Adv. Appl. Prob. (SGSA) **39,** 326–342 (2007)
Printed in Northern Ireland

EXPLICIT LAWS OF LARGE NUMBERS FOR RANDOM NEAREST-NEIGHBOUR-TYPE GRAPHS

ANDREW R. WADE,* *University of Bristol*

Abstract

Under the unifying umbrella of a general result of Penrose and Yukich (*Annals of Applied Probability* **13** (2003), 277–303) we give laws of large numbers (in the L^p sense) for the total power-weighted length of several nearest-neighbour-type graphs on random point sets in $\mathbb{R}^d$, $d \in \mathbb{N}$. Some of these results are known; some are new. We give limiting constants explicitly, where previously they have been evaluated in less generality or not at all. The graphs we consider include the k-nearest-neighbours graph, the Gabriel graph, the minimal directed spanning forest, and the on-line nearest-neighbour graph.

Keywords: Nearest-neighbour-type graph; law of large numbers; spanning forest; spatial network evolution

2000 Mathematics Subject Classification: Primary 60D05
Secondary 60F25

1. Introduction

Graphs constructed on random point sets in $\mathbb{R}^d$, $d \in \mathbb{N}$, formed by joining nearby points according to some deterministic rule, have recently received considerable interest [19], [29], [31]. Such graphs include the geometric graph, the minimal spanning tree, and (as studied in this paper) the nearest-neighbour graph and its relatives. Applications include the modelling of spatial networks, as well as statistical procedures.

The graphs in this paper are based on edges between nearest neighbours, sometimes in some restricted sense. A unifying characteristic of these graphs is *stabilization*: roughly speaking, the configuration of edges around any particular vertex is not affected by changes to the vertex set outside some sufficiently large (but finite) ball. Thus, these graphs are locally determined in some sense.

A functional of particular interest is the total edge length of the graph, or, more generally, the total power-weighted edge length (i.e. the sum of the edge lengths each raised to a given power $\alpha \geq 0$). The large-sample asymptotic theory for the power-weighted length of stabilizing graphs is now well understood; see, e.g. [15], [19], [20], [24], [25], [29], and [31].

In the present paper we collect several laws of large numbers (LLNs) for the total power-weighted length for the family of nearest-neighbour-type graphs, defined on independent random points on $\mathbb{R}^d$. We present these results as corollaries to a general umbrella theorem of Penrose and Yukich [25]. Some of the results (for the most common graphs) are known to various extents in the literature; others are new. We take a unified approach which highlights the connections between these results.

In particular, all our results are explicit: we give explicit expressions for limiting constants. In some cases these constants have been seen previously in the literature.

Received 3 May 2006; revision received 13 November 2006.
* Postal address: Department of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, UK.
Email address: andrew.wade@bristol.ac.uk

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

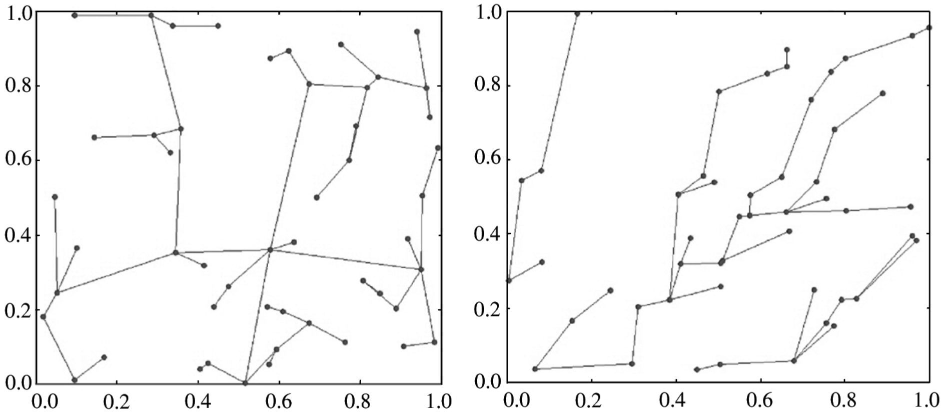

FIGURE 1: Realizations of the ONG (*left*) and MDSF under '$\preccurlyeq^*$' (*right*), each on 50 simulated uniform random points in the unit square.

Nearest-neighbour graphs and nearest-neighbour distances in $\mathbb{R}^d$ are of interest in several areas of applied science, including the social sciences, geography, and ecology, where proximity data are often important (see, e.g. [16] and [27]). Ad hoc networks, in which nodes scattered in space are connected according to some geometric rule, are of interest with respect to various types of communications networks. Quantities of interest such as the overall network throughput may be related to the power-weighted length.

In the analysis of multivariate data, in particular via nonparametric statistics, nearest-neighbour graphs and near-neighbour distances have found many applications, including goodness-of-fit tests, classification, regression, noise estimation, density estimation, dimension identification, cluster analysis, and the two-sample and multisample problems; see, e.g. [6], [7], [8], [10], [12], [13], and [30], and references therein.

In this paper we give a new LLN for the total power-weighted length of the *on-line nearest-neighbour graph* (*ONG*), which is one of the simplest models of network evolution. We give a detailed description later. In the ONG on a sequence of points arriving in $\mathbb{R}^d$, each point after the first is joined by an edge to its nearest predecessor. The ONG appeared in [4] as a simple model for the evolution of the Internet graph. Figure 1 shows a sample realization of an ONG.

Recently, graphs with an 'on-line' structure, in which vertices are added one by one and connected to existing vertices via some rule, have been the subject of considerable study in relation to the modelling of real-world networks. The ONG is one of the simplest network evolution models that captures some of the observed characteristics of real-world networks, such as spatial structure and sequential growth.

We also consider the *minimal directed spanning forest* (*MDSF*). The MDSF is constructed on a partially ordered point set in $\mathbb{R}^d$ by connecting each point to its nearest neighbour amongst those points (if any) that precede it in the partial order. If an MDSF is a tree, it is called a *minimal directed spanning tree* (*MDST*). The MDST was introduced by Bhatt and Roy in [5] as a model for drainage or communications networks, in $d = 2$, with the 'coordinatewise' partial order '$\preccurlyeq^*$', such that $(x_1, y_1) \preccurlyeq^* (x_2, y_2)$ if and only if $x_1 \leq x_2$ and $y_1 \leq y_2$. In this version of the MDSF, each point is joined by an edge to its nearest neighbour in its 'south-western' quadrant. In the present paper we give new LLNs for the total power-weighted length for a family of MDSFs indexed by partial orderings on $\mathbb{R}^2$, which include '$\preccurlyeq^*$' as a special case. Figure 1 shows an example of an MDSF under '$\preccurlyeq^*$'.

2. Notation and results

Notions of *stabilizing* functionals of point sets have recently proved to be a useful basis for establishing limit theorems for functionals of random point sets in $\mathbb{R}^d$. In particular, Penrose and Yukich [24], [25] proved general central limit theorems and laws of large numbers for stabilizing functionals.

The LLNs we give in the present paper are all derived ultimately from Theorem 2.1 of [25], which we restate as Theorem 1 below before we present our results.

In order to describe the result of [25], we need to introduce some notation. Let $d \in \mathbb{N}$. Let $\|\cdot\|$ be the Euclidean norm on $\mathbb{R}^d$, and let $|\cdot|$ denote d-dimensional Lebesgue measure. Write $\operatorname{card}(\mathcal{X})$ for the cardinality of a finite set $\mathcal{X} \subset \mathbb{R}^d$. For a locally finite point set $\mathcal{X} \subset \mathbb{R}^d$, $a > 0$, and $\boldsymbol{y} \in \mathbb{R}^d$, let $\boldsymbol{y} + a\mathcal{X}$ denote the set $\{\boldsymbol{y} + a\boldsymbol{x} : \boldsymbol{x} \in \mathcal{X}\}$. Let $B(\boldsymbol{x}; r)$ denote the closed Euclidean ball with centre $\boldsymbol{x} \in \mathbb{R}^d$ and radius $r > 0$. Let $\mathbf{0}$ denote the origin in $\mathbb{R}^d$.

Let $\xi(\boldsymbol{x}; \mathcal{X})$ be a measurable, $[0, \infty)$-valued function defined for all pairs $(\boldsymbol{x}, \mathcal{X})$, where $\mathcal{X} \subset \mathbb{R}^d$ is finite and $\boldsymbol{x} \in \mathcal{X}$. Assume that ξ is translation invariant, i.e. that, for all $\boldsymbol{y} \in \mathbb{R}^d$, $\xi(\boldsymbol{y} + \boldsymbol{x}; \boldsymbol{y} + \mathcal{X}) = \xi(\boldsymbol{x}; \mathcal{X})$. When $\boldsymbol{x} \notin \mathcal{X}$, we abbreviate the notation $\xi(\boldsymbol{x}; \mathcal{X} \cup \{\boldsymbol{x}\})$ to $\xi(\boldsymbol{x}; \mathcal{X})$. For our applications, ξ will be *homogeneous of order* $\alpha \geq 0$, i.e. such that $\xi(r\boldsymbol{x}; r\mathcal{X}) = r^\alpha \xi(\boldsymbol{x}; \mathcal{X})$ for all $r > 0$, all finite point sets $\mathcal{X}$, and all $\boldsymbol{x} \in \mathcal{X}$.

