_{C(x,\mathcal{P}_{\lambda f})} du \xrightarrow{P} \frac{g(x)}{f(x)} \Gamma_1. \tag{4.21}$$

Combining (4.19- 4.21) with the continuity of ϕ gives as $\lambda \to \infty$,

$$\phi^2 \left(\lambda \int_{C(x,\mathcal{P}_{\lambda f})} g(u) du \right) \xrightarrow{P} \phi^2 \left(\frac{g(x)}{f(x)} \Gamma_1 \right).$$

Now for all $x \in A$ and $\phi \in \mathcal{F}$ we have

$$\sup_{\lambda \geq 1} \mathbb{E} \left[\phi^2 \left(\lambda \int_{C(x,\mathcal{P}_{\lambda f})} g(u) du \right) \right] < \infty.$$

It follows that for each x, $\phi^2 \left(\lambda \int_{C(x,\mathcal{P}_{\lambda f})} g(u) du \right)$, $\lambda > 0$, are uniformly integrable, showing that (4.18) holds and so Lemma 4.2(i) holds.

Now we prove implications (ii) and (iii). Clearly, there is a function $\delta := \delta(\lambda) \to \infty$ as $\lambda \to \infty$ such that implication (iii) is true and, by exponential decay of c^g_λ, that

$$\delta(\lambda) \sup_{|x-y| \geq (\delta/\lambda)^{1/d}} c^g_\lambda(x,y) \to 0 \quad \text{and} \quad \delta(\lambda) \sup_{|x-y| \geq (\delta/\lambda)^{1/d}} \tilde{c}^g_\lambda(x,y) \to 0.$$

From the definition of c^g_λ and $\tilde{c}^g_\lambda$, it is enough to show that

$$\sup_{|x-y| < (\delta/\lambda)^{1/d}} \delta(\lambda) |\mathbb{E}\,[\Phi^g_\lambda(x, \mathcal{P}_{\lambda f} \cup y) \Phi^g_\lambda(y, \mathcal{P}_{\lambda f} \cup x) - \Phi^{g(x)}_\lambda(x, \mathcal{P}_{\lambda f(x)} \cup y) \Phi^{g(x)}_\lambda(y, \mathcal{P}_{\lambda f(x)} \cup x)]| \to 0 \tag{4.22}$$

and

$$\sup_{|x-y| < (\delta/\lambda)^{1/d}} \delta(\lambda) |\mathbb{E}\,[\Phi^g_\lambda(x, \mathcal{P}_{\lambda f}) \Phi^g_\lambda(y, \mathcal{P}'_{\lambda f}) - \Phi^{g(x)}_\lambda(x, \mathcal{P}_{\lambda f(x)}) \Phi^{g(x)}_\lambda(y, \mathcal{P}'_{\lambda f(x)})]| \to 0 \tag{4.23}$$

as $\lambda \to \infty$. We will show (4.22); (4.23) has a similar proof.

Now bound the absolute value in (4.22) as

$$|\mathbb{E}\,[\Phi^g_\lambda(x, \mathcal{P}_{\lambda f} \cup y) \Phi^g_\lambda(y, \mathcal{P}_{\lambda f} \cup x) - \Phi^{g(x)}_\lambda(x, \mathcal{P}_{\lambda f(x)} \cup y) \Phi^{g(x)}_\lambda(y, \mathcal{P}_{\lambda f(x)} \cup x)]|$$

$$\leq |\mathbb{E}\,[(\Phi^g_\lambda(x, \mathcal{P}_{\lambda f} \cup y) - \Phi^{g(x)}_\lambda(x, \mathcal{P}_{\lambda f(x)} \cup y)) \Phi^g_\lambda(y, \mathcal{P}_{\lambda f} \cup x)]|$$

$$+ \, |\mathbb{E}\,[(\Phi_\lambda^g(y, \mathcal{P}_{\lambda f} \cup x) - \Phi_\lambda^{g(x)}(y, \mathcal{P}_{\lambda f(x)} \cup x))\Phi_\lambda^{g(x)}(x, \mathcal{P}_{\lambda f(x)} \cup y)]|.$$

Since $\mathbb{E}\,[(\Phi_\lambda^g(y, \mathcal{P}_{\lambda f} \cup x))^2] < \infty$ and $\mathbb{E}\,[(\Phi_\lambda^{g(x)}(x, \mathcal{P}_{\lambda f(x)} \cup y))^2] < \infty$ for all x, y, it suffices by Cauchy-Schwarz to show for all $x \in A$ that as $\lambda \to \infty$

$$\delta(\lambda)\mathbb{E}\left[\left|\Phi_\lambda^g(x, \mathcal{P}_{\lambda f} \cup y) - \Phi_\lambda^{g(x)}(x, \mathcal{P}_{\lambda f(x)} \cup y)\right|^2\right] \to 0 \tag{4.24}$$

and that

$$\sup_{|x-y|<(\delta/\lambda)^{1/d}} \delta(\lambda)\mathbb{E}\left[\left|\Phi_\lambda^g(y, \mathcal{P}_{\lambda f} \cup x) - \Phi_\lambda^{g(x)}(y, \mathcal{P}_{\lambda f(x)} \cup x)\right|^2\right] \to 0. \tag{4.25}$$

