Normal Approximation in Geometric Probability

Mathew D. Penrose[1] and J. E. Yukich[2]

University of Bath and Lehigh University

February 2004

Abstract

Statistics arising in geometric probability can often be expressed as sums of stabilizing functionals, that is functionals which satisfy a local dependence structure. In this note we show that stabilization leads to nearly optimal rates of convergence in the CLT for statistics such as total edge length and total number of edges of graphs in computational geometry and the total number of particles accepted in random sequential packing models. These rates also apply to the 1-dimensional marginals of the random measures associated with these statistics.

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

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

[2] Research supported in part by NSA grant MDA904-01-1-0029 and NSF grant DMS-0203720

1 Introduction and main results

In the study of limit theorems for functionals on Poisson or binomial spatial point processes, the notion of *stabilization* has recently proved to be a useful unifying concept [4, 9, 11]. Laws of large numbers and central limit theorems can be proved in the general setting of functionals satisfying an abstract 'stabilization' property whereby the insertion of a point into a Poisson process has only a local effect in some sense. These results can then be applied to deduce limit laws for a great variety of particular functionals, including those concerned with minimal spanning tree, nearest neighbor graph, Voronoi and Delaunay graph, and packing (see Section 2).

Several different techniques are available for proving general central limit theorems for stabilizing functionals. These include a martingale approach [9] and a method of moments [4]. In the present work, we revisit a third technique for proving central limit theorems for stabilizing functionals on Poisson point processes, which was introduced by Avram and Bertsimas [1]. This method is based on the normal approximation of sums of random variables which are 'mostly independent of one another' in a sense made precise via dependency graphs, which in turn is proved via Stein's method [12]. It has the advantage of providing the possibility of explicit error bounds and rates of convergence.

We extend the work of Avram and Bertsimas in several directions. First, whereas in [1] attention was restricted to certain particular functionals, here we derive a general result holding for arbitrary functionals satisfying a stabilization condition which can then be checked rather easily for many special cases. Second, we consider non-uniform point process intensities and do not require the functionals to be translation invariant. Third, we improve on the rates of convergence in [1] by making use of the recent refinement by Chen and Shao [6] of previous normal approximation results for sums of 'mostly independent' variables. Finally, we apply the methods not only to random variables obtained by summing some quantity over Poisson points, but to the associated *random point measures*, thereby recovering many of the results of Baryshnikov and Yukich [4] on convergence of measures, with extra information about the rate of convergence and without requiring higher order moment calculations.

Let $\xi(x;\mathcal{X})$ be a measurable $\mathbb{R}$-valued function defined for all pairs $(x,\mathcal{X})$, where $\mathcal{X} \subset \mathbb{R}^d$ is finite and where $x \in \mathcal{X}$. When $x \notin \mathcal{X}$, we abbreviate notation and write $\xi(x;\mathcal{X})$ instead of $\xi(x;\mathcal{X} \cup \{x\})$. For all $\lambda > 0$ let

$$\xi_\lambda(x;\mathcal{X}) := \xi(x; x + \lambda^{1/d}(-x + \mathcal{X}))$$

where given $a > 0$ and $y \in \mathbb{R}^d$, we let $a\mathcal{X} := \{ax : x \in \mathcal{X}\}$ and $y + \mathcal{X} := \{y + x : x \in \mathcal{X}\}$.

We say ξ is *translation invariant* if $\xi(x, \mathcal{X}) = \xi(y+x, y+\mathcal{X})$ for all $y \in \mathbb{R}^d$. When ξ is translation invariant, the functional ξ_λ simplifies to $\xi_\lambda(x; \mathcal{X}) = \xi(\lambda^{1/d}x; \lambda^{1/d}\mathcal{X})$.

Given a probability density function κ with compact support $A \subset \mathbb{R}^d$, for all $\lambda > 0$ we let $\mathcal{P}_\lambda := \mathcal{P}_{\lambda\kappa}$ denote a Poisson point process with intensity $\lambda\kappa$ on A. We shall assume throughout that κ is bounded with supremum denoted $\|\kappa\|_\infty$.

The following notion of exponential stabilization, introduced in [4], plays a central role in all that follows. For $x \in \mathbb{R}^d$ and $r > 0$, let $B_r(x)$ denote the Euclidean ball centered at x of radius r.

Definition 1.1 *ξ is exponentially stabilizing for κ if for all $\lambda \geq 1$ and all $x \in A$, there exists an a.s. finite random variable $R := R(x, \lambda)$ (*a radius of stabilization *for ξ at x) such that*

$$\xi_\lambda\left(x; [\mathcal{P}_\lambda \cap B_{\lambda^{-1/d}R}(x)] \cup \mathcal{X}\right)$$

is independent of $\mathcal{X}$ for all finite $\mathcal{X} \subset A \setminus B_{\lambda^{-1/d}R}(x)$ and there exists a constant $C > 0$ such that for all $t > 0$

$$\sup_{\lambda \geq 1,\ x \in A} P[R(x, \lambda) > t] \leq C \exp(-t/C).$$

Definition 1.2 *ξ has a moment of order $p > 0$ if*

$$\sup_{\lambda \geq 1,\ x \in \mathbb{R}^d} \mathbb{E}\left[|\xi_\lambda(x; \mathcal{P}_\lambda)|^p\right] < \infty. \tag{1.1}$$

For $\lambda > 0$, define the random weighted point measure

$$\mu_\lambda^\xi := \sum_{x \in \mathcal{P}_\lambda} \xi_\lambda(x; \mathcal{P}_\lambda)\delta_x.$$

and the centered version $\overline{\mu}_\lambda^\xi := \mu_\lambda^\xi - \mathbb{E}[\mu_\lambda^\xi]$. Let $B(A)$ denote the set of bounded Borel-measurable functions on A, and let $B_c(A)$ denote the set of continuous functions in $B(A)$. Given $f \in B(A)$, let $\langle f, \mu_\lambda^\xi \rangle := \int_A f d\mu_\lambda^\xi$ and $\langle f, \overline{\mu}_\lambda^\xi \rangle := \int_A f d\overline{\mu}_\lambda^\xi$.

Let Φ denote the distribution function of the standard normal. Our main result is a normal approximation result for $\langle f, \overline{\mu}_\lambda^\xi \rangle$, suitably scaled.