For any locally finite point set $\mathcal{X} \subset \mathbb{R}^d$ and any $\ell \in \mathbb{N}$, define

$$\xi^+(\mathcal{X}; \ell) := \sup_{k \in \mathbb{N}} \operatorname{ess\,sup}\{\xi(\mathbf{0}; (\mathcal{X} \cap B(\mathbf{0}; \ell)) \cup \mathcal{A}^*) : \mathcal{A} \in (\mathbb{R}^d \setminus B(\mathbf{0}; \ell))^k\},$$
$$\xi^-(\mathcal{X}; \ell) := \inf_{k \in \mathbb{N}} \operatorname{ess\,inf}\{\xi(\mathbf{0}; (\mathcal{X} \cap B(\mathbf{0}; \ell)) \cup \mathcal{A}^*) : \mathcal{A} \in (\mathbb{R}^d \setminus B(\mathbf{0}; \ell))^k\},$$

where for $\mathcal{A} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k) \in (\mathbb{R}^d)^k$ we let $\mathcal{A}^* = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ (provided that all k vectors are distinct). Define the *limit* of ξ on $\mathcal{X}$ by

$$\xi_\infty(\mathcal{X}) := \limsup_{\ell \to \infty} \xi^+(\mathcal{X}; \ell).$$

We say that the functional ξ *stabilizes* on $\mathcal{X}$ if

$$\lim_{\ell \to \infty} \xi^+(\mathcal{X}; \ell) = \lim_{\ell \to \infty} \xi^-(\mathcal{X}; \ell) = \xi_\infty(\mathcal{X}).$$

Stabilization can be interpreted loosely as the property that the value of the functional at a point is unaffected by changes in the configuration of points at a sufficiently large distance from that point.

Let f be a probability density function on $\mathbb{R}^d$. For $n \in \mathbb{N}$, let $\mathcal{X}_n := (\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_n)$ be the point process consisting of n independent, random d-vectors with common density f. With probability 1, $\mathcal{X}_n$ has distinct interpoint distances; hence, all the nearest-neighbour-type graphs on $\mathcal{X}_n$ that we consider are almost surely unique.

Let $\mathcal{H}_1$ be a homogeneous Poisson point process of unit intensity on $\mathbb{R}^d$. The following general LLN is due to Penrose and Yukich, and is obtained from Theorem 2.1 of [25] together with Equation (2.9) there (the homogeneous case). Here and subsequently we write $\operatorname{supp}(f)$ for the support of function f and $\xrightarrow{L^p}$ for convergence in L^p, $p \geq 1$.

Theorem 1. *Let $q \in \{1, 2\}$. Suppose that ξ is homogeneous of order α and almost surely stabilizes on $\mathcal{H}_1$, with limit $\xi_\infty(\mathcal{H}_1)$. If ξ satisfies the moments condition*

$$\sup_{n \in \mathbb{N}} \mathrm{E}[\xi(n^{1/d}\boldsymbol{X}_1; n^{1/d}\mathcal{X}_n)^p] < \infty \tag{1}$$

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

for some $p > q$, then, as $n \to \infty$,

$$n^{-1} \sum_{x \in \mathcal{X}_n} \xi(n^{1/d} x; n^{1/d} \mathcal{X}_n) \xrightarrow{L^q} \mathrm{E}[\xi_\infty(\mathcal{H}_1)] \int_{\mathrm{supp}(f)} f(x)^{(d-\alpha)/d} \, \mathrm{d}x,$$

and the limit is finite.

From this result we will derive LLNs for the total power-weighted length for a collection of nearest-neighbour-type graphs. Let $j \in \mathbb{N}$. A point $x \in \mathcal{X}$ has a *jth-nearest neighbour* $y \in \mathcal{X} \setminus \{x\}$ if $\mathrm{card}(\{z : z \in \mathcal{X} \setminus \{x\}, \|z - x\| < \|y - x\|\}) = j - 1$. For all $x, y \in \mathbb{R}^d$, we define the weight function

$$w_\alpha(x, y) := \|x - y\|^\alpha,$$

for some fixed parameter $\alpha \geq 0$. By the total power-weighted edge length of a graph with edge set E (where edges may be directed or undirected), we mean the functional

$$\sum_{(u,v) \in E} w_\alpha(u, v) = \sum_{(u,v) \in E} \|u - v\|^\alpha.$$

We will often assume one of the following conditions on the function f: either

(C1) f is supported by a convex polyhedron in $\mathbb{R}^d$ and is bounded away from 0 and infinity on its support; or,

(C2) for weight exponent $\alpha \in [0, d)$, we require that $\int_{\mathbb{R}^d} f(x)^{(d-\alpha)/d} \, \mathrm{d}x < \infty$ and $\int_{\mathbb{R}^d} \|x\|^r f(x) \, \mathrm{d}x < \infty$ for some $r > d/(d - \alpha)$.

In some cases we take $f(x) = 1$ for $x \in (0, 1)^d$ and $f(x) = 0$ otherwise, and denote by $\mathcal{X}_n = \mathcal{U}_n = (U_1, U_2, \ldots, U_n)$ the binomial point process consisting of n independent, uniform random vectors on $(0, 1)^d$.

In the remainder of this section we present our LLNs derived from Theorem 1. Theorems 2, 3, and 6 follow directly from Theorem 1 and results of [25], up to evaluation of constants, while Theorems 4 and 5 need some more work. These results are natural companions, as are their proofs, which we present in Section 3 below; in particular, the proof of Theorem 2 is useful for the other proofs.

2.1. The k-nearest-neighbours and jth-nearest-neighbour graphs

Let $j \in \mathbb{N}$. In the jth-nearest-neighbour (directed) graph on $\mathcal{X}$, denoted by jth-NNG$'(\mathcal{X})$, a directed edge joins each point of $\mathcal{X}$ to its jth-nearest neighbour. Let $k \in \mathbb{N}$. In the k-nearest-neighbours (directed) graph on $\mathcal{X}$, denoted by k-NNG$'(\mathcal{X})$, a directed edge joins each point of $\mathcal{X}$ to each of its first k nearest neighbours in $\mathcal{X}$ (i.e. each of its jth-nearest neighbours, for $j = 1, 2, \ldots, k$). Clearly the 1st-NNG$'$ and the 1-NNG$'$ coincide, giving the standard nearest-neighbour (directed) graph. See Figure 2 for realizations of particular examples of the jth-NNG$'$ and the k-NNG$'$. We also consider the k-nearest-neighbours (undirected) graph on $\mathcal{X}$, denoted by k-NNG$(\mathcal{X})$, in which an undirected edge joins $x, y \in \mathcal{X}$ if x is one of the first k nearest neighbours of y or y is one of the first k nearest neighbours of x (or both).

From now on we take the point set $\mathcal{X}$ to be *random*; in particular, for $n \in \mathbb{N}$, we take $\mathcal{X} = \mathcal{X}_n$. For $d \in \mathbb{N}$ and $\alpha \geq 0$, let $\mathcal{L}_j^{d,\alpha}(\mathcal{X}_n)$ and $\mathcal{L}_{\leq k}^{d,\alpha}(\mathcal{X}_n)$ respectively denote the total

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

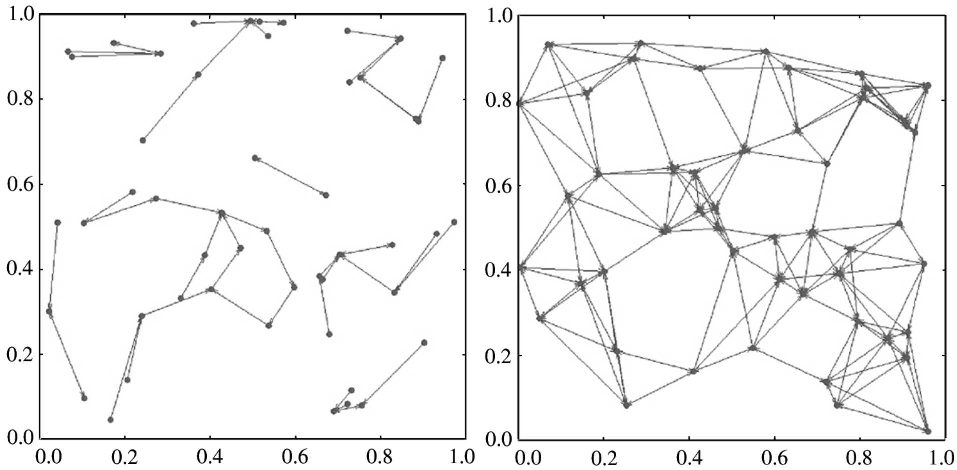

FIGURE 2: Realizations of 3rd-NNG′ (*left*) and 5-NNG′ (*right*), each on 50 simulated uniform random points in the unit square.