Let $x \in A$ be distant at least $(\delta/\lambda)^{1/d}$ from ∂A. Let $A(x, \lambda, \delta)$ be the event that $\mathcal{P}_{\lambda f} = \mathcal{P}_{\lambda f(x)}$ on the ball $B_x((\delta/\lambda)^{1/d})$ *and* that the radius of stabilization $R_{x,\lambda}^{f,g}$ for Φ^g with respect to $\mathcal{P}_{\lambda f}$ at x is less than δ. The coupling estimate (4.7) and the exponential decay of $R_{x,\lambda}^{f,g}$ imply that

$$P[A(x, \lambda, \delta)^c] \le K[\omega_d \delta t_f((\delta/\lambda)^{1/d}) + \exp(-K\delta)]. \tag{4.26}$$

Now we show (4.24). Now the left-hand side of (4.24) equals

$$\delta(\lambda)\mathbb{E}\left[\left|\phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f} \cup y)} g(u)du\right) - \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f(x)} \cup y)} g(x)du\right)\right|^2\right]$$

which is bounded by the sum of

$$\delta(\lambda)\mathbb{E}\left[\left|\phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f} \cup y)} g(u)du\right) - \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f(x)} \cup y)} g(x)du\right)\right|^2 1_{A(x,\lambda,\delta)}\right]$$

and

$$\delta(\lambda)\mathbb{E}\left[\left|\phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f} \cup y)} g(u)du\right) - \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f(x)} \cup y)} g(x)du\right)\right|^2 1_{A^c(x,\lambda,\delta)}\right].$$

The second expectation goes to zero by Cauchy-Schwarz, $\phi \in \mathcal{F}$, and the estimate (4.26).

Since $C(x, \mathcal{P}_{\lambda f}) \subset B_x((\delta/\lambda)^{1/d})$ on the set $A(x, \lambda, \delta)$, the definition of $A(x, \lambda, \delta)$ shows that the first term is bounded by

$$\delta(\lambda)\mathbb{E}\left[\left|\phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f(x)} \cup y)} g(u)du\right) \; - \; \phi\left(\lambda \int_{C(x,\mathcal{P}_{\lambda f(x)} \cup y)} g(x)du\right)\right|^2\right].$$

The mean value theorem for integrals implies the existence of some $x'_\lambda \in C(x, \mathcal{P}_{\lambda f(x) \cup y})$ such that the above equals

$$\delta(\lambda)\mathbb{E}\left[\left|\phi\left(g(x'_\lambda)|C(x, \mathcal{P}_{f(x)} \cup y)|\right) - \phi\left(g(x)|C(x, \mathcal{P}_{f(x)} \cup y)|\right)\right|^2\right].$$

Since $x'_\lambda \xrightarrow{P} x$, the continuity of ϕ and g imply that

$$\left|\phi\left(g(x'_\lambda)|C(x,\mathcal{P}_{f(x)} \cup y)|\right) - \phi(g(x)|C(x,\mathcal{P}_{f(x)} \cup y)|)\right|^2 \xrightarrow{P} 0. \tag{4.27}$$

Since the left hand side of (4.27) is uniformly integrable, it follows that

$$\lim_{\lambda\to 0} \mathbb{E}\left[\left|\phi\left(g(x'_\lambda)|C(x,\mathcal{P}_{f(x)} \cup y)|\right) - \phi(g(x)\cdot|C(x,\mathcal{P}_{f(x)} \cup y)|)\right|^2\right] = 0 \tag{4.28}$$

and therefore there is some $\delta_o(\lambda) \to \infty$ such that

$$\delta_o(\lambda)\mathbb{E}\left[\left|\phi\left(g(x'_\lambda)\cdot|C(x,\mathcal{P}_{f(x)} \cup y)|\right) - \phi(g(x)|C(x,\mathcal{P}_{f(x)} \cup y)|)\right|^2\right] \to 0.$$

This shows (4.24), changing our choice of $\delta(\lambda)$, if necessary.

Now we prove (4.25), i.e. we show

$$\lim_{\lambda\to\infty} \sup_{|x-y|<(\delta/\lambda)^{1/d}} \delta(\lambda)\mathbb{E}\left[\left|\phi\left(\lambda\int_{C(y,\mathcal{P}_{\lambda f}\cup x)} g(u)du\right) - \phi\left(\lambda\int_{C(y,\mathcal{P}_{\lambda f(x)}\cup x)} g(x)du\right)\right|^2\right] = 0.$$

Let $y \in B_x((\frac{\delta}{\lambda})^{1/d})$. Let $A(x,\lambda,\delta,y)$ be the event that $\mathcal{P}_{\lambda f} = \mathcal{P}_{\lambda f(x)}$ on the ball $B_x(2(\delta/\lambda)^{1/d})$ *and* that the radius of stabilization $R^{f,g}_{y,\lambda}$ for Φ^g with respect to $\mathcal{P}_{\lambda f}$ at y is less than δ. Thus $C(y,\mathcal{P}_{\lambda f} \cup x)) = C(y,\mathcal{P}_{\lambda f(x)} \cup x))$ on the set $A(x,\lambda,\delta,y)$. Following the proof of (4.24) it is therefore enough to show

$$\lim_{\lambda\to\infty} \sup_{|x-y|<(\delta/\lambda)^{1/d}} \delta(\lambda)\mathbb{E}\left[\left|\phi\left(\lambda\int_{C(y,\mathcal{P}_{\lambda f(x)}\cup x)} g(y)du\right) - \phi\left(\lambda\int_{C(y,\mathcal{P}_{\lambda f(x)}\cup x)} g(x)du\right)\right|^2\right] = 0.$$