Theorem 1.1 *Let ξ be exponentially stabilizing and assume that ξ satisfies the moment condition (1.1) for some $p > 3$. Let $f \in B(A)$ and put $T_\lambda := \langle f, \mu_\lambda^\xi \rangle$. There exists a finite constant C depending on d, ξ, κ and f, such that for all $\lambda > 1$*

$$\sup_{t \in \mathbb{R}} \left| P\left[\frac{T_\lambda - \mathbb{E}\, T_\lambda}{(\mathrm{Var} T_\lambda)^{1/2}} \leq t\right] - \Phi(t) \right| \leq C(\log \lambda)^{3d} \lambda (\mathrm{Var} T_\lambda)^{-3/2}. \tag{1.2}$$

Remarks

(i) For many functionals of interest $\mathrm{Var}\langle f, \mu_\lambda^\xi \rangle = \Theta(\lambda)$ (see Remark (v) below). Whenever $\mathrm{Var}\langle f, \mu_\lambda^\xi \rangle = \Theta(\lambda)$, Theorem 1.1 yields a rate of convergence $O((\log \lambda)^{3d} \lambda^{-1/2})$ to the normal distribution. We are not sure if the logarithmic factors can be removed. The rate in [1] is $O((\log \lambda)^{1+3/(2d)} \lambda^{-1/4})$.

(ii) If in Theorem 1.1 we assume only that (1.1) holds for some $p > 2$ instead of some $p > 3$, and if $\mathrm{Var} T_\lambda = \Theta(\lambda)$, then the proof of Theorem 1.1 can be adapted to give a rate of convergence of $O(\lambda^{1+\varepsilon-p/2})$, for arbitrary $\varepsilon > 0$.

(iii) We do not have rate of convergence results in the binomial (non-Poisson) setting. For central limit theorems in the binomial setting, we refer to [9] and [4], which treat uniform and non-uniform samples respectively.

(iv) Some functionals, such as those defined in terms of the minimal spanning tree, satisfy a weaker form of stabilization but are not known to satisfy exponential stabilization. In these cases univariate and multivariate central limit theorems hold [8, 9] but our Theorem 1.1 does not apply and explicit rates of convergence are not known.

(v) In many cases, combining Theorem 1.1 with known results on the asymptotic behavior of $\mathrm{Var}(T_\lambda)$ yields central limit theorems. More precisely, it is established in Theorem 2.4(i) of [4], using methods developed in [9, 11], that if A is convex, κ is continuous, and ξ lies in a certain class of 'slowly varying' functionals SV(4/3) which includes all translation invariant functionals as a special case, and if certain exponential stabilization and pth moment conditions hold which are similar in spirit to those given in Definitions 1.1 and 1.2 above, for some $p > 2$, then then for all $f \in B_c(A)$,

$$\lim_{\lambda \to \infty} \lambda^{-1} \mathrm{Var}\langle f, \mu_\lambda^\xi \rangle = \int_A f(x)^2 V^\xi(\kappa(x)) \kappa(x) dx \tag{1.3}$$

with $V^\xi(\cdot)$ given explicitly in terms of ξ in [4]. Combining (1.3) with Theorem 1.1 yields

$$\langle f, \lambda^{-1/2} \overline{\mu}_\lambda^\xi \rangle \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \int_A f(x)^2 V^\xi(\kappa(x)) \kappa(x) dx\right),$$

where $\mathcal{N}(0, \sigma^2)$ denotes a centered normal distribution with variance σ^2 if $\sigma^2 > 0$, and a unit point mass at 0 if $\sigma^2 = 0$. Thus, when (1.3) holds we can use Theorem 1.1 to recover the conclusions of Theorem 2.4 (ii) of [4] (a central limit theorem for the finite-dimensional distributions of the random field $(\langle f, \lambda^{-1/2} \overline{\mu}_\lambda^\xi \rangle, f \in B_c(A))$), without any computation of higher order moments. This is characteristic of Stein's method.

(vi) Our Theorem 1.1 requires neither the underlying density function κ nor the test function f to be continuous (both of these conditions are imposed in [4]). In particular, Theorem 1.1 applies when f is the indicator function of a Borel subset B of A, giving normal approximation for $\bar{\mu}_\lambda^\xi(B)$.

2 Applications

Applications of Theorem 1.1 to geometric probability include functionals of proximity graphs, Boolean models, and random sequential packing models. The following examples are for illustrative purposes only and are not meant to be encyclopedic. For simplicity we will assume that $\mathbb{R}^d$ is equipped with the usual Euclidean metric. However, since we do not assume that ξ is translation invariant, the examples can be modified to treat the situation where $\mathbb{R}^d$ has a local metric structure.

2.1 k-nearest neighbor graph

Let k be a positive integer. Given a locally finite point set $\mathcal{X} \subset \mathbb{R}^d$, the k-nearest neighbors (undirected) graph on $\mathcal{X}$, denoted $kNG(\mathcal{X})$, is the graph with vertex set $\mathcal{X}$ obtained by including $\{x, y\}$ as an edge whenever y is one of the k nearest neighbors of x and/or x is one of the k nearest neighbors of y. The k-nearest neighbors (directed) graph on $\mathcal{X}$, denoted $kNG'(\mathcal{X})$, is the graph with vertex set $\mathcal{X}$ obtained by placing a directed edge between each point and its k nearest neighbors.

Let $N^k(\mathcal{X})$ denote the total edge length of the (undirected) k-nearest neighbors graph on $\mathcal{X}$. Note that $N^k(\mathcal{X}) = \sum_{x \in \mathcal{X}} \xi^k(x; \mathcal{X})$, where $\xi^k(x; \mathcal{X})$ denotes the sum of the edge lengths in $kNG(\mathcal{X})$ incident to x. If A is convex or polyhedral and κ is bounded away from 0 on A, then ξ^k is exponentially stabilizing (cf. Lemma 6.1 of [9]) and has moments of all orders. Moreover $\mathrm{Var}[N^k(\lambda^{1/d}\mathcal{P}_\lambda)] \geq C\lambda$. We thus obtain the following rates in the CLT for the total edge length of $N^k(\lambda^{1/d}\mathcal{P}_\lambda)$ improving upon Avram and Bertsimas [1] and Bickel and Breiman [5]. A similar CLT holds for the total edge length of the k-nearest neighbors directed graph.