power-weighted edge length of the jth-nearest-neighbour (directed) graph and the k-nearest-neighbours (directed) graph on $\mathcal{X}_n \subset \mathbb{R}^d$. Note that

$$\mathcal{L}^{d,\alpha}_{\leq k}(\mathcal{X}_n) = \sum_{j=1}^{k} \mathcal{L}^{d,\alpha}_{j}(\mathcal{X}_n). \tag{2}$$

For $d \in \mathbb{N}$, we denote the volume of the unit d-ball (see, e.g. Equation (6.50) of [14]) by

$$v_d := \pi^{d/2}[\Gamma(1 + d/2)]^{-1}. \tag{3}$$

Theorems 2 and 4 below feature constants $C(d, \alpha, k)$ defined, for $d, k \in \mathbb{N}$ and $\alpha \geq 0$, by

$$C(d, \alpha, k) := v_d^{-\alpha/d} \frac{d}{d + \alpha} \frac{\Gamma(k + 1 + \alpha/d)}{\Gamma(k)}. \tag{4}$$

Our first result is Theorem 2, which gives LLNs for $\mathcal{L}^{d,\alpha}_{j}(\mathcal{X}_n)$ and $\mathcal{L}^{d,\alpha}_{\leq k}(\mathcal{X}_n)$, with explicit expressions for the limiting constants; it is the natural starting point for our LLNs for nearest-neighbour-type graphs. Recall that under condition (C1) $\operatorname{supp}(f)$ is a convex polyhedron, and under condition (C2) $\operatorname{supp}(f)$ is $\mathbb{R}^d$.

Theorem 2. *Let $d \in \mathbb{N}$. The following results hold, with $p = 2$, for $\alpha \geq 0$ if f satisfies condition (C1), and, with $p = 1$, for $\alpha \in [0, d)$ if f satisfies condition (C2).*

(a) *For the jth-NNG′ on $\mathbb{R}^d$, as $n \to \infty$ we have*

$$n^{(\alpha-d)/d} \mathcal{L}^{d,\alpha}_{j}(\mathcal{X}_n) \xrightarrow{L^p} v_d^{-\alpha/d} \frac{\Gamma(j + \alpha/d)}{\Gamma(j)} \int_{\operatorname{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d} \, \mathrm{d}\boldsymbol{x}. \tag{5}$$

(b) *For the k-NNG′ on $\mathbb{R}^d$, as $n \to \infty$ we have*

$$n^{(\alpha-d)/d} \mathcal{L}^{d,\alpha}_{\leq k}(\mathcal{X}_n) \xrightarrow{L^p} C(d, \alpha, k) \int_{\operatorname{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d} \, \mathrm{d}\boldsymbol{x}. \tag{6}$$

In particular, as $n \to \infty$,

$$n^{(\alpha-d)/d} \mathcal{L}^{d,\alpha}_{\le k}(\mathcal{U}_n) \xrightarrow{L^p} C(d, \alpha, k). \tag{7}$$

Remark 1. (a) If we use a norm on $\mathbb{R}^d$ different from the Euclidean, Theorem 2 remains valid with v_d redefined as the volume of the unit d-ball in the chosen norm.

(b) Theorem 2 is essentially contained in Theorem 2.4 of [25], but with the constants evaluated explicitly. There are several related LLN results in the literature. Theorem 8.3 of [31] gives LLNs (with complete convergence) for $\mathcal{L}^{d,1}_{\le k}(\mathcal{X}_n)$ (see also [17]); the limiting constants are not given. Avram and Bertsimas (in Theorem 7 of [2]) stated a result on the limiting expectation (and, hence, the constant in the LLN) for $\mathcal{L}^{2,1}_j(\mathcal{U}_n)$, which they attribute to Miles [18] (see also [31, p. 101]). The constant in [2] was given as

$$\frac{1}{2}\pi^{-1/2} \sum_{i=1}^{j} \frac{\Gamma(i-\frac{1}{2})}{\Gamma(i)},$$

which simplifies (by induction on j) to $\pi^{-1/2}\Gamma(j+\frac{1}{2})/\Gamma(j)$, the $d = 2$, $\alpha = 1$ case of (5) for $\mathcal{X}_n = \mathcal{U}_n$.

(c) Related results are the asymptotic expectations of jth-nearest-neighbour distances in finite point sets given in [9] and [26]. The results of [26] are consistent with the $\alpha = 1$ case of (7) here. The result of [9] includes general-α and certain nonuniform densities, although the conditions on f there are more restrictive than condition (C1); the result is consistent with (6). Also, [9] contains (in Equation (6.4) there) a weak LLN for the empirical mean k-nearest-neighbour *distance*. With Theorem 2.4 of [25], the results of [9] yield LLNs for the total weight of the jth-NNG′ and the k-NNG′ only for $d - 1 < \alpha < d$ (due to the rates of convergence given in [9]).

(d) Smith [28] gave, in some sense, expectations of randomly selected edge lengths for nearest-neighbour-type graphs on the homogeneous Poisson point process of unit intensity in $\mathbb{R}^d$, including the jth-NNG′, the nearest-neighbour (undirected) graph, and the Gabriel graph. His results coincide with ours only for the jth-NNG′, since here each vertex contributes a fixed number (j) of directed edges: Equation (5.4.1) of [28] matches our expression for $C(d, 1, k)$.

From the results on nearest-neighbour (directed) graphs, we may obtain results for nearest-neighbour (undirected) graphs, in which if $\boldsymbol{x}$ is a nearest neighbour of $\boldsymbol{y}$ and vice versa, then the edge between $\boldsymbol{x}$ and $\boldsymbol{y}$ is counted only once. As an example, we give the following result, where for $d \in \mathbb{N}$ and $\alpha \ge 0$ we let $\mathcal{N}^{d,\alpha}(\mathcal{X}_n)$ denote the total power-weighted edge length of the nearest-neighbour (undirected) graph on $\mathcal{X}_n \subset \mathbb{R}^d$ and we let ω_d be the volume of the union of two unit d-balls with centres a unit distance apart in $\mathbb{R}^d$.

Theorem 3. *Suppose that* $d \in \mathbb{N}$, *that* $\alpha \ge 0$, *and that* f *satisfies condition (C1). As* $n \to \infty$,

$$n^{(\alpha-d)/d} \mathcal{N}^{d,\alpha}(\mathcal{X}_n) \xrightarrow{L^2} \Gamma\left(1+\frac{\alpha}{d}\right)\left(v_d^{-\alpha/d} - \frac{1}{2} v_d \omega_d^{-1-\alpha/d}\right) \int_{\operatorname{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d}\, \mathrm{d}\boldsymbol{x}. \tag{8}$$

In particular, for $d = 2$ *and* $\alpha \ge 0$ *we have*

$$n^{(\alpha-2)/2} \mathcal{N}^{2,\alpha}(\mathcal{U}_n) \xrightarrow{L^2} \Gamma\left(1+\frac{\alpha}{2}\right)\left(\pi^{-\alpha/2} - \frac{\pi}{2}\left(\frac{6}{8\pi + 3\sqrt{3}}\right)^{1+\alpha/2}\right), \tag{9}$$

and for $d = 2$ and $\alpha = 1$, we have

$$n^{-1/2}\mathcal{N}^{2,1}(\mathcal{U}_n) \xrightarrow{L^2} \frac{1}{2} - \frac{1}{4}\left(\frac{6\pi}{8\pi + 3\sqrt{3}}\right)^{3/2} \approx 0.377508. \tag{10}$$

Finally, for $d = 1$ and $\alpha = 1$ we have $\mathcal{N}^{1,1}(\mathcal{U}_n) \xrightarrow{L^2} \frac{7}{18}$ as $n \to \infty$.

Remark 2. A pair of points, each of which is the other's nearest neighbour, is known as a reciprocal pair. Reciprocal pairs are of interest in ecology (see [27, pp. 158–159]. For $\alpha = 0$, $\mathcal{N}^{d,\alpha}(\mathcal{X}_n)$ counts the number of vertices minus one half of the number of reciprocal pairs. In this case (8) says that $n^{-1}\mathcal{N}^{d,0}(\mathcal{X}_n) \xrightarrow{L^2} 1 - (v_d/(2\omega_d))$. This is consistent with results of Henze [12] for the fraction of points that are the ℓth-nearest neighbours of their own kth-nearest neighbours; in particular (see [12] and references therein), as $n \to \infty$ the probability that a point is in a reciprocal pair tends to v_d/ω_d.

2.2. The on-line nearest-neighbour graph

We now consider the on-line nearest-neighbour graph (ONG). Let $d \in \mathbb{N}$. Suppose that $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ are points in $(0,1)^d$, arriving sequentially; for $n \in \mathbb{N}$, form a graph on vertex set $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ by connecting each point $\boldsymbol{x}_i$, $i = 2, 3, \ldots, n$, to its nearest neighbour amongst its predecessors (i.e. $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}$), using the lexicographic ordering on $\mathbb{R}^d$ to break any ties. The resulting tree is the ONG on $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$.