The continuity of g and ϕ ensure that

$$\lim_{\lambda\to\infty} \sup_{|x-y|<(\delta/\lambda)^{1/d}} \mathbb{E}\left[\left|\phi\left(\lambda\int_{C(y,\mathcal{P}_{\lambda f(x)}\cup x)} g(y)du\right) - \phi\left(\lambda\int_{C(y,\mathcal{P}_{\lambda f(x)}\cup x)} g(x)du\right)\right|^2\right] = 0.$$

Decreasing our choice of $\delta(\lambda)$ if necessary, (4.25) follows. This completes the proof of Lemma 4.2.
□

4.3 Step 2: de-Poissonization: variance convergence over samples of fixed size

We show here how to pass from (4.3) to the de-Poissonized limit (4.4). Let $\mathcal{X}_m$ be the point process consisting of m i.i.d. random variables $X_1, ..., X_m$ with density f on A. De-Poissonization involves coupling $\mathcal{X}_m$, m large, with a Poisson point process [31, 30, 5]. It will suffice to establish the following two coupling lemmas.

We first fix the terminology. For all $h \in C(A)$, and $m, n \in \mathbb{N}^+$, let

$$H_n^h(\mathcal{X}_m) := \sum_{i=1}^{m} \phi(nD_{i,n}^g)h(X_i)$$

and put $R_{m,n}^h := H_n^h(\mathcal{X}_{m+1}) - H_n^h(\mathcal{X}_m)$. Define for all $\beta > 0$

$$\Delta'_\phi(\beta) := \lim_{r\to\infty}[N_{1,\phi}^{\beta}(\mathcal{P} \cap B_r(\mathbf{0}) \cup \mathbf{0}) - N_{1,\phi}^{\beta}(\mathcal{P} \cap B_r(\mathbf{0}))].$$

The existence of $\Delta'_\phi(\beta)$ is guaranteed by Lemma 6.1 and Definition 2.1 of [31]. Note that $\mathbb{E}\left[\Delta'_\phi(\beta)\right] = \Delta_\phi(\beta)$ where Δ_ϕ is given by Definition 2.2.

Lemma 4.4 *Fix $\phi \in \mathcal{F}$ and let $\varepsilon > 0$. Let h, g and f be fixed. There exists $\delta > 0$ and $n_0 \geq 1$ such that for all $n \geq n_0$ and all $m, m' \in [(1-\delta)n, (1+\delta)n]$ with $m < m'$, there exists random variables X, X' with density f and a coupled family of variables $D := D(X)$, $D' := D(X')$, $R := R(X, X')$, $R' := R'(X, X')$ with the following properties:*

(i) D and D' each have the same distribution as $h(X)\Delta'_\phi(\frac{g(X)}{f(X)})$,

(ii) D and D' are independent,

(iii) (R, R') have the same joint distribution as $(R_{m,n}^h, R_{m',n}^h)$,

(iv) $P[\{|D - R| > \varepsilon\} \cup \{|D' - R'| > \varepsilon\}] < \varepsilon$.

Proof. The existence of D, D', R, and R' as well as properties (i)-(iii) follow from modifications of the proof of Lemma 6.1 of [5] (alternatively, see Theorem 2.16 of [30]). Suppose we are given n. Let $X, X', Y_1, Y_2, ...$ be i.i.d. random variables with density f on A. On a suitable probability space, let $\mathcal{P} := \mathcal{P}_{nf}$ and $\mathcal{P}' := \mathcal{P}'_{nf}$ be independent Poisson processes on A with intensity $nf : A \to \mathbb{R}^+$.

Let $\mathcal{P}''$ be the point process consisting of those points of $\mathcal{P}$ which lie closer to X than to X' (in the Euclidean norm), together with those points of $\mathcal{P}'$ which lie closer to X' than to X. Clearly $\mathcal{P}''$ is a Poisson process also having intensity measure nf on A, and moreover it is independent of X and of X'.

Let N denote the number of points of $\mathcal{P}''$ (a Poisson variable with mean $n|A|$). Choose an ordering on the points of $\mathcal{P}''$, uniformly at random from all $N!$ possible such orderings. Use this ordering to list the points of $\mathcal{P}''$ as $W_1, W_2, \ldots, W_N$. Also, set $W_{N+1} = Y_1, W_{N+2} = Y_2, W_{N+3} = Y_3$ and so on.

Let

$$R := R(X, X') := H_n^h(\{W_1, \ldots, W_m, X\}) - H_n^h(\{W_1, \ldots, W_m\})$$

and

$$R' := R(X, X') := H_n^h(\{W_1, \dots, W_{m'-1}, X, X'\}) - H_n^h(\{W_1, \dots, W_{m'-1}, X\}).$$

$X, X', W_1, W_2, W_3, \dots,$ are i.i.d. variables on A with density f, and therefore the pairs (R, R') and $(R_{m,n}^h, R_{m',n}^h)$ have the same joint distribution as claimed.