Theorem 2.1 *Suppose A is convex or polyhedral and κ is bounded away from 0 on A. Let $N_\lambda := N^k(\lambda^{1/d}\mathcal{P}_\lambda)$ denote the total edge length of the k-nearest neighbors graph on $\lambda^{1/d}\mathcal{P}_\lambda$. There exists a finite constant C depending on d, ξ^k, and κ such that*

$$\sup_{t \in \mathbb{R}} \left| P\left[\frac{N_\lambda - \mathbb{E}\, N_\lambda}{(\mathrm{Var} N_\lambda)^{1/2}} \leq t \right] - \Phi(t) \right| \leq C(\log \lambda)^{3d} \lambda^{-1/2}. \tag{2.1}$$

Similarly, letting $\xi^s(x; \mathcal{X})$ be one or zero according to whether the distance between x and its nearest neighbor in $\mathcal{X}$ is less than s or not, we can verify that ξ^s is exponentially stabilizing and that the variance of $\sum_{x \in \lambda^{1/d}\mathcal{P}_\lambda} \xi^s(x; \lambda^{1/d}\mathcal{P}_\lambda)$ is bounded below by a positive multiple of λ. We thus obtain rates of convergence of $O((\log \lambda)^{3d}\lambda^{-1/2})$ in the CLT for the empirical distribution function of k nearest neighbor distances on $\lambda^{1/d}\mathcal{P}_\lambda$, improving upon those implicit on p. 88 of [7].

Using the results from section 6.2 of [9], we could likewise obtain the same rates of convergence in the CLT for the number of vertices of fixed degree in the k nearest neighbors graph.

2.2 Voronoi and sphere of influence graphs

We will consider the Voronoi graph for $d = 2$ and the sphere of influence graph for all d (see sections 7 and 8 of [9] for definitions). From the results of [4, 9, 11], we know that the total edge length of the Voronoi and sphere of influence graphs on $\mathcal{X}$ both admit the representation $\sum_{x \in \mathcal{X}} \xi(x; \mathcal{X})$; moreover, if κ is bounded away from 0 and infinity and A is convex, then ξ is exponentially stabilizing and satisfies the moment condition (1.1) for all $p > 1$. Also, the variance of the total edge length of these graphs on $\mathcal{P}_\lambda$ is bounded below by a multiple of λ. We thus obtain $O((\log \lambda)^{3d}\lambda^{-1/2})$ rates of convergence in the CLT for the total edge length functionals of these graphs on $\mathcal{P}_\lambda$, thereby improving and generalizing the results of Avram and Bertsimas [1].

For the sphere of influence graph we may draw on the results of sections 7.1 and 7.3 of [9], to obtain $O((\log \lambda)^{3d}\lambda^{-1/2})$ rates of convergence in the CLT for the total number of edges and the number of vertices of fixed degree in the sphere of influence graph on $\mathcal{P}_\lambda$.

2.3 Random sequential packing models

The following prototypical random sequential packing model is of considerable scientific interest; see [10] for references to the vast literature.

With $N(\lambda)$ standing for a Poisson random variable with parameter λ, we let $B_{\lambda,1}, B_{\lambda,2}, ..., B_{\lambda,N(\lambda)}$ be a sequence of d-dimensional balls of volume λ^{-1} whose centers are i.i.d. random d-vectors $X_1, ..., X_{N(\lambda)}$ with probability density function $\kappa : A \to [0, \infty)$. Without loss of generality, assume that the balls are sequenced in the order determined by marks (time coordinates) in $[0, 1]$. Let the first ball $B_{\lambda,1}$ be *packed*, and recursively for $i = 2, 3, \ldots$, let the i-th ball $B_{\lambda,i}$ be packed iff $B_{\lambda,i}$ does not overlap any ball in $B_{\lambda,1}, ..., B_{\lambda,i-1}$ which has already been packed. If not packed, the i-th ball is discarded. In much of the literature, the time coordinates are assumed independent of the spatial coordinates but since we do not need to confine attention to translation invariant models, we do not require this assumption here.

Packing models of this type arise in diverse disciplines, including physical, chemical, and biological processes. See [10] for a discussion of the many applications, the many references, and the widespread use. Penrose and Yukich [10] establish the asymptotic normality of the number of accepted balls when the spatial distribution is uniform and also show [11] a LLN for the number of accepted balls when the spatial distribution is non-uniform.

For any finite point set $\mathcal{X} \subset \mathbb{R}^d$, assume the points $x \in \mathcal{X}$ have time coordinates which are independent and uniformly distributed over the interval $[0, 1]$. Assume balls of volume λ^{-1} centered at the points of $\mathcal{X}$ arrive sequentially in an order determined by the time coordinates, and assume as before that each ball is packed or discarded according to whether or not it overlaps a previously packed ball. Let $\xi(x; \mathcal{X})$ be either 1 or 0 depending on whether the ball centered at x is packed or discarded. Consider the re-scaled packing functional $\xi_\lambda(x; \mathcal{X}) = \xi(\lambda^{1/d}x; \lambda^{1/d}\mathcal{X})$, where $\lambda^{1/d}x$ denotes scalar multiplication of x but *not* the mark associated with x and where balls centered at points of $\lambda^{1/d}\mathcal{X}$ have volume one. The random measure

$$\mu_\lambda^\xi := \sum_{i=1}^{N(\lambda)} \xi_\lambda(X_i; \{X_i\}_{i=1}^{N(\lambda)})\ \delta_{X_i},$$

is called the random sequential packing measure induced by balls with centers arising from κ. The convergence of the finite dimensional distributions of the packing measures μ_λ^ξ is established in [3, 4]. ξ is exponentially stabilizing [10, 3] and for any $f \in B_c([0, 1]^d)$ and κ uniform, the variance of $\langle f, \mu_\lambda^\xi \rangle$ is bounded below by a positive multiple of λ [4], showing that $\langle f, \mu_\lambda^\xi \rangle$ satisfies a CLT with an $O((\log \lambda)^{3d}\lambda^{-1/2})$ rate of convergence.

It follows easily from the stabilization analysis of [10] that many variants of the above basic packing model satisfy similar rates of convergence in the CLT. For example, the number of balls accepted in the cooperative sequential adsorption models and the monolayer ballistic deposition models of [10] both satisfy the CLT with an $O((\log \lambda)^{3d}\lambda^{-1/2})$ rate of convergence. The same comment applies for the number of seeds accepted in the spatial birth-growth models [10].