Again, we take our sequence of points to be random. We restrict our analysis to the case in which we have independent, uniformly distributed points $\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots$ on $(0,1)^d$. For $d \in \mathbb{N}$ and $\alpha \geq 0$, let $\mathcal{O}^{d,\alpha}(\mathcal{U}_n)$ denote the total power-weighted edge length of the ONG on sequence $\mathcal{U}_n = (\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n)$. The next result gives a new LLN for $\mathcal{O}^{d,\alpha}(\mathcal{U}_n)$ for $\alpha < d$.

Theorem 4. *Suppose that $d \in \mathbb{N}$ and $\alpha \in [0, d)$. With $C(d, \alpha, k)$ as given in (4), as $n \to \infty$ we have*

$$n^{(\alpha-d)/d}\mathcal{O}^{d,\alpha}(\mathcal{U}_n) \xrightarrow{L^1} \frac{d}{d-\alpha}C(d,\alpha,1) = \frac{d}{d-\alpha}v_d^{-\alpha/d}\Gamma\left(1 + \frac{\alpha}{d}\right).$$

Related results include those on the convergence in distribution of $\mathcal{O}^{d,\alpha}(\mathcal{U}_n)$, given in [23] for $\alpha > d$ ($\alpha > \frac{1}{2}$ in the case in which $d = 1$) and in [20] in the form of a central limit theorem for $\alpha \in (0, \frac{1}{4})$. Also, the ONG in $d = 1$ is related to the 'directed linear tree' considered in [22].

2.3. The minimal directed spanning forest

The minimal directed spanning forest (MDSF) is related to the standard nearest-neighbour (directed) graph, with the additional constraint that edges can only lie in a given direction. In general, the MDSF can be defined as a global optimization problem for directed graphs on partially ordered sets endowed with a weight function, and it also admits a local construction; see [5], [21], and [22]. As above, we consider the Euclidean setting, where our points lie in $\mathbb{R}^d$.

Suppose that $\mathcal{X} \in \mathbb{R}^d$ is a finite set bearing a partial order '$\preccurlyeq$'. A *minimal element*, or *sink*, of $\mathcal{X}$ is a vertex $\boldsymbol{v}_0 \in \mathcal{X}$ for which there exists no $\boldsymbol{v} \in \mathcal{X} \setminus \{\boldsymbol{v}_0\}$ such that $\boldsymbol{v} \preccurlyeq \boldsymbol{v}_0$. Let $\mathcal{S}$ denote the set of all sinks of $\mathcal{X}$. (Note that $\mathcal{S}$ cannot be empty.)

For $\boldsymbol{v} \in \mathcal{X}$, we say that $\boldsymbol{u} \in \mathcal{X} \setminus \{\boldsymbol{v}\}$ is a *directed nearest neighbour* of $\boldsymbol{v}$ if $\boldsymbol{u} \preccurlyeq \boldsymbol{v}$ and $\|\boldsymbol{v} - \boldsymbol{u}\| \leq \|\boldsymbol{v} - \boldsymbol{u}'\|$ for all $\boldsymbol{u}' \in \mathcal{X} \setminus \{\boldsymbol{v}\}$ such that $\boldsymbol{u}' \preccurlyeq \boldsymbol{v}$. For each $\boldsymbol{v} \in \mathcal{X} \setminus \mathcal{S}$, let $n_{\boldsymbol{v}}$ be a directed nearest neighbour of $\boldsymbol{v}$ (chosen arbitrarily if $\boldsymbol{v}$ has more than one). Then (see [21]) the directed graph on $\mathcal{X}$ obtained by taking edge set $E := \{(\boldsymbol{v}, n_{\boldsymbol{v}}) \colon \boldsymbol{v} \in \mathcal{X} \setminus \mathcal{S}\}$ is an MDSF of $\mathcal{X}$. Thus, if all edge-weights are distinct, the MDSF is unique and is obtained by connecting each nonminimal vertex to its directed nearest neighbour. In the case in which there is a single sink, the MDSF is a tree (ignoring directedness of edges) and it is called the minimal directed spanning tree (MDST).

In what follows, we consider a general type of partial order on $\mathbb{R}^2$, denoted '$\overset{\theta,\phi}{\preccurlyeq}$', specified by the angles $\theta \in [0, 2\pi)$ and $\phi \in (0, \pi]$. For $\boldsymbol{x} \in \mathbb{R}^2$, let $C_{\theta,\phi}(\boldsymbol{x})$ be the closed half-cone of angle ϕ with vertex $\boldsymbol{x}$ and boundaries given by the rays from $\boldsymbol{x}$ at angles θ and $\theta+\phi$, measuring anticlockwise from the upwards vertical. The partial order is such that, for $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^2$,

$$\boldsymbol{x}_1 \overset{\theta,\phi}{\preccurlyeq} \boldsymbol{x}_2 \quad \Longleftrightarrow \quad \boldsymbol{x}_1 \in C_{\theta,\phi}(\boldsymbol{x}_2). \tag{11}$$

We shall use '$\preccurlyeq^*$' as shorthand for the special case of this partial order with $\theta = \phi = \pi/2$, which is of particular interest, as in [5]. In this case $(u_1, u_2) \preccurlyeq^* (v_1, v_2)$ if and only if $u_1 \le v_1$ and $u_2 \le v_2$. The symbol '$\preccurlyeq$' will denote a general partial order on $\mathbb{R}^2$. Note that for $\phi = \pi$ (11) does not, in fact, define a partial order on the whole of $\mathbb{R}^2$, since the antisymmetric property ($\boldsymbol{x} \preccurlyeq \boldsymbol{y}$ and $\boldsymbol{y} \preccurlyeq \boldsymbol{x}$ imply that $\boldsymbol{x} = \boldsymbol{y}$) fails; however, with probability 1 it is a true partial order (in fact a total order) on the random point sets that we consider.

We do not permit here the choice $\phi = 0$, which would almost surely give us a disconnected point set. Nor do we allow $\phi \in (\pi, 2\pi]$, since in this case the transitivity property ($\boldsymbol{u} \preccurlyeq \boldsymbol{v}$ and $\boldsymbol{v} \preccurlyeq \mathbf{w}$ imply that $\boldsymbol{u} \preccurlyeq \boldsymbol{w}$) fails and, so, the directional relation (11) is not a partial order.

Again we take $\mathcal{X}$ to be random; set $\mathcal{X} = \mathcal{X}_n$, where (as before) $\mathcal{X}_n$ is a point process consisting of n independent, random points on $(0, 1)^2$ with common density f. When the partial order is '$\preccurlyeq^*$', as in [5], we also consider the point set $\mathcal{X}_n^0 := \mathcal{X}_n \cup \{\mathbf{0}\}$ (where $\mathbf{0}$ is the origin in $\mathbb{R}^2$) on which the MDSF is an MDST rooted at $\mathbf{0}$.

In this random setting, almost surely each nonminimal point of $\mathcal{X}$ has a unique directed nearest neighbour, meaning that $\mathcal{X}$ has a unique MDSF. For $\alpha > 0$, denote by $\mathcal{M}^\alpha(\mathcal{X})$ the total power-weighted edge length, with weight exponent α, of the MDSF on $\mathcal{X}$.

Theorem 5 presents LLNs for $\mathcal{M}^\alpha(\mathcal{X}_n)$ in the uniform case, in which $\mathcal{X}_n = \mathcal{U}_n$. However, the proof carries through for other distributions. In particular, if the points of $\mathcal{X}_n$ are distributed in $\mathbb{R}^2$ with a density f that satisfies condition (C1), then (12) holds with a factor of $\int_{\mathrm{supp}(f)} f(\boldsymbol{x})^{(2-\alpha)/2}\,\mathrm{d}\boldsymbol{x}$ introduced into the right-hand side.

Theorem 5. *Let $d \in \mathbb{N}$ and $\alpha \in (0, 2)$. Under partial order '$\overset{\theta,\phi}{\preccurlyeq}$' with $\theta \in [0, 2\pi)$ and $\phi \in (0, \pi]$, as $n \to \infty$ we have*

$$n^{(\alpha-2)/2}\mathcal{M}^\alpha(\mathcal{U}_n) \xrightarrow{L^1} \left(\frac{2}{\phi}\right)^{\alpha/2} \Gamma\left(1 + \frac{\alpha}{2}\right). \tag{12}$$

Moreover, when the partial order is '$\preccurlyeq^*$', (12) remains true with $\mathcal{U}_n$ replaced by $\mathcal{U}_n^0$.

2.4. The Gabriel graph

In the *Gabriel graph* (see [11]) on point set $\mathcal{X} \subset \mathbb{R}^d$, two points are joined by an edge if and only if the ball that has the line segment joining those two points as a diameter contains no other points of $\mathcal{X}$. The Gabriel graph has been applied in many of the same contexts as nearest-neighbour graphs; see, e.g. [30].

For $d \in \mathbb{N}$ and $\alpha \ge 0$, let $\mathcal{G}^{d,\alpha}(\mathcal{X})$ denote the total power-weighted edge length of the Gabriel graph on $\mathcal{X} \subset \mathbb{R}^d$. As before, we consider the random point set $\mathcal{X}_n$ with underlying

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

density f. An LLN for $\mathcal{G}^{d,\alpha}(\mathcal{X}_n)$ was given in [25]; in the present paper we give the limiting constant explicitly.