To prove property (iv), we need to first fix the notation. Fix $h \in C(A)$. We will show $P[|D - R| > \varepsilon] < \varepsilon$; the proof of $P[|D' - R'| > \varepsilon] < \varepsilon$ is identical.

For all $\tau > 0$, recall that $\mathcal{P}_\tau$ is a homogeneous Poisson point process with intensity τ. Given $\mathcal{P}_\tau$, for all $x \in \mathbb{R}^d$ and $\beta > 0$, let $B(x, \tau)$ denote a ball (depending on $\mathcal{P}_\tau$) such that

$$N_{1,\phi}^{\beta}(\mathcal{P}_\tau \cap B(x,\tau) \cup x) - N_{1,\phi}^{\beta}(\mathcal{P}_\tau \cap B(x,\tau)) = \Delta'_\phi(\frac{\beta}{\tau}).$$

For all $x \in A$ define

$$D^g(x) := [N_{1,\phi}^{g}(\mathcal{P}_{f(x)} \cap B(x, f(x)) \cup x) - N_{1,\phi}^{g}(\mathcal{P}_{f(x)} \cap B(x, f(x)))]h(x)$$

and

$$D(x) := D^{g(x)}(x) := [N_{1,\phi}^{g(x)}(\mathcal{P}_{f(x)} \cap B(x, f(x)) \cup x) - N_{1,\phi}^{g(x)}(\mathcal{P}_{f(x)} \cap B(x, f(x)))]h(x) = h(x)\Delta'_\phi(\frac{g(x)}{f(x)}).$$

By following Lemma 6.1 of [5] we can show the following approximations for n large:

(i) the set of points of $W_1, W_2, ..., W_m$ on $n^{-1/d}B(x, f(x))$ differs little from the set of points $\mathcal{P}_{nf}$ on $n^{-1/d}B(x, f(x))$, and

(ii) the set of points $\mathcal{P}_{nf}$ on $n^{-1/d}B(x, f(x))$ differs little from the set of points of $\mathcal{P}_{nf(x)}$ on $n^{-1/d}B(x, f(x))$.

For all $n = 1, 2, ...$ we have

$$D^g(x) \overset{\mathcal{D}}{=} D_n^g(x) := [N_{n,\phi}^{g}(\mathcal{P}_{nf(x)} \cap n^{-1/d}B(x, f(x)) \cup x) - N_{n,\phi}^{g}(\mathcal{P}_{nf(x)} \cap n^{-1/d}B(x, f(x)))]h(x). \tag{4.29}$$

Combining (i) and (ii) one might expect that $|R(x, x') - D_n^g(x)|$ is small in probability for n large. Indeed, by following Lemma 6.1 of [5] we can show for all $\varepsilon > 0$ that there exists $\delta > 0$, $n_0 \geq 1$ such that for all $n \geq n_0$ and all $m, m' \in [(1-\delta)n, (1+\delta)n]$ with $m < m'$ and all pairs $(x, x') \in A \times A$ that

$$P[|R(x, x') - D_n^g(x)| > \varepsilon/2] < \varepsilon/2.$$

To complete the proof that $P[|D - R| > \varepsilon] < \varepsilon$, it remains to show for all $x \in A$ and all $\varepsilon > 0$ that for n large

$$P[|D_n^g(x) - D^{g(x)}(x)| > \varepsilon/2] < \varepsilon/2. \tag{4.30}$$

By (4.29), it follows that the difference $D_n^g(x) - D^{g(x)}(x)$ can be estimated by considering differences of the type

$$\sum_{X_i \in B(x,f(x))\cap \mathcal{P}_{f(x)}} \phi\left(\int_{C(X_i,B(x,f(x))\cap\mathcal{P}_{f(x)})} g(u)du\right) -$$

$$- \sum_{X_i\in n^{-1/d}B(x,f(x))\cap\mathcal{P}_{nf(x)}} \phi\left(n\int_{C(X_i,\ n^{-1/d}B(x,f(x))\cap\mathcal{P}_{nf(x)})} g(x)du\right)$$

$$\overset{\mathcal{D}}{=} \sum_{X_i\in n^{-1/d}B(x,f(x))\cap\mathcal{P}_{nf(x)}} \phi\left(n\int_{C(X_i,\mathcal{P}_{nf(x)})} g(u)du\right) - \phi\left(n\int_{C(X_i,\mathcal{P}_{nf(x)})} g(x)du\right).$$

Using the continuity of ϕ and g as well as the boundedness of f and g, standard arguments (see (4.27) - (4.28)) imply that the summands are uniformly small in probability for large n. Since the number of summands is a.s. finite, the above difference is small in probability, showing (4.30). Thus condition (iv) holds and Lemma 4.4 is proved. □

The proof of the next lemma follows that of Lemma 4.3 in [31]. Recall that the random variable X has a density f.