2.4 Independence number, off-line packing

An *independent set* of vertices in a graph G is a set of vertices in G, no two of which are connected by an edge. The *independence number* of G, which we denote $\beta(G)$, is defined to be the maximum cardinality of all independent sets of vertices in G.

For $r > 0$, and for finite or countable $\mathcal{X} \subset \mathbb{R}^d$, let $G(\mathcal{X}, r)$ denote the *geometric graph* with vertex set $\mathcal{X}$ and with edges between each pair of vertices distant at most r apart. Then the independence number $\beta(G(\mathcal{X}, r))$ is the maximum number of disjoint closed balls of radius $r/2$ that can be centered at points of $\mathcal{X}$; it is an 'off-line' version of the packing functionals considered in the previous section.

Let $b > 0$ be a constant, and consider the graph $G(\mathcal{P}_\lambda, b\lambda^{-1/d})$ (or equivalently, $G(\lambda^{1/d}\mathcal{P}_\lambda, b)$). Random geometric graphs of this type are the subject of [7], although independence number is considered only briefly there (on page 135). A law of large numbers for the independence number is described in Theorem 2.7 (iv) of [11].

For $\mu > 0$, let $\mathcal{H}_\mu$ denote a homogeneous Poisson point process of intensity μ on $\mathbb{R}^d$, and let $\mathcal{H}_\mu^0$ be the point process $\mathcal{H}_\mu$ with a point inserted at the origin. As on page 189 of [7], let λ_c be the infimum of all μ such that the origin has a non-zero probability of being in an infinite component of $G(\mathcal{H}_\mu, 1)$.

If $b^d\|\kappa\|_\infty < \lambda_c$, we can use Theorem 1.1 to obtain a central limit theorem for the independence number $\beta(G(\mathcal{P}_\lambda, b\lambda^{-1/d}))$.

We only sketch the proof. The graph $G(\mathcal{P}_\lambda, b\lambda^{-1/d})$ is isomorphic to $G(b^{-1}\lambda^{1/d}\mathcal{P}_\lambda, 1)$ and the point process $b^{-1}\lambda^{1/d}\mathcal{P}_\lambda$ is dominated by $\mathcal{H}_{b^d\|\kappa\|_\infty}$ (in the sense of [7], page 189). By exponential decay for subcritical continuum percolation (Lemma 10.2 of [7]) the probability that the component of $G(\mathcal{H}^0_{b^d\|\kappa\|_\infty}, 1)$ containing the origin includes a point distant more than r from the origin decays exponentially in r, and one can deduce exponential stabilization from this.

3 Proof of Theorem 1.1

3.1 A CLT for dependency graphs

We shall prove Theorem 1.1 by showing that exponential stabilization implies that a modification of $\langle f, \overline{\mu}_\lambda^\xi \rangle$ has a *dependency graph* structure, whose definition we now recall (see e.g. Chapter 2 of [7]). Let $X_\alpha,\ \alpha \in \mathcal{V}$, be a collection of random variables. The graph $G := (\mathcal{V}, \mathcal{E})$ is a *dependency graph* for $X_\alpha,\ \alpha \in \mathcal{V}$, if for any pair of disjoint sets $A_1, A_2 \subset \mathcal{V}$ such that no edge in $\mathcal{E}$ has one endpoint in A_1 and the other in A_2, the sigma-fields $\sigma\{X_\alpha, \alpha \in A_1\}$, and $\sigma\{X_\alpha, \alpha \in A_2\}$, are mutually independent. Let D denote the maximal degree of the dependency graph.

It is well known that sums of random variables indexed by the vertices of a dependency graph admit rates of convergence to a normal. The rates of Baldi and Rinott [2] and those in Penrose [7] are particularly useful; Avram and Bertsimas [1] use the former to obtain rate results for the total edge length of the nearest neighbor, Voronoi, and Delaunay graphs.

In many cases, the following theorem of Chen and Shao [6] provides superior rate results. We shall apply this result when $p = 3$. For any random variable X and any $p > 0$, let $||X||_p = (\mathbb{E}[|X|^p])^{1/p}$.

Lemma 3.1 *(see Thm 2.7 of [6]) Let $2 < p \leq 3$. Let $W_i,\ i \in \mathcal{V}$, be random variables indexed by the vertices of a dependency graph. Let $W = \sum_{i \in \mathcal{V}} W_i$. Assume that $\mathbb{E}[W^2] = 1, \mathbb{E}[W_i] = 0$, and $||W_i||_p \leq \theta$ for all $i \in \mathcal{V}$ and for some $\theta > 0$. Then*

$$\sup_t |P[W \leq t] - \Phi(t)| \leq 75 D^{5(p-1)} |\mathcal{V}| \theta^p. \tag{3.1}$$

3.2 Auxiliary lemmas

To prepare for the proof of Theorem 1.1 we will need some auxiliary lemmas. We assume throughout that $A \subseteq [0,1]^d$, but all of our results can be easily modified to treat the case of arbitrary compact sets $A \subset \mathbb{R}^d$. Throughout, C denotes a generic constant depending possibly on d, ξ, and κ and whose value may vary at each occurrence. We assume $\lambda > 1$ throughout.

Let $\alpha > 0$ be a constant to be chosen later. Given $\lambda > 0$, let $s_\lambda := \lambda^{-1/d} \alpha \log \lambda$, and cover $[0,1)^d$ by cubes of side s_λ of the form $\prod_{i=1}^d [j_i s_\lambda, (j_i + 1) s_\lambda)$, with all $j_i \in \mathbb{Z}$. Let the cubes in the covering be denoted $Q_1, Q_2, ..., Q_V$, where $V := V(\lambda)$ is the number of cubes in the covering, i.e. $V(\lambda) := \lceil s_\lambda^{-1} \rceil^d = \lceil \lambda^{1/d} / (\alpha \log \lambda) \rceil^d$.