Theorem 6. *Let $d \in \mathbb{N}$ and $\alpha \geq 0$, and suppose that f satisfies condition (C1). As $n \to \infty$,*

$$n^{(\alpha-d)/d}\mathcal{G}^{d,\alpha}(\mathcal{X}_n) \xrightarrow{L^2} v_d^{-\alpha/d}2^{d+\alpha-1}\Gamma\left(1+\frac{\alpha}{d}\right)\int_{\operatorname{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d}\,\mathrm{d}\boldsymbol{x}.$$

3. Proofs

3.1. Proof of Theorems 2 and 3

For $j \in \mathbb{N}$, let $d_j(\boldsymbol{x}; \mathcal{X})$ be the (Euclidean) distance from $\boldsymbol{x}$ to its jth-nearest neighbour in $\mathcal{X} \setminus \{\boldsymbol{x}\}$, if such a neighbour exists, and let it equal 0 otherwise. We will use the following form of Euler's gamma integral (see Equation 6.1.1 of [1]): for $a, b, c \geq 0$,

$$\int_0^\infty r^a \exp(-cr^b)\,\mathrm{d}r = \frac{1}{b}c^{-(a+1)/b}\Gamma((a+1)/b). \tag{13}$$

Proof of Theorem 2. In applying Theorem 1 to the jth-NNG′ and k-NNG′ functionals, we take $\xi(\boldsymbol{x}; \mathcal{X}_n)$ to be $d_j(\boldsymbol{x}; \mathcal{X}_n)^\alpha$, where $\alpha \geq 0$. Then ξ is translation invariant and homogeneous of order α. It was shown in Theorem 2.4 of [25] that the jth-NNG′ total weight functional ξ satisfies the conditions of Theorem 1 in the following two cases: (i) with $q = 2$, if f satisfies condition (C1), and $\alpha \geq 0$; and (ii) with $q = 1$, if f satisfies condition (C2), and $0 \leq \alpha < d$. (In fact, in [25] this was proved for the k-NNG′ functional $\sum_{j=1}^k d_j(\boldsymbol{x}; \mathcal{X}_n)^\alpha$, but this implies that the conditions also hold for the jth-NNG′ functional $d_j(\boldsymbol{x}; \mathcal{X}_n)^\alpha$.)

The functional $\xi(\boldsymbol{x}; \mathcal{X}_n) = d_j(\boldsymbol{x}; \mathcal{X}_n)^\alpha$ stabilizes on $\mathcal{H}_1$, with limit $\xi_\infty(\mathcal{H}_1) = d_j(\boldsymbol{0}; \mathcal{H}_1)^\alpha$. Also, the moments condition in (1) is satisfied for some $p > 1$ (if f satisfies condition (C2) and $\alpha < d$) or $p > 2$ (if f satisfies condition (C1)), so Theorem 1 with $q = 1$ or $q = 2$, respectively, yields

$$\begin{aligned}
n^{\alpha/d-1}\mathcal{L}_j^{d,\alpha}(\mathcal{X}_n) &= n^{-1}\sum_{\boldsymbol{x}\in\mathcal{X}_n} n^{\alpha/d}\xi(\boldsymbol{x}; \mathcal{X}_n) \\
&= n^{-1}\sum_{\boldsymbol{x}\in\mathcal{X}_n} \xi(n^{1/d}\boldsymbol{x}; n^{1/d}\mathcal{X}_n) \\
&\xrightarrow{L^q} \mathrm{E}[\xi_\infty(\mathcal{H}_1)]\int_{\operatorname{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d}\,\mathrm{d}\boldsymbol{x}
\end{aligned} \tag{14}$$

(using the fact that ξ is homogeneous of order α). We now need to evaluate the expectation on the right-hand side of (14). For $r > 0$,

$$\begin{aligned}
\mathrm{P}[\xi_\infty(\mathcal{H}_1) > r] &= \mathrm{P}[d_j(\boldsymbol{0}; \mathcal{H}_1) > r^{1/\alpha}] \\
&= \sum_{i=0}^{j-1} \mathrm{P}[\operatorname{card}(B(\boldsymbol{0}; r^{1/\alpha}) \cap \mathcal{H}_1) = i] \\
&= \sum_{i=0}^{j-1} \frac{(v_d r^{d/\alpha})^i}{i!}\exp(-v_d r^{d/\alpha}),
\end{aligned}$$

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

where v_d is as given in (3). Thus,

$$\mathrm{E}[\xi_\infty(\mathcal{H}_1)] = \int_0^\infty \mathrm{P}[\xi_\infty(\mathcal{H}_1) > r]\,\mathrm{d}r = \int_0^\infty \sum_{i=0}^{j-1} \frac{(v_d r^{d/\alpha})^i}{i!} \exp(-v_d r^{d/\alpha})\,\mathrm{d}r.$$

Interchanging the order of summation and integration, and using (13), we obtain

$$\mathrm{E}[\xi_\infty(\mathcal{H}_1)] = v_d^{-\alpha/d} \frac{\alpha}{d} \sum_{i=0}^{j-1} \frac{\Gamma(i+\alpha/d)}{\Gamma(i+1)} = v_d^{-\alpha/d} \frac{\Gamma(j+\alpha/d)}{\Gamma(j)}, \tag{15}$$

where the final equality follows by induction on j. Then, from (3), (14), and (15), we obtain the jth-NNG$'$ result in (5). By (2), the k-NNG$'$ result in (6) follows from (5) with

$$C(d,\alpha,k) = v_d^{-\alpha/d} \sum_{j=1}^{k} \frac{\Gamma(j+\alpha/d)}{\Gamma(j)} = v_d^{-\alpha/d} \frac{d}{d+\alpha} \frac{\Gamma(k+1+\alpha/d)}{\Gamma(k)}.$$

Proof of Theorem 3. The nearest-neighbour (directed) graph counts the weights of edges from points that are nearest neighbours of their own nearest neighbours twice, while the nearest-neighbour (undirected) graph counts such weights only once.

Let $q(\boldsymbol{x};\mathcal{X})$ be the functional denoting the distance from $\boldsymbol{x}$ to its nearest neighbour in $\mathcal{X}\setminus\{\boldsymbol{x}\}$ if $\boldsymbol{x}$ is a nearest neighbour of its own nearest neighbour, and let it equal 0 otherwise. Recall that $d_1(\boldsymbol{x};\mathcal{X})$ is the distance from $\boldsymbol{x}$ to its nearest neighbour in $\mathcal{X}\setminus\{\boldsymbol{x}\}$. For $\alpha \geq 0$, define

$$\xi'(\boldsymbol{x};\mathcal{X}) := d_1(\boldsymbol{x};\mathcal{X})^\alpha - \tfrac{1}{2} q(\boldsymbol{x};\mathcal{X})^\alpha.$$

Then $\sum_{\boldsymbol{x}\in\mathcal{X}} \xi'(\boldsymbol{x};\mathcal{X})$ is the total weight of the nearest-neighbour (undirected) graph on $\mathcal{X}$. Note that ξ' is translation invariant and homogeneous of order α. We can check that ξ' is stabilizing on the Poisson process $\mathcal{H}_1$, using arguments similar to those for the jth-NNG$'$ and k-NNG$'$ functionals. Also (see [25]) if condition (C1) holds then ξ' satisfies the moments condition (1) for some $p > 2$ and for all $\alpha \geq 0$.

Let $\boldsymbol{e}_1$ be a vector of unit length in $\mathbb{R}^d$, and recall that $\omega_d := |B(\mathbf{0};1) \cup B(\boldsymbol{e}_1;1)|$ denotes the volume of the union of two unit d-balls with centres a unit distance apart.