Lemma 4.5 *Let $(k(n))_{n\geq 1}$ be a sequence with $k(n)/n \to 0$ as $n\to\infty$. Then for all $h\in C(A)$*

$$\lim_{n\to\infty} \sup_{n-k(n)\leq m\leq n+k(n)} \left| \mathbb{E}\, R_{m,n}^h - \mathbb{E}\, h(X)\Delta_\phi'\left(\frac{g(X)}{f(X)}\right)\right| = 0.$$

Also

$$\lim_{n\to\infty} \sup_{n-k(n)\leq m< m'\leq n+k(n)} \left| \mathbb{E}\, R_{m,n}^h R_{m',n}^h - \left(\mathbb{E}\, h(X)\Delta_\phi'\left(\frac{g(X)}{f(X)}\right)\right)^2\right| = 0,$$

and

$$\lim_{n\to\infty} \sup_{n-k(n)\leq m\leq n+k(n)} |\mathbb{E}\,(R_{m,n}^h)^2| < \infty.$$

We now show the limit (4.4). Given $h\in C(A)$, let $H_n^h := \langle h, \mu_{n,\phi}^g\rangle$ and $H_\lambda^{'h} := \langle h, \mu_{\lambda,\phi}^g\rangle$. We assume that $\mathcal{P}_{\lambda f}$ is coupled to $\{X_1,...,X_n\}$ by setting $\mathcal{P}_{\lambda f} := \{X_1,...,X_{N_n}\}$ with N_n an independent Poisson random variable with mean n.

We show for all $h\in C(A)$ that

$$\lim_{n\to\infty} \frac{\mathrm{Var}[H_n^h]}{n} \to \tau_h^2$$

where

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

Letting $\alpha := \mathbb{E}\,[h(X)\Delta'_\phi(\frac{g(X)}{f(X)})]$ first note that as $n \to \infty$

$$\mathbb{E}\left[(n^{-1/2}({H'}^h_n - H^h_n - (N_n - n)\alpha))^2\right] \to 0. \tag{4.31}$$

To show (4.31) we employ the coupling Lemmas 4.4 and 4.5 and follow pp. 1019-1020 of [31] verbatim. Now consider the identity

$$n^{-1/2}{H'}^h_n = n^{-1/2}H^h_n + n^{-1/2}(N_n - n)\alpha + n^{-1/2}({H'}^h_n - H^h_n - (N_n - n)\alpha).$$

The last summand in the above has variance tending to zero by (4.31) and the second summand has variance α^2 and is independent of the first term. Letting $\sigma_h^2 := \int_A h^2(x)V_\phi\left(\frac{g(x)}{f(x)}\right)f(x)dx$, it follows that

$$\sigma_h^2 = \lim_{n\to\infty} n^{-1}\mathrm{Var}[{H'}^h_n] = \lim_{n\to\infty} n^{-1}\mathrm{Var}[H^h_n] + \alpha^2,$$

i.e., (4.4) holds as desired.

4.4 Step 3: Completion of the proof of Theorem 2.1

We consider

$$\mu^g_{\lambda,\phi} := \sum_{x\in\mathcal{P}_{\lambda f}} \Phi^g_\lambda(x, \mathcal{P}_{\lambda f})\delta_x,$$

where we recall from (4.1) that Φ^g is a stabilizing functional with radius of stabilization $R^{f,g}_{x,\lambda}$ with respect to $\mathcal{P}_{\lambda f}$ at x. $R^{f,g}_{x,\lambda}$ has exponentially decaying tails uniformly in $x \in A$ and $\lambda > 0$, i.e., there exists $K > 0$ such that for all $t > 0$

$$\sup_{\lambda>0,\ x\in A} P[R^{f,g}_{x,\lambda} > t] \le K\exp(-t/K).$$

By hypothesis, $\sup_{\lambda>0,x\in A}\mathbb{E}\,[|\Phi^g_\lambda(x, \mathcal{P}_{\lambda f}\cup x)|^p] < \infty$ for some $p > 3$. By Corollary 2.4 of [33], for all $h \in C(A)$ there is a constant $K > 0$ such that for all $\lambda \ge 3$

$$\sup_{t\in\mathbb{R}}\left|P\left[\frac{\langle h, \overline{\mu}^g_{\lambda,\phi}\rangle}{(\mathrm{Var}[\langle h, \overline{\mu}^g_{\lambda,\phi}\rangle])^{1/2}} \le t\right] - P[N(0,1) \le t]\right| \le K(\log\lambda)^{3d}\lambda^{-1/2}.$$

Thus, combining with Step 1, we obtain that $\lambda^{-1/2}\langle h, \overline{\mu}^g_{\lambda,\phi}\rangle$ converges as $\lambda \to \infty$ to a normal random variable with mean zero and variance $\int_A h^2(x)V_\phi\left(\frac{g(x)}{f(x)}\right)f(x)dx$, completing the proof of (4.5).