For all $1 \leq i \leq V(\lambda)$, the number of points in $\mathcal{P}_\lambda \cap Q_i$ is a Poisson random variable $N_i := N(\tau_i)$, where $\tau_i := \lambda \int_{Q_i} \kappa(x) dx$. Assuming $\tau_i > 0$, choose an ordering on the points of $\mathcal{P}_\lambda \cap Q_i$ uniformly at random from all $(N_i)!$ possible such orderings. Use this ordering to list the points as $X_{i,1}, ..., X_{i,N_i}$, where conditional on the value of N_i, the random variables $X_{i,j},\ j = 1, 2, ...$ are i.i.d. on Q_i with a density $\kappa_i(\cdot) := \kappa(\cdot) / \int_{Q_i} \kappa(x) dx$. Thus we have the representation $\mathcal{P}_\lambda = \cup_{i=1}^{V(\lambda)} \{X_{i,j}\}_{j=1}^{N_i}$. For all $1 \leq i \leq V(\lambda)$, let $\mathcal{P}_i := \mathcal{P}_\lambda \setminus \{X_{i,j}\}_{j=1}^{N_i}$ and note that $\mathcal{P}_i$ is a Poisson point process on $[0,1]^d$ with intensity density $\lambda \kappa$ on $[0,1]^d \setminus Q_i$ and intensity zero on Q_i.

We show that the condition (1.1), which bounds the moments of the value of ξ at points inserted into $\mathcal{P}_\lambda$, also yields bounds on $\mathbb{E}[|\xi_\lambda(X_{i,j}; \mathcal{P}_\lambda) \cdot 1_{j \leq N_i}|^p]$. More precisely, we have

Lemma 3.2 *Let $p > 0$. If (1.1) holds, then there is a constant C such that for all $\lambda > 1$ and all $1 \leq i \leq V(\lambda)$*

$$\mathbb{E}[|\xi_\lambda(X_{i,j}; \mathcal{P}_\lambda) \cdot 1_{j \leq N_i}|^p] \leq C (\log \lambda)^d. \tag{3.2}$$

Proof. If $N_i = n$, then denote $\{X_{i,1}, ..., X_{i,N_i}\}$ by $\mathcal{X}_n$. We have by definition

$$\mathbb{E}\left[|\xi_\lambda(X_{i,j};\mathcal{P}_\lambda)\cdot 1_{j\le N_i}|^p\right] = \sum_{n=j}^{\infty}\int_{Q_i}\mathbb{E}\left[|\xi_\lambda(x;\mathcal{X}_{n-1}\cup\mathcal{P}_i)|^p\right]\kappa_i(x)dx\cdot P[N_i=n],$$

where the expectation on the right hand side is with respect to $\mathcal{X}_{n-1}$ and $\mathcal{P}_i$. The above is bounded by

$$\le \tau_i\sum_{n=1}^{\infty}\int_{Q_i}\mathbb{E}\left[|\xi(x;\mathcal{X}_{n-1}\cup\mathcal{P}_i)|^p\right]\kappa_i(x)dx\cdot\frac{e^{-\tau_i}\tau_i^{n-1}}{(n-1)!}$$

$$= \tau_i\sum_{m=0}^{\infty}\int_{Q_i}\mathbb{E}\left[|\xi_\lambda(x;\mathcal{P}_\lambda)|^p \;\;|\;\; |\mathcal{P}_\lambda\cap Q_i| = m\right]\kappa_i(x)dx\cdot P[|\mathcal{P}_\lambda\cap Q_i|=m]$$

$$= \tau_i\int_{Q_i}\mathbb{E}\left[|\xi_\lambda(x;\mathcal{P}_\lambda\cup x)|^p\right]\kappa_i(x)dx \le \text{const.}\times\tau_i,$$

where the last inequality follows by (1.1). Since $\tau_i = \lambda\int_{Q_i}\kappa(x)dx \le \|\kappa\|_\infty(\alpha\log\lambda)^d$, this shows (3.2). ∎

Fix $1 \le i \le V$. For all $j = 1, 2, ...$ we define

$$\xi_j := \xi_{i,j} := \xi_\lambda(X_{i,j};\mathcal{P}_\lambda)$$

when $1 \le j \le N_i$ and otherwise we set $\xi_j = 0$.

Lemma 3.3 *If (1.1) holds for some* $p > 3$*, then*

$$\left\|\sum_{j=1}^{\infty}|\xi_j|\right\|_3^3 \le C(\log\lambda)^{4d}.$$

Proof. Clearly, with $N := N_i$ and $\tau := \tau_i$,

$$\left\|\sum_{j=1}^{\infty}|\xi_j|\right\|_3 = \left\|\sum_{j=1}^{\infty}|\xi_j|\left(1_{N\le\tau}+\sum_{k=0}^{\infty}1_{2^k\tau<N\le 2^{k+1}\tau}\right)\right\|_3$$

$$\le \left\|\sum_{j=1}^{\infty}\sum_{k=0}^{\infty}|\xi_j|\cdot 1_{2^k\tau<N\le 2^{k+1}\tau}\right\|_3 + \left\|\sum_{j=1}^{\infty}|\xi_j|\cdot 1_{N\le\tau}\right\|_3.$$

Since a.s. only finitely many summands in the double sum are non-zero, by subadditivity of the norm, the above is bounded by

$$\le \sum_{k=0}^{\infty}\left\|\sum_{j=1}^{\infty}|\xi_j|\cdot 1_{2^k\tau<N\le 2^{k+1}\tau}\right\|_3 + \left\|\sum_{j=1}^{\tau}|\xi_j|\cdot 1_{N\le\tau}\right\|_3$$

$$\leq \sum_{k=0}^{\infty} \left|\left| \sum_{j=1}^{2^{k+1}\tau} |\xi_j| \cdot 1_{N \geq 2^k \tau} \right|\right|_3 + \left|\left| \sum_{j=1}^{\tau} |\xi_j| \cdot 1_{N \leq \tau} \right|\right|_3$$

$$\leq \sum_{k=0}^{\infty} \sum_{j=1}^{2^{k+1}\tau} \left|\left| \xi_j \cdot 1_{N \geq 2^k \tau} \right|\right|_3 + \sum_{j=1}^{\tau} \left|\left| \xi_j \cdot 1_{N \leq \tau} \right|\right|_3 . \tag{3.3}$$