Now we apply Theorem 1 with $q = 2$. We have

$$n^{\alpha/d-1}\mathcal{N}^{d,\alpha}(\mathcal{X}_n) = n^{-1} \sum_{\boldsymbol{x}\in\mathcal{X}_n} \xi'(n^{1/d}\boldsymbol{x}; n^{1/d}\mathcal{X}_n) \xrightarrow{L^2} \mathrm{E}[\xi'_\infty(\mathcal{H}_1)] \int_{\mathrm{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d}\,\mathrm{d}\boldsymbol{x}, \tag{16}$$

where $\mathrm{E}[\xi'_\infty(\mathcal{H}_1)] = \mathrm{E}[d_1(\mathbf{0};\mathcal{H}_1)^\alpha] - \frac{1}{2}\mathrm{E}[q(\mathbf{0};\mathcal{H}_1)^\alpha]$. Now we need to evaluate $\mathrm{E}[q(\mathbf{0};\mathcal{H}_1)^\alpha]$. With $\boldsymbol{X}$ denoting the nearest point of $\mathcal{H}_1$ to $\mathbf{0}$,

$$\begin{aligned}\mathrm{P}[q(\mathbf{0};\mathcal{H}_1) \in \mathrm{d}r] &= \mathrm{P}[\{\|\boldsymbol{X}\| \in \mathrm{d}r\} \cap \{\mathcal{H}_1 \cap (B(\mathbf{0};r) \cup B(\boldsymbol{X};r)) = \{\boldsymbol{X}\}\}] \\ &= d v_d r^{d-1} \exp(-v_d r^d) \exp(-(\omega_d - v_d) r^d)\,\mathrm{d}r \\ &= d v_d r^{d-1} \exp(-\omega_d r^d)\,\mathrm{d}r.\end{aligned}$$

Using (13) we thus obtain

$$\mathrm{E}[q(\mathbf{0};\mathcal{H}_1)^\alpha] = \int_0^\infty d v_d r^{d-1+\alpha} \exp(-\omega_d r^d)\,\mathrm{d}r = v_d \omega_d^{-1-\alpha/d} \Gamma\left(1+\frac{\alpha}{d}\right). \tag{17}$$

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

Then, from (16), using (17) and the $j = 1$ case of (15), we obtain (8). From some calculus we find that $\omega_2 = 4\pi/3 + \sqrt{3}/2$, which with the $d = 2$ case of (8) yields (9); in verifying (10) note that $\Gamma(\frac{3}{2}) = \pi^{1/2}/2$ (see Equation 6.1.9 of [1]). Finally, the statement for $\mathcal{N}^{1,1}(\mathcal{U}_n)$ follows from the $d = 1$ case of (8) since $\omega_1 = 3$.

3.2. Proof of Theorem 4

In order to obtain our LLN (Theorem 4), we modify the setup of the ONG slightly. Let $\mathcal{U}_n$ be a *marked* random finite point process in $\mathbb{R}^d$ consisting of n independent, uniform random vectors in $(0, 1)^d$, where each point $\boldsymbol{U}_i$ of $\mathcal{U}_n$ carries a random mark $T(\boldsymbol{U}_i)$ which is uniformly distributed on [0, 1], independent of the other marks and of the point process $\mathcal{U}_n$. We join each point $\boldsymbol{U}_i$ of $\mathcal{U}_n$ to its nearest neighbour amongst those points of $\mathcal{U}_n$ with mark less than $T(\boldsymbol{U}_i)$, if there are any such points, to obtain a graph that we call the *ONG on the marked point set* $\mathcal{U}_n$. This definition extends to infinite but locally finite point sets. Clearly the ONG on the marked point process $\mathcal{U}_n$ has the same distribution as the ONG (with the first definition) on a sequence $\boldsymbol{U}_1, \boldsymbol{U}_2, \dots, \boldsymbol{U}_n$ of independent uniform points on $(0, 1)^d$.

We apply Theorem 1 to obtain an LLN for $\mathcal{O}^{d,\alpha}(\mathcal{U}_n)$, $\alpha \in [0, d)$. Once again, the method enables us to evaluate the limit explicitly. We take f to be the indicator function of $(0, 1)^d$. Define $D(\boldsymbol{x}; \mathcal{X})$ to be the distance from point $\boldsymbol{x}$, with mark $T(\boldsymbol{x})$, to its nearest neighbour in $\mathcal{X}$ amongst those points $\boldsymbol{y} \in \mathcal{X}$ that have mark $T(\boldsymbol{y})$ such that $T(\boldsymbol{y}) < T(\boldsymbol{x})$, if such a neighbour exists, and let it equal 0 otherwise. We take $\xi(\boldsymbol{x}; \mathcal{X})$ to be $D(\boldsymbol{x}; \mathcal{X})^\alpha$. Again, ξ is translation invariant and homogeneous of order α.

Lemma 1. *The ONG functional ξ almost surely stabilizes on $\mathcal{H}_1$.*

Proof. Although the notion of stabilization there is somewhat different, the same argument as given at the start of the proof of Theorem 3.6 of [20] applies.

Lemma 2. *Let $d \in \mathbb{N}$ and $\alpha \in [0, d)$, and let $p > 1$ with $\alpha p < d$. Then the ONG functional ξ satisfies the moments condition (1).*

Proof. Let T_n denote the rank of the mark of $\boldsymbol{U}_1$ amongst the marks of all the points of $\mathcal{U}_n$; thus, T_n is distributed uniformly over the integers $1, 2, \dots, n$. We have, by conditioning on T_n,

$$\begin{aligned} \mathrm{E}[\xi(n^{1/d}\boldsymbol{U}_1; n^{1/d}\mathcal{U}_n)^p] &= n^{-1}\sum_{i=1}^{n} \mathrm{E}[d_1(n^{1/d}\boldsymbol{U}_1; n^{1/d}\mathcal{U}_i)^{p\alpha}] \\ &= n^{-1}\sum_{i=1}^{n} \left(\frac{n}{i}\right)^{p\alpha/d} \mathrm{E}[d_1(i^{1/d}\boldsymbol{U}_1; i^{1/d}\mathcal{U}_i)^{p\alpha}]. \end{aligned} \tag{18}$$

It was shown in [23] that there exists a $C \in (0, \infty)$ such that, for all $r > 0$,

$$\sup_{i \ge 1} \mathrm{P}[d_1(i^{1/d}\boldsymbol{U}_1; i^{1/d}\mathcal{U}_i) > r] \le C \exp(-r^{1/d}/C).$$

Thus, the last expectation in (18) is bounded by a constant independent of i, meaning that the final expression in (18) is bounded by a constant times

$$n^{(p\alpha-d)/d}\sum_{i=1}^{n} i^{-p\alpha/d},$$

which is uniformly bounded by a constant for $\alpha p < d$.

Proof of Theorem 4. Let $d \in \mathbb{N}$. Let f be the indicator function of $(0, 1)^d$ and let ξ be the ONG functional $\xi(\boldsymbol{x}; \mathcal{U}_n) = D(\boldsymbol{x}; \mathcal{U}_n)^\alpha$. By Lemmas 1 and 2, ξ is homogeneous of order α, stabilizing on $\mathcal{H}_1$ with limit $\xi_\infty(\mathcal{H}_1) = D(\mathbf{0}; \mathcal{H}_1)^\alpha$, and satisfies the moments condition (1) for some $p > 1$, provided that $\alpha < d$. Theorem 1 with $q = 1$ thus implies that

$$n^{\alpha/d-1}\mathcal{O}^{d,\alpha}(\mathcal{U}_n) = n^{-1}\sum_{\boldsymbol{x}\in\mathcal{U}_n} D(n^{1/d}\boldsymbol{x}; n^{1/d}\mathcal{U}_n)^\alpha \xrightarrow{L^1} \mathrm{E}[\xi_\infty(\mathcal{H}_1)].$$

For $u \in (0, 1)$, the points of $\mathcal{H}_1$ with marks lower than u form a homogeneous Poisson point process of intensity u, so by conditioning on the mark of the point at $\mathbf{0}$ we obtain

$$\mathrm{E}[\xi_\infty(\mathcal{H}_1)] = \int_0^1 \mathrm{E}[d_1(\mathbf{0}; \mathcal{H}_u)^\alpha]\,\mathrm{d}u = \int_0^1 u^{-\alpha/d}\,\mathrm{E}[d_1(\mathbf{0}; \mathcal{H}_1)^\alpha]\,\mathrm{d}u = \frac{d}{d-\alpha}C(d, \alpha, 1),$$

since, as we saw in the proof of Theorem 2, $\mathrm{E}[d_1(\mathbf{0}; \mathcal{H}_1)^\alpha] = C(d, \alpha, 1)$.

3.3. Proof of Theorem 5

In applying Theorem 1 to the MDSF, we take f to be the indicator function of $(0, 1)^2$. We take $\xi(\mathbf{x}; \mathcal{X})$ to be $d(\boldsymbol{x}; \mathcal{X})^\alpha$, where $d(\boldsymbol{x}; \mathcal{X})$ is the distance from point $\mathbf{x}$ to its directed nearest neighbour in $\mathcal{X} \setminus \{\boldsymbol{x}\}$, if such a point exists, and equals 0 otherwise, i.e.

$$\xi(\mathbf{x}; \mathcal{X}) = d(\mathbf{x}; \mathcal{X})^\alpha \quad \text{with} \quad d(\boldsymbol{x}; \mathcal{X}) := \min\{\|\boldsymbol{x} - \boldsymbol{y}\| : \boldsymbol{y} \in \mathcal{X} \setminus \{\boldsymbol{x}\},\ \boldsymbol{y} \overset{\theta,\phi}{\preccurlyeq} \boldsymbol{x}\} \tag{19}$$

with the convention that $\min \varnothing = 0$.

We consider the random point set $\mathcal{U}_n$, the binomial point process consisting of n independent, uniformly distributed points on $(0, 1)^2$. However, as remarked before the statement of Theorem 5, the result in (12) carries through (with virtually the same proof) to more general point sets $\mathcal{X}_n$.

We need to show that ξ as given in (19) satisfies the conditions of Theorem 1. As before, $\mathcal{H}_1$ denotes a homogeneous Poisson process on $\mathbb{R}^2$.