The limit (4.5) shows that $n^{-1/2}({H'}^h_n - \mathbb{E}\,{H'}^h_n) \xrightarrow{\mathcal{D}} N(0, \sigma_h^2)$. Combined with (4.31) this gives

$$n^{-1/2}(H^h_n - \mathbb{E}\,{H'}^h_n + (N_n - n)\alpha) \xrightarrow{\mathcal{D}} N(0, \sigma_h^2).$$

Since $n^{-1/2}(N_n - n)\alpha$ is independent of H_n^h and is asymptotically normal with mean zero and variance α^2, it follows by considering characteristic functions that

$$n^{-1/2}(H_n^h - \mathbb{E}\, {H'}_n^h) \xrightarrow{\mathcal{D}} N(0, \sigma_h^2 - \alpha^2). \tag{4.32}$$

By (4.31), the expectation of $n^{-1/2}({H'}_n^h - H_n^h - (N_n - n)\alpha)$ tends to zero, so in (4.32) we can replace $\mathbb{E}\, {H'}_n^h$ by $\mathbb{E}\, H_n$, which gives

$$n^{-1/2}(H_n^h - \mathbb{E}\, H_n^h) \xrightarrow{\mathcal{D}} N(0, \tau_h^2).$$

Taking $h = h_1 + h_2$ and using simple algebra shows that the limiting Gaussian field has the desired covariance kernel (2.8), completing the proof of Theorem 2.1. □

5 Proof of Theorem 2.2

The proof of Theorem 2.2 follows exactly along the lines of the proof of Theorem 2.1. We only need to show (2.14), i.e. show for all $\beta > 0$ that

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[S^{\beta}_{\lambda,\phi,k}(\mathcal{P}_\lambda \cap [0,1])]}{\lambda} = V^S_{\phi,k}(\beta) \tag{5.1}$$

where $V^S_{\phi,k}$ is given by (2.13). Throughout $\mathcal{P}$ denotes a rate one Poisson point process on $\mathbb{R}$ and $0 < S_1 < S_2 < ...$ denote the order statistics for the restriction of $\mathcal{P}$ to $\mathbb{R}^+$. Let $\overline{S}_k := \{S_1, ..., S_k\}$, and put $S_0 := 0$.

Simple modifications of the proof of Step 1 of Theorem 2.1 show that for all $\beta > 0$,

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[S^{\beta}_{\lambda,\phi,k}(\mathcal{P}_\lambda \cap [0,1])]}{\lambda} = \mathbb{E}\,[\phi^2(\beta\Gamma_k)] + \int_{-\infty}^{\infty} c^{\beta}(\mathbf{0}, y)dy$$

where

$$c^{\beta}(\mathbf{0}, y) = \mathbb{E}\,[\phi(\beta C_{\mathbf{0}})\phi(\beta C_y)] - \mathbb{E}\,[\phi(\beta S_k)]\mathbb{E}\,[\phi(\beta S_k)],$$

where $C_{\mathbf{0}}$ (respectively, C_y) denotes the length of the k-spacings starting at the origin (respectively, starting at y) with respect to the augmented point set $\mathcal{P} \cup \mathbf{0} \cup y$. Write $e_k := \mathbb{E}\,[\phi(\beta S_k)]$. Now

$$c^{\beta}(\mathbf{0}, y) = \mathbb{E}\,[(\phi(\beta C_{\mathbf{0}})\phi(\beta C_y) - \phi(\beta S_k)e_k)(1(y < S_k) + 1(y > S_k))].$$

The expectation over the latter term vanishes since

$$\mathbb{E}\,[(\phi(\beta C_{\mathbf{0}})\phi(\beta C_y) - \phi(\beta S_k)e_k)(1(y > S_k))] = \mathbb{E}\,[\phi(\beta S_k)(\phi(\beta C_y) - e_k)(1(y > S_k))]$$

$$= \mathbb{E}_{\overline{S}_k}[\phi(\beta S_k)(1(y > S_k))\mathbb{E}\left(\phi(\beta C_y) - e_k|\overline{S}_k)\right)] = 0,$$

as the restriction of $\mathcal{P}$ to $[y, \infty)$ is independent of $\overline{S}_k$.

Therefore for $y > 0$

$$\int_0^\infty c^\beta(\mathbf{0}, y)dy = \int_0^\infty \mathbb{E}\left[\phi(\beta C_{\mathbf{0}})\phi(\beta C_y) - \phi(\beta S_k)e_k)(1(y < S_k))\right]dy$$

$$= \sum_{l=1}^{k} \mathbb{E}\left[\phi(\beta C_{\mathbf{0}})\phi(\beta C_y) - \phi(\beta S_k)e_k) \cdot (1(S_{l-1} < y < S_l)\right]dy$$

$$= \sum_{l=1}^{k} \mathbb{E}\left[\phi(\beta S_k)\phi(\beta S_{k+l} - \beta S_l) - \phi(\beta S_{k+1})e_k\right].$$

The above clearly equals

$$\sum_{l=1}^{k} \mathbb{E}\left[\phi(\beta S_k)\phi(\beta S_{k+l} - \beta S_l) - \phi(\beta S_k)e_k + (\phi(\beta S_k) - \phi(\beta S_{k+1})e_k\right]$$

$$= \sum_{l=1}^{k} \mathrm{Cov}(\phi(\beta S_k), \phi(\beta S_{k+l} - \beta S_l)) + k(e_k - e_{k+1})e_k$$

$$= \sum_{l=1}^{k-1} \mathrm{Cov}(\phi(\beta S_k), \phi(\beta S_{k+l} - \beta S_l)) + k(e_k - e_{k+1})e_k,$$

since when $l = k$ the covariance vanishes. By symmetry, $\int_{-\infty}^0 c^\beta(\mathbf{0}, y)dy = \int_0^\infty c^\beta(\mathbf{0}, y)dy$ and thus

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[S^\beta_{\lambda,\phi,k}(\mathcal{P} \cap [0, 1])]}{\lambda} = V^S_{\phi,k}(\beta)$$

where $V^S_{\phi,k}(\beta)$ is defined in (2.13). This completes the proof of Theorem 2.2. □

6 Proof of Corollary 3.5

Corollary 3.5 follows by modifying the proof of Theorem 2.1 as follows.