Hölder's inequality yields for all $1 \leq j \leq 2^{k+1}\tau$ and any $0 < \delta < 1/9$:

$$||\xi_j \cdot 1_{N \geq 2^k \tau}||_3 \leq ||\xi_j||_{3+3\delta} \cdot (P[N \geq 2^k \tau])^{\delta/(3+3\delta)}$$

and therefore replacing 3δ with δ gives

$$||\xi_j \cdot 1_{N \geq 2^k \tau}||_3 \leq ||\xi_j||_{3+\delta} \cdot (P[N \geq 2^k \tau])^{\delta/10}. \tag{3.4}$$

Now by (3.2) we have

$$||\xi_j||_{3+\delta} \leq C(\log \lambda)^{d/(3+\delta)}. \tag{3.5}$$

Substituting (3.4) and (3.5) into (3.3) we obtain

$$\left|\left| \sum_{j=1}^{\infty} |\xi_j| \right|\right|_3 \leq C(\log \lambda)^{d/(3+\delta)} \sum_{k=0}^{\infty} \tau 2^{k+1} \cdot (P[N \geq 2^k \tau])^{\delta/10} + \sum_{j=1}^{\tau} \left|\left| \xi_j \cdot 1_{N \leq \tau} \right|\right|_3 . \tag{3.6}$$

Now for $\tau > 1$ we have

$$\sum_{k=0}^{\infty} \tau 2^{k+1} \cdot (P[N \geq 2^k \tau])^{\delta/10} \leq \sum_{k=0}^{\infty} \tau 2^{k+1} \cdot (P[N \geq 2^k])^{\delta/10} \leq C\tau$$

whereas for $0 < \tau < 1$ we have

$$\sum_{k=0}^{\infty} \tau 2^{k+1} \cdot (P[N \geq 2^k \tau])^{\delta/10}$$

$$\leq \sum_{k=0}^{\lceil \log_2 \frac{1}{\tau} \rceil + 2} \tau 2^{k+1} + \sum_{k=\lceil \log_2 \frac{1}{\tau} \rceil + 3}^{\infty} \tau 2^{k+1} \cdot (P[N \geq 2^k \tau])^{\delta/10}$$

$$\leq C + \sum_{k=\lceil \log_2 \frac{1}{\tau} \rceil + 3}^{\infty} \tau 2^{k+1} \cdot (\exp(-2^{k-1}\tau \cdot k))^{\delta/10} \leq C,$$

where the penultimate inequality follows from bounds for the tail of a Poisson (see e.g. (1.12) in [7]). Thus the first sum on the right hand side of (3.6) is at most $C(\log \lambda)^{4d/3}$ since $\tau \leq C(\log \lambda)^d$. Since $||\xi_j||_3 \leq C(\log \lambda)^{d/3}$ we find that (3.6) implies

$$\left|\left| \sum_{j=1}^{\infty} |\xi_j| \right|\right|_3 \leq C(\log \lambda)^{4d/3}, \tag{3.7}$$

which concludes the proof of Lemma 3.3. □

3.3 Conclusion of proof of Theorem 1.1

Throughout this section, we fix $f \in B(A)$ and set $T_\lambda := \langle f, \mu_\lambda^\xi \rangle$. For all $1 \le i \le V$ and all $j = 1, 2, ...$, let $R_{i,j}$ denote the radius of stabilization of ξ at $\lambda^{1/d} X_{i,j}$ if $1 \le j \le N_i$ and let $R_{i,j}$ be zero otherwise.

Let $E_{i,j} := \{R_{i,j} \le \alpha \log \lambda\}$. Then by standard Palm theory (e.g. Theorem 1.6 in [7]) $\mathbb{E}\left[\sum_{i=1}^V \sum_{j=1}^{N_i} 1_{E_{i,j}^c}\right] = \int_{[0,\lambda^{1/d}]^d} P[R(x,\lambda) > \alpha \log \lambda] \kappa(\lambda^{-1/d} x) dx \le C\lambda^{-33}$ by exponential stabilization if α is large enough. Let $E_\lambda := \cap_{i=1}^V \cap_{j=1}^\infty E_{i,j}$ and note that $P[E_\lambda^c] \le C\lambda^{-33}$.

Recalling the representation $\mathcal{P}_\lambda = \cup_{i=1}^{V(\lambda)} \{X_{i,j}\}_{j=1}^{N_i}$, we have

$$T_\lambda = \sum_{i=1}^{V(\lambda)} \sum_{j=1}^{N_i} \xi_\lambda(X_{i,j}; \mathcal{P}_\lambda) \cdot f(X_{i,j}).$$

To obtain rates of normal approximation for T_λ, it will be be convenient to consider a closely related sum enjoying more independence structure, namely

$$T'_\lambda := \sum_{i=1}^{V(\lambda)} \sum_{j=1}^{N_i} \xi_\lambda(X_{i,j}; \mathcal{P}_\lambda) \cdot 1_{E_{i,j}} \cdot f(X_{i,j}).$$

For all $1 \le i \le V(\lambda)$ define

$$S_i := S_{Q_i} := (\mathrm{Var} T'_\lambda)^{-1/2} \sum_{j=1}^{N_i} \xi_\lambda(X_{i,j}; \mathcal{P}_\lambda) \cdot 1_{E_{i,j}} \cdot f(X_{i,j})$$

and put $S := (\mathrm{Var} T'_\lambda)^{-1/2}(T'_\lambda - \mathbb{E}\, T'_\lambda) = \sum_{i=1}^{V(\lambda)} (S_i - \mathbb{E}\, S_i)$. Clearly $\mathrm{Var} S = \mathbb{E}\, S^2 = 1$.

We define a graph $G_\lambda := (\mathcal{V}_\lambda, \mathcal{E}_\lambda)$ as follows. The set $\mathcal{V}_\lambda$ consists of the subcubes $Q_1, ..., Q_{V(\lambda)}$ and edges (Q_i, Q_j) belong to $\mathcal{E}_\lambda$ if $d(Q_i, Q_j) \le 2\alpha\lambda^{-1/d} \log \lambda$, where $d(Q_i, Q_j) := \inf\{|x - y| : x \in Q_i, y \in Q_j\}$. By definition of exponential stabilization, we note that if A_1 and A_2 are disjoint collections of cubes in $\mathcal{V}_\lambda$ such that no edge in $\mathcal{E}_\lambda$ has one endpoint in A_1 and one endpoint in A_2, then the random variables $\{S_{Q_i}, Q_i \in A_1\}$ and $\{S_{Q_j}, Q_j \in A_2\}$ are independent. Thus G_λ is a dependency graph.