Lemma 3. *The MDSF ξ given in (19) almost surely stabilizes on $\mathcal{H}_1$, with limit $\xi_\infty(\mathcal{H}_1) = d(\mathbf{0}; \mathcal{H}_1)^\alpha$.*

Proof. Set $R := d(\mathbf{0}; \mathcal{H}_1)$. Since $\phi > 0$, we have $0 < R < \infty$ almost surely. However, for any $\ell > R$ we then have $\xi(\mathbf{0}; (\mathcal{H}_1 \cap B(\mathbf{0}; \ell)) \cup \mathcal{A}) = R^\alpha$ for any finite $\mathcal{A} \subset \mathbb{R}^d \setminus B(\mathbf{0}; \ell)$. Thus, ξ stabilizes on $\mathcal{H}_1$ with limit $\xi_\infty(\mathcal{H}_1) = R^\alpha$.

We now give a geometrical lemma. Let ∂B denote the boundary of $B \subset \mathbb{R}^d$. For $B \subset \mathbb{R}^2$ with B bounded, and for $\boldsymbol{x} \in B$, we write $\mathrm{dist}(\boldsymbol{x}; \partial B)$ for $\sup\{r : B(\boldsymbol{x}; r) \subseteq B\}$, and for $s > 0$ define the region

$$A_{\theta,\phi}(\boldsymbol{x}, s; B) := B(\boldsymbol{x}; s) \cap B \cap C_{\theta,\phi}(\boldsymbol{x}). \tag{20}$$

Lemma 4. *Let B be a convex, bounded set in $\mathbb{R}^2$, and let $\boldsymbol{x} \in B$. If $A_{\theta,\phi}(\boldsymbol{x}, s; B) \cap \partial B(\boldsymbol{x}; s) \neq \varnothing$ and $s > \mathrm{dist}(\boldsymbol{x}; \partial B)$, then*

$$|A_{\theta,\phi}(\boldsymbol{x}, s; B)| \geq s \sin\left(\frac{\phi}{2}\right)\frac{\mathrm{dist}(\boldsymbol{x}; \partial B)}{2}.$$

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

Proof. The condition $A_{\theta,\phi}(\boldsymbol{x}, s; B) \cap \partial B(\boldsymbol{x}; s) \neq \varnothing$ says that there exists a $\boldsymbol{y} \in B \cap C_{\theta,\phi}(\boldsymbol{x})$ with $\|\boldsymbol{y} - \boldsymbol{x}\| = s$. The line segment $\boldsymbol{x}\boldsymbol{y}$ is contained in the cone $C_{\theta,\phi}(\boldsymbol{x})$. Take a half-line $\boldsymbol{h}$ starting from $\boldsymbol{x}$, at an angle $\phi/2$ to the line segment $\boldsymbol{x}\boldsymbol{y}$ and such that $\boldsymbol{h}$ is also contained in $C_{\theta,\phi}(\boldsymbol{x})$, and let $\boldsymbol{z}$ be the point on $\boldsymbol{h}$ at a distance $\operatorname{dist}(\boldsymbol{x}; \partial B)$ from $\boldsymbol{x}$. Then the interior of the triangle $\boldsymbol{x}\boldsymbol{y}\boldsymbol{z}$ is entirely contained in $A_{\theta,\phi}(\boldsymbol{x}, s; B)$, and has area $s \sin(\phi/2) \operatorname{dist}(\boldsymbol{x}; \partial B)/2$.

Lemma 5. *Suppose that $\alpha > 0$. Then the MDSF ξ given in (19) satisfies the moments condition (1) for any $p \leq 2/\alpha$.*

Proof. With $R_n := (0, n^{1/2})^2$, and conditioning on the position of $\boldsymbol{U}_1$, we have

$$\mathrm{E}[\xi(n^{1/2}\boldsymbol{U}_1; n^{1/2}\mathcal{U}_n)^p] = n^{-1} \int_{R_n} \mathrm{E}[\xi(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1})^p] \, \mathrm{d}\boldsymbol{x}. \tag{21}$$

For $\boldsymbol{x} \in R_n$, let $m(\boldsymbol{x}) := \operatorname{dist}(\boldsymbol{x}, \partial R_n)$. We divide R_n into three regions:

$$\begin{aligned} R_n(1) &:= \{\boldsymbol{x} \in R_n : m(\boldsymbol{x}) \leq n^{-1/2}\}, \\ R_n(2) &:= \{\boldsymbol{x} \in R_n : m(\boldsymbol{x}) > 1\}, \\ R_n(3) &:= \{\boldsymbol{x} \in R_n : n^{-1/2} < m(\boldsymbol{x}) \leq 1\}. \end{aligned}$$

For all $\boldsymbol{x} \in R_n$, we have $\xi(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1}) \leq (2n)^{\alpha/2}$ and, hence, since $R_n(1)$ has area at most 4, we can bound the contribution to (21) from $\boldsymbol{x} \in R_n(1)$ by

$$n^{-1} \int_{R_n(1)} \mathrm{E}[\xi(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1})^p] \, \mathrm{d}\boldsymbol{x} \leq 4n^{-1}(2n)^{p\alpha/2} = 2^{2+p\alpha/2} n^{(p\alpha-2)/2}, \tag{22}$$

which is bounded if $p\alpha \leq 2$. Now, for $\boldsymbol{x} \in R_n$, with $A_{\theta,\phi}$ as defined in (20), we have

$$\begin{aligned} \mathrm{P}[d(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1}) > s] &\leq \mathrm{P}[n^{1/2}\mathcal{U}_{n-1} \cap A_{\theta,\phi}(\boldsymbol{x}, s; R_n) = \varnothing] \\ &= \left(1 - \frac{|A_{\theta,\phi}(\boldsymbol{x}, s; R_n)|}{n}\right)^{n-1} \\ &\leq \exp(1 - |A_{\theta,\phi}(\boldsymbol{x}, s; R_n)|), \end{aligned} \tag{23}$$

since $|A_{\theta,\phi}(\boldsymbol{x}, s; R_n)| \leq n$. For $\boldsymbol{x} \in R_n$ and $s > m(\boldsymbol{x})$, by Lemma 4 we have

$$|A_{\theta,\phi}(\boldsymbol{x}, s; R_n)| \geq s \sin\left(\frac{\phi}{2}\right) \frac{m(\boldsymbol{x})}{2} \quad \text{if} \quad A_{\theta,\phi}(\boldsymbol{x}, s; R_n) \cap \partial B(\boldsymbol{x}; s) \neq \varnothing,$$

and also

$$\mathrm{P}[d(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1}) > s] = 0 \quad \text{if} \quad A_{\theta,\phi}(\boldsymbol{x}, s; R_n) \cap \partial B(\boldsymbol{x}; s) = \varnothing.$$

For $s \leq m(\boldsymbol{x})$, we have $|A_{\theta,\phi}(\boldsymbol{x}, s; R_n)| = s^2\phi/2 \geq s^2 \sin(\phi/2)$. Combining these observations with (23), for all $\boldsymbol{x} \in R_n$ and $s > 0$ we obtain

$$\mathrm{P}[d(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1}) > s] \leq \exp\left(1 - \frac{s}{2} \min(s, m(\boldsymbol{x})) \sin\left(\frac{\phi}{2}\right)\right), \quad \boldsymbol{x} \in R_n.$$

With $c = \frac{1}{2}\sin(\phi/2)$, for $\boldsymbol{x} \in R_n$ we therefore have

$$\begin{aligned}
\mathrm{E}[\xi(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1})^p] &= \int_0^\infty \mathrm{P}[d(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1}) > r^{1/(\alpha p)}]\,\mathrm{d}r \\
&\le \int_0^{m(\boldsymbol{x})^{\alpha p}} \exp(1 - cr^{2/(\alpha p)})\,\mathrm{d}r + \int_{m(\boldsymbol{x})^{\alpha p}}^\infty \exp(1 - cm(\boldsymbol{x})r^{1/(\alpha p)})\,\mathrm{d}r \\
&= O(1) + \alpha p m(\boldsymbol{x})^{-p\alpha} \int_{m(\boldsymbol{x})^2}^\infty \mathrm{e}^{1-cu} \alpha p u^{\alpha p - 1}\,\mathrm{d}u \\
&= O(1) + O(m(\boldsymbol{x})^{-\alpha p}).
\end{aligned} \tag{24}$$

For $\boldsymbol{x} \in R_n(2)$ this bound is $O(1)$, and the area of $R_n(2)$ is less than n, so the contribution to (21) from $R_n(2)$ satisfies

$$\limsup_{n\to\infty} n^{-1} \int_{R_n(2)} \mathrm{E}[\xi(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1})^p]\,\mathrm{d}\boldsymbol{x} < \infty. \tag{25}$$

Finally, by (24), there exists a constant $C \in (0, \infty)$ such that the contribution to (21) from $R_n(3)$ satisfies

$$\begin{aligned}
n^{-1} \int_{R_n(3)} \mathrm{E}[\xi(\boldsymbol{x}; n^{1/2}\mathcal{U}_{n-1})^p]\,\mathrm{d}\boldsymbol{x} &\le Cn^{-1/2} \int_{n^{-1/2}}^1 y^{-\alpha p}\,\mathrm{d}y \\
&\le Cn^{-1/2} \max(\log n, n^{(\alpha p - 1)/2}),
\end{aligned}$$

which is bounded if $\alpha p \le 2$. Combined with the bounds in (22) and (25), this shows that the expression in (21) is uniformly bounded, provided that $\alpha p \le 2$.