Step 1. Step 1 of the proof of Theorem 2.1 showed that for all $\beta > 0$

$$\lim_{\lambda \to \infty} \frac{\mathrm{Var}[N^1_{\lambda,\phi}(\mathcal{P}_{\lambda\beta} \cap [0, 1]^d)]}{\beta\lambda} = V_\phi(1/\beta) := \overline{V}_\phi(\beta)$$

where we put $\overline{V}(\beta) := V(1/\beta)$. Likewise, we now find $V_t : \mathbb{R}^+ \to \mathbb{R}^+$ such that for all $\beta > 0$

$$V_t(\beta) := \lim_{\lambda \to \infty} \frac{\mathrm{Var}[H^t_\lambda(\mathcal{P}_{\beta\lambda} \cap [0, 1]^d)]}{\lambda\beta}. \tag{6.1}$$

Following Step 1 of the proof of Theorem 2.1 with $A = [0,1]^d,\ f \equiv \beta,\ h \equiv 1$ shows that the limit (6.1) equals

$$\int_{[0,1]^d} \mathbb{E}\,(\Phi^t(\mathbf{0}, \mathcal{P}_\beta))^2 dx +$$

$$+ \int_{[0,1]^d} \int_{\mathbb{R}^d} \mathbb{E}\,[\Phi^t(\mathbf{0}, \mathcal{P}_\beta \cup y)\Phi^t(y, \mathcal{P}_\beta \cup \mathbf{0})] - \mathbb{E}\,[\Phi^t(\mathbf{0}, \mathcal{P}_\beta)]\mathbb{E}\,[\Phi^t(y, \mathcal{P}_\beta)] dy dx.$$

Recalling $\Phi^t(x, \mathcal{X}) := \mathrm{card}(B_t(x) \cap \mathcal{X})$ and letting

$$c^\beta(\mathbf{0}, y) := \mathbb{E}\,[|B_t(\mathbf{0}) \cap \mathcal{P}_\beta \cup \mathbf{0} \cup \{y\}||B_t(y) \cap \mathcal{P}_\beta \cup \mathbf{0} \cup \{y\}|] - \mathbb{E}\,[|B_t(\mathbf{0}) \cap \mathcal{P}_\beta|]\mathbb{E}\,[|B_t(y) \cap \mathcal{P}_\beta|]$$

be the two point correlation function for Φ^t on $\mathcal{P}_\beta$ we get that (6.1) equals

$$\mathbb{E}\,[(\Phi^t(\mathbf{0}, \mathcal{P}_\beta))^2] + \int_{\mathbb{R}^d} c^\beta(\mathbf{0}, y) dy.$$

Writing $B_t(\mathbf{0}) \cap \mathcal{P}_\beta$ as the union of disjoint sets $B_t(\mathbf{0}) \cap B_t^c(y) \cap \mathcal{P}_\beta$ and $B_t(\mathbf{0}) \cap B_t(y) \cap \mathcal{P}_\beta$ and similarly for $B_t(y) \cap \mathcal{P}_\beta$, it follows that

$$\mathbb{E}\,[(\Phi^t(\mathbf{0}, \mathcal{P}_\beta))^2] = \mathbb{E}\,[|B_t(\mathbf{0}) \cap \mathcal{P}_\beta|^2]$$

and

$$c^\beta(\mathbf{0}, y) = \mathrm{Var}[|B_t(\mathbf{0}) \cap B_t(y) \cap \mathcal{P}_\beta|] = \beta|B_t(\mathbf{0}) \cap B_t(y)|.$$

Letting $v_t := t^d \omega_d$ we thus obtain

$$\int_{\mathbb{R}^d} c(\mathbf{0}, y) dy = \beta \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 1_{B_t(\mathbf{0})}(u) 1_{B_t(y)}(u) du dy$$

$$= \beta \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} 1_{B_t(\mathbf{0})}(u) 1_{B_t(y)}(u) dy du = \beta \int_{\mathbb{R}^d} 1_{B_t(\mathbf{0})}(u) \int_{\mathbb{R}^d} 1_{B_t(y)}(u) dy du = \beta v_t^2.$$

Notice also that $\mathbb{E}\,[|B_t(\mathbf{0}) \cap \mathcal{P}_\beta|^2] = \beta v_t(1 + \beta v_t)$. Collecting terms, it follows that for all $\beta > 0$

$$V_t(\beta) := \lim_{\lambda \to \infty} \frac{\mathrm{Var}[H_\lambda^t(\mathcal{P}_{\beta\lambda} \cap [0,1]^d)]}{\lambda\beta} = \mathbb{E}\,[|B_t(\mathbf{0}) \cap \mathcal{P}_\beta|^2] + \beta v_t^2 = v_t(1 + 2\beta v_t).$$

Step 2. We identify the average add one cost

$$\Delta_t(\beta) := \mathbb{E}\,\left[\lim_{r \to \infty} [H^t(\mathcal{P}_\beta \cap B_r(\mathbf{0}) \cup \mathbf{0}) - H^t(\mathcal{P}_\beta \cap B_r(\mathbf{0}))\right].$$

First, note that since the difference in the inside braces is the number of points in $\mathcal{P}_\beta$ within t of the origin it follows that

$$\Delta_t(\beta) = \mathbb{E}\,[N(\beta v_t)] = \beta v_t,$$

where $N(\tau)$ denotes an independent Poisson random variable with parameter τ.