To prepare for an application of Lemma 3.1, we make four observations:

(i) $V(\lambda) := |\mathcal{V}_\lambda| = \lceil \lambda^{1/d} / (\alpha \log \lambda) \rceil^d$.

(ii) Since the number of cubes in $Q_1, ..., Q_V$ distant at most $2\alpha\lambda^{-1/d} \log \lambda$ from a given cube is bounded by 5^d, it follows that the maximal degree D satisfies $D := D_\lambda \le 5^d$.

(iii) The definitions of S_i and $\xi_{i,j}$ and Lemma 3.3 tell us that for all $1 \le i \le V(\lambda)$

$$\mathbb{E}\left[|S_i|^3\right] \le C(\mathrm{Var} T'_\lambda)^{-3/2} \mathbb{E}\left[\left(\sum_{j=1}^{\infty} |\xi_{i,j}|\right)^3\right] \le C(\mathrm{Var} T'_\lambda)^{-3/2} (\log \lambda)^{4d}.$$

(iv) $\mathrm{Var}[T'_\lambda]$ is close to $\mathrm{Var}[T_\lambda]$ for λ large. We require a few estimations to show this. Note that $|T'_\lambda - T_\lambda| = 0$ except possibly on the set E^c_λ which has probability less than $C\lambda^{-33}$. Lemma 3.3, along with Minkowski's inequality, yields the upper bound

$$\mathbb{E}\left[\left(\sum_{i=1}^{V(\lambda)}\sum_{j=1}^{N_i}|\xi_\lambda(X_{i,j};\mathcal{P}_\lambda)|\right)^3\right] \leq CV(\lambda)^3(\log\lambda)^{4d} \leq C\lambda^4. \tag{3.8}$$

Thus Hölder's inequality implies that

$$\begin{aligned}
&\mathbb{E}\left[|T_\lambda - T'_\lambda|^2\right] \\
&\leq \mathbb{E}\left[|T_\lambda - T'_\lambda|^2 \cdot 1_{E^c_\lambda}\right] \leq 4\mathbb{E}\left[(T^2_\lambda + {T'}^2_\lambda)\cdot 1_{E^c_\lambda}\right] \\
&\leq 8\left(\mathbb{E}\left[\sum_{i=1}^{V(\lambda)}\left(\sum_{j=1}^{N_i}|\xi_\lambda(X_{i,j};\mathcal{P}_\lambda)|\right)^3\right]\right)^{2/3}\cdot(P[E^c_\lambda])^{1/3} \leq C\lambda^{-8}
\end{aligned} \tag{3.9}$$

and thus

$$\mathbb{E}\left[|T'_\lambda - T_\lambda|\right] \leq C\lambda^{-4}. \tag{3.10}$$

Additionally, (3.8) and Jensen's inequality yield

$$\mathbb{E}\left[\left(\sum_{i=1}^{V(\lambda)}\sum_{j=1}^{N_i}|\xi_\lambda(X_{i,j};\mathcal{P}_\lambda)|\right)^2\right] \leq C\lambda^3. \tag{3.11}$$

Since

$$\mathrm{Var}[T'_\lambda] = \mathrm{Var}[T_\lambda] + \mathrm{Var}(T'_\lambda - T_\lambda) + 2\mathrm{Cov}(T_\lambda, T'_\lambda - T_\lambda),$$

by (3.9), (3.11) and the Cauchy-Schwarz inequality we obtain

$$|\mathrm{Var}[T'_\lambda] - \mathrm{Var}[T_\lambda]| \leq C\lambda^{-2}. \tag{3.12}$$

Given the four observations (i)-(iv), we are now ready to apply Lemma 3.1 to prove Theorem 1.1. Trivially, (1.2) holds for large enough λ when $\mathrm{Var}[T_\lambda] < 1$, and so without loss of generality we now assume $\mathrm{Var}[T_\lambda] \geq 1$. To establish the rate of convergence (1.2) in this case, we apply the bound (3.1) to $W_i := S_i - \mathbb{E}\, S_i,\ 1 \leq i \leq V_\lambda$, with $p = 3$ and with

$$\theta := C(\mathrm{Var}T'_\lambda)^{-1/2}(\log\lambda)^{4d/3}.$$

Our choice of θ is applicable because of observation (iii). We clearly have $\mathbb{E}[W_i] = 0$ and $\mathbb{E}[(\sum_{i=1}^{V_\lambda} W_i)^2] = 1$. With $S = \sum_{i=1}^{V_\lambda} W_i$, Lemma 3.1 yields

$$\sup_t |P[S \leq t] - \Phi(t)| \leq C\lceil\frac{\lambda^{1/d}}{\alpha\log\lambda}\rceil^d \cdot\ (\mathrm{Var}T'_\lambda)^{-3/2}(\log\lambda)^{4d}$$

$$\leq C\lambda(\mathrm{Var}T_\lambda)^{-3/2}(\log\lambda)^{3d}, \tag{3.13}$$

where the last line makes use of the fact that $\mathrm{Var}[T'_\lambda] \geq \mathrm{Var}[T_\lambda]/2$, which follows from (3.12).

Now if $\beta > 0$ is a constant and Z any random variable then by (3.13) we have for all $t \in \mathbb{R}$

$$P[Z \leq t] \leq P[S \leq t+\beta] + P[|Z-S| \geq \beta]$$

$$\leq \Phi(t+\beta) + C\lambda(\mathrm{Var}T_\lambda)^{-3/2}(\log\lambda)^{3d} + P[|Z-S| \geq \beta]$$

$$\leq \Phi(t) + C\beta + C\lambda(\mathrm{Var}T_\lambda)^{-3/2}(\log\lambda)^{3d} + P[|Z-S| \geq \beta]$$

by the Lipschitz property of Φ. Similarly for all $t \in \mathbb{R}$

$$P[Z \leq t] \geq \Phi(t) - C\beta - C\lambda(\mathrm{Var}T_\lambda)^{-3/2}(\log\lambda)^{3d} - P[|Z-S| \geq \beta].$$

In other words

$$\sup_t |P[Z \leq t] - \Phi(t)| \leq C\beta + C\lambda(\mathrm{Var}T_\lambda)^{-3/2}(\log\lambda)^{3d} + P[|Z-S| \geq \beta]. \tag{3.14}$$

Now by definition of S,

$$|(\mathrm{Var}T'_\lambda)^{-1/2}(T_\lambda - \mathbb{E}\,T_\lambda) - S| = |(\mathrm{Var}T'_\lambda)^{-1/2}\{(T_\lambda - \mathbb{E}\,T_\lambda) - (T'_\lambda - \mathbb{E}\,T'_\lambda)\}|$$

$$\leq (\mathrm{Var}T'_\lambda)^{-1/2}\{|T_\lambda - T'_\lambda| + \mathbb{E}\,[|T_\lambda - T'_\lambda|]\}$$

which by (3.10) is bounded by $C\lambda^{-4}$ except possibly on the set E_λ^c which has probability less than $C\lambda^{-33}$. Thus by (3.14) with $Z = (\mathrm{Var}T'_\lambda)^{-1/2}(T_\lambda - \mathbb{E}\,T_\lambda)$ and $\beta = C\lambda^{-4}$

$$\sup_t |P[(\mathrm{Var}T'{}_\lambda)^{-1/2}(T_\lambda - \mathbb{E}\,T_\lambda) \leq t] - \Phi(t)| \leq C\lambda(\mathrm{Var}T_\lambda)^{-3/2}(\log\lambda)^{3d} + C\lambda^{-4} + C\lambda^{-33}.$$

Moreover, by the triangle inequality

$$\sup_t \left| P[(\mathrm{Var}T_\lambda)^{-1/2}(T_\lambda - \mathbb{E}\,T_\lambda) \leq t] - \Phi(t) \right| \leq$$

$$\leq \sup_t \left| P\left[(\mathrm{Var}T'_\lambda)^{-1/2}(T_\lambda - \mathbb{E}\,T_\lambda) \leq t \cdot (\frac{\mathrm{Var}T_\lambda}{\mathrm{Var}T'_\lambda})^{1/2}\right] - \Phi\left(t(\frac{\mathrm{Var}T_\lambda}{\mathrm{Var}T'_\lambda})^{1/2}\right)\right| +$$

$$+ \sup_t \left|\Phi\left(t(\frac{\mathrm{Var}T_\lambda}{\mathrm{Var}T'_\lambda})^{1/2}\right) - \Phi(t)\right|.$$

Since for all $s \leq t$, we have $|\Phi(s) - \Phi(t)| \leq (t-s)\max_{s \leq u \leq t}\phi(u)$ where ϕ denotes the standard normal density, and since by (3.12) there is a constant $0 < C < \infty$ such that for all $\lambda > 0$ and all $t \in \mathbb{R}$

$$\left|t(\frac{\mathrm{Var}T_\lambda}{\mathrm{Var}T'_\lambda})^{1/2} - t\right| \leq |t| \left|\frac{\mathrm{Var}T_\lambda}{\mathrm{Var}T'_\lambda} - 1\right| \leq \frac{C|t|}{\lambda^2}$$

we get

$$\sup_t \left| \Phi \left(t(\frac{\mathrm{Var} T_\lambda}{\mathrm{Var} T'_\lambda})^{1/2} \right) - \Phi(t) \right| \le C \sup_t \left(\left(\frac{|t|}{\lambda^2} \right) \left(\max_{u \in [t - tC/\lambda^2,\ t + tC/\lambda^2]} \phi(u) \right) \right) \le \frac{C}{\lambda^2}.$$

Thus,

$$\sup_t |P[(\mathrm{Var} T_\lambda)^{-1/2}(T_\lambda - \mathbb{E}\, T_\lambda) \le t] - \Phi(t)| \le C\lambda (\mathrm{Var} T_\lambda)^{-3/2} (\log \lambda)^{3d} + C\lambda^{-2}. \tag{3.15}$$

Finally we assert that

$$\mathrm{Var} T_\lambda = O((\log \lambda)^{8d/3} \lambda). \tag{3.16}$$

To see this, observe that T'_λ is the sum of $V(\lambda)$ random variables, which by Jensen's inequality and Lemma 3.3 each have a second moment bounded by a constant multiple of $(\log \lambda)^{8d/3}$. Thus the variance of each of the $V(\lambda)$ random variables is also bounded by a constant multiple of $(\log \lambda)^{8d/3}$. Moreover, the covariance of any pair of the $V(\lambda)$ random variables is zero when the indices of the random variables correspond to non-adjacent sub-cubes. For adjacent sub-cubes, the covariance is also bounded by a constant multiple of $(\log \lambda)^{8d/3}$. This shows that $\mathrm{Var} T'_\lambda = O((\log \lambda)^{8d/3} \lambda)$, and combined with (3.12) this yields (3.16).

By (3.16), in (3.15) the first term in the right hand side dominates, thus yielding the desired bound (1.2), and the proof of Theorem 1.1 is complete. □

Acknowledgments. We began this work while visiting the Institute for Mathematical Sciences at the National University of Singapore, and continued it while visiting the Isaac Newton Institute for Mathematical Sciences at Cambridge. We thank both institutions for their hospitality.

References

[1] F. Avram and D. Bertsimas (1993), On central limit theorems in geometrical probability. *Ann. Appl. Probab.* **3,** 1033-1046.

[2] P. Baldi and Y. Rinott (1989), Asymptotic normality of some graph related statistics. *J. Appl. Probab.* **26**, 171-175.

[3] Yu. Baryshnikov and J.E. Yukich (2003), Gaussian fields and random packing. *J. Statist. Phys.* **111**, 443-463.

[4] Yu. Baryshnikov and J.E. Yukich (2004), Gaussian limits for random measures in geometric probability. *Ann. Appl. Probab.*, to appear, Electronically available via `http://www.lehigh.edu/~jey0/publications.html`

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

[6] L. Chen and Q.-M. Shao (2003), Normal approximation under local dependence, preprint, *Ann. Probab.*, to appear.

[7] M.D. Penrose (2003), Random Geometric Graphs, Oxford University Press.

[8] M.D. Penrose (2004). Multivariate spatial central limit theorems with applications to percolation and spatial graphs. Preprint, University of Bath. M.D. Penrose. Electronically available via `http://www.bath.ac.uk/math-sci/`

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

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

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

[12] C. Stein (1972), Approximate Computation of Expectations. IMS, Hayward, CA.