It follows as in Step 2 of the proof of Theorem 2.1 that

$$\lim_{n\to\infty}\frac{\mathrm{Var}[H_n^t(X_1,...,X_n)]}{n}=\int_A V_t(f(x))f(x)dx-\left(\int_A \Delta_t(f(x))f(x)dx\right)^2$$

$$=v_t\int_A f^2(x)dx+v_t^2\left(2\int_A f^3(x)dx-(\int_A f^2(x)dx)^2\right),$$

which is the desired limit (3.2).

Step 3. We follow Step 3 of the proof of Theorem 2.1 to obtain Corollary 3.5. □

7 Proof of Proposition 3.1

We deduce Proposition 3.1 from the identity (2.5) as follows. From (2.5) we obtain

$$V_\phi(\beta)=\mathbb{E}\left[\phi^2(\beta\Gamma_1)\right]+\int_{\mathbb{R}^d}c(\mathbf{0},y)dy,$$

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

$$c(\mathbf{0},y)=\int_0^\infty\int_0^\infty\phi(\beta s)\phi(\beta t)\frac{\partial^2 p}{\partial s\partial t}dsdt-\left(\int_0^\infty\phi(\beta s)e^{-s}ds\right)^2$$

$$=\beta^2\int_0^\infty\int_0^\infty\phi'(\beta s)\phi'(\beta t)[p(s,t)-e^{-(s+t)}]dsdt,$$

using integration by parts and $\phi(0)=0$. Since $p(s,t)$ vanishes whenever $(s,t)\in[0,|y|^d\omega_d]$ we obtain

$$c(\mathbf{0},y)=\beta^2\int_0^{|y|^d\omega_d}\int_0^{|y|^d\omega_d}\phi'(\beta s)\phi'(\beta t)[e^{-(s+t)+I(s,t,|y|)}-e^{-(s+t)}]dsdt-$$

$$-\beta^2\int\int_{\max(s,t)\geq|y|^d\omega_d}\phi'(\beta s)\phi'(\beta t)e^{-(s+t)}dsdt.$$

Therefore

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

$$-\beta^2 \int_{\mathbb{R}^d} \int \int_{\max(s,t) \geq |y|^d \omega_d} \phi'(\beta s)\phi'(\beta t) e^{-(s+t)} ds dt dy$$

and making a change of variable on the outside integral yields

$$\int_{\mathbb{R}^d} c(\mathbf{0}, y) dy = \beta^2 \int_0^\infty \int_0^{|y|^d \omega_d} \int_0^{|y|^d \omega_d} \phi'(\beta s)\phi'(\beta t)[e^{-(s+t)+I(s,t,|y|)} - e^{-(s+t)}] ds dt d(|y|^d \omega_d) -$$

$$-\beta^2 \int_0^\infty \int \int_{\max(s,t) \geq |y|^d \omega_d} \phi'(\beta s)\phi'(\beta t) e^{-(s+t)} ds dt d(|y|^d \omega_d).$$

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

$$\int_0^\infty c(\mathbf{0}, y) dy = \beta^2 \int_0^\infty \int_0^u \int_0^u \phi'(\beta s)\phi'(\beta t) e^{-(s+t)} [e^{I(s,t,(u/\omega_d)^{1/d})} - 1] ds dt du -$$

$$-\beta^2 \int_0^\infty \int \int_{\max(s,t) \geq u} \phi'(\beta s)\phi'(\beta t) e^{-(s+t)} ds dt du.$$

Finally, change the order of integration to obtain

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

$$-\beta^2 \int_0^\infty \int_0^\infty \phi'(\beta s)\phi'(\beta t) e^{-(s+t)} \int_0^{\max(s,t)} du ds dt,$$

which is exactly the desired limit. This proves Proposition 3.1. □.

Acknowledgements: Y. 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. Statist. 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 (2004), Gaussian limits for random measures in geometric probability, Ann. Appl. Prob., to appear, Electronically available via `http://www.lehigh.edu/∼jey0/publications.html`

[6] 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.

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

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

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

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

[11] 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.

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

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

[14] 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.

[15] 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.

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

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

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

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

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

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

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

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

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

[25] Jimenez, R. and J. E. Yukich (2004), Statistical distances based on Euclidean graphs. Recent Advances in Probability and its Applications, Kluwer Acad. Pubs., to appear, ed. J. Glaz and H. Gyzl.

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

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

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

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

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

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

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

[33] Penrose, M. D. and J.E. Yukich (2004), Normal approximation in geometric probability, Proceedings of the Workshop on 'Stein's method and applications', Institute for Mathematical Sciences Lecture Notes Series, World Scientific Press, Singapore, to appear. Electronically available via `http://www.lehigh.edu/∼jey0/publications.html`.

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

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

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

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

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

[39] 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.

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

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

[42] 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.

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

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