For $k \in \mathbb{N}$, $a < b$, and $c < d$, let $\mathcal{U}_{k,(a,b]\times(c,d]}$ denote the point process consisting of k independent, random vectors uniformly distributed on the rectangle $(a, b] \times (c, d]$. Before proceeding further, we recall that if $M(\mathcal{X})$ denotes the number of minimal elements, under the ordering '$\preccurlyeq^*$', of a point set $\mathcal{X} \subset \mathbb{R}^2$, then

$$\mathrm{E}[M(\mathcal{U}_{k,(a,b]\times(c,d]})] = \mathrm{E}[M(\mathcal{U}_k)] = 1 + \frac{1}{2} + \cdots + \frac{1}{k} \le 1 + \log k. \tag{26}$$

The first equality in (26) comes from some obvious scaling which shows that the distribution of $M(\mathcal{U}_{k,(a,b]\times(c,d]})$ does not depend on a, b, c, or d. For a proof of the second equality in (26), see, e.g. [3].

Proof of Theorem 5. Suppose that $\alpha \in (0, 2)$ and let f be the indicator function of $(0, 1)^2$. By Lemmas 3 and 5, the functional ξ, given in (19), satisfies the conditions of Theorem 1 with $p = 2/\alpha$ and $q = 1$. Thus, by Theorem 1 we have

$$n^{\alpha/2-1}\mathcal{M}^\alpha(\mathcal{U}_n) = n^{-1} \sum_{\boldsymbol{x}\in\mathcal{U}_n} \xi(n^{1/2}\boldsymbol{x}; n^{1/2}\mathcal{U}_n) \xrightarrow{L^1} \mathrm{E}[\xi_\infty(\mathcal{H}_1)]. \tag{27}$$

Since the disk sector $C_{\theta,\phi}(\boldsymbol{x}) \cap B(\boldsymbol{x}; r)$ has area $(\phi/2)r^2$, by Lemma 3 we have

$$\mathrm{P}[\xi_\infty(\mathcal{H}_1) > s] = \mathrm{P}[\mathcal{H}_1 \cap C_{\theta,\phi}(\mathbf{0}) \cap B(\mathbf{0}; s^{1/\alpha}) = \varnothing] = \exp\biggl(-\frac{\phi}{2}s^{2/\alpha}\biggr).$$

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

Hence, using (13), the limit in (27) is

$$\mathrm{E}[\xi_\infty(\mathcal{H}_1)] = \int_0^\infty \mathrm{P}[\xi_\infty(\mathcal{H}_1) > s]\,\mathrm{d}s = \alpha 2^{(\alpha-2)/2}\phi^{-\alpha/2}\Gamma\left(\frac{\alpha}{2}\right),$$

and this gives us (12). Finally, in the case where '$\overset{\theta,\phi}{\preccurlyeq}$' is '$\preccurlyeq^*$', (12) remains true if $\mathcal{U}_n$ is replaced by $\mathcal{U}_n^0$, since

$$\mathrm{E}[n^{\alpha/2-1}|\mathcal{M}^\alpha(\mathcal{U}_n^0) - \mathcal{M}^\alpha(\mathcal{U}_n)|] \le 2^{\alpha/2}n^{\alpha/2-1}\,\mathrm{E}[M(\mathcal{U}_n)]. \tag{28}$$

By (26), $\mathrm{E}[M(\mathcal{U}_n)] \le 1 + \log n$ and, hence, the right-hand side of (28) tends to 0 as n tends to ∞, for $\alpha < 2$.

3.4. Proof of Theorem 6

In applying Theorem 1 to the Gabriel graph, we take $\xi(\boldsymbol{x}; \mathcal{X}_n)$ to be *one-half* of the total power-α-weighted length of all the edges incident at $\boldsymbol{x}$ in the Gabriel graph on $\mathcal{X}_n \cup \{\boldsymbol{x}\}$; the factor of one-half accounts for double counting. As stated in Section 2.3(e) of [25], ξ is translation invariant, homogeneous of order α, and stabilizing on $\mathcal{H}_1$, and if the function f satisfies condition (C1) then the moments condition (1) is satisfied for some $p > 2$. Thus, by Theorem 1 with $q = 2$,

$$n^{\alpha/d-1}\mathcal{G}^{d,\alpha}(\mathcal{X}_n) = n^{-1}\sum_{\boldsymbol{x}\in\mathcal{X}_n}\xi(n^{1/d}\boldsymbol{x}; n^{1/d}\mathcal{X}_n) \xrightarrow{L^2} \mathrm{E}[\xi_\infty(\mathcal{H}_1)]\int_{\mathrm{supp}(f)} f(\boldsymbol{x})^{(d-\alpha)/d}\,\mathrm{d}\boldsymbol{x}. \tag{29}$$

We need to evaluate the expectation on the right-hand side of (29). The net contribution from a vertex at $\mathbf{0}$ to the total weight of the Gabriel graph on $\mathcal{H}_1$ is

$$\frac{1}{2}\sum_{k=1}^\infty d_k(\mathbf{0}; \mathcal{H}_1)^\alpha\, 1_{E_k}, \tag{30}$$

where the factor of one-half accounts for the fact that edges are not counted twice, $d_k(\mathbf{0}; \mathcal{H}_1)$ is the distance from $\mathbf{0}$ to its kth-nearest neighbour in $\mathcal{H}_1$, E_k denotes the event that $\mathbf{0}$ and its kth-nearest neighbour in $\mathcal{H}_1$ are joined by an edge in the Gabriel graph, and 1_{E_k} denotes the indicator function of this event.

Given that the point $\boldsymbol{x} \in \mathcal{H}_1$ is the kth-nearest neighbour of $\mathbf{0}$, an edge between $\boldsymbol{x}$ and $\mathbf{0}$ exists in the Gabriel graph if and only if the ball upon which $\mathbf{0}$ and $\boldsymbol{x}$ are diametrically opposed contains none of the other $k-1$ points of $\mathcal{H}_1$ that are uniformly distributed in the ball of centre $\mathbf{0}$ and radius $\|\boldsymbol{x}\|$. Thus, for $k \in \mathbb{N}$,

$$\mathrm{P}[E_k] = \left(\frac{v_d r^d - v_d(r/2)^d}{v_d r^d}\right)^{k-1} = (1-2^{-d})^{k-1}, \tag{31}$$

and from (30) and (31) we have

$$\begin{aligned}\mathrm{E}[\xi_\infty(\mathcal{H}_1)] &= \frac{1}{2}\sum_{k=1}^\infty (1-2^{-d})^{k-1}\,\mathrm{E}[d_k(\mathbf{0}; \mathcal{H}_1)^\alpha] \\ &= \frac{1}{2}\sum_{k=1}^\infty (1-2^{-d})^{k-1} v_d^{-\alpha/d}\frac{\Gamma(k+\alpha/d)}{\Gamma(k)},\end{aligned}$$

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press

by (15). However, owing to the properties of the Gauss hypergeometric series (see Equations 15.1.1 and 15.1.8 of [1]),

$$\sum_{k=1}^{\infty}(1-2^{-d})^{k-1}\frac{\Gamma(k+\alpha/d)}{\Gamma(k)}=\Gamma\left(1+\frac{\alpha}{d}\right)2^{d+\alpha}.$$

Substituting the resulting expression for $\mathrm{E}[\xi_\infty(\mathcal{H}_1)]$ into (29) completes the proof.

Acknowledgements

Some of this work was done when the author was at the University of Durham, supported by an EPSRC doctoral training account, and at the University of Bath. The author thanks Mathew Penrose for very helpful discussions and comments, and two anonymous referees, whose comments have led to an improved presentation.

References

[1] Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (National Bureau Standards Appl. Math. Ser. **55**). U.S. Government Printing Office, Washington, DC.

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

[3] Barndorff-Nielsen, O. and Sobel, M. (1966). On the distribution of the number of admissible points in a vector random sample. *Theory Prob. Appl.* **11,** 249–269.

[4] Berger, N. *et al.* (2003). Degree distribution of the FKP model. In *Automata, Languages and Programming* (*Lecture Notes Comput. Sci.* **2719**), eds J. C. M. Baeten, J. K. Lenstra, J. Parrow and G. J. Woeginger, Springer, Berlin, pp. 725–738.

[5] Bhatt, A. G. and Roy, R. (2004). On a random directed spanning tree. *Adv. Appl. Prob.* **36,** 19–42.

[6] Bickel, P. J. and Breiman, L. (1983). Sums of functions of nearest neighbour distances, moment bounds, and a goodness of fit test. *Ann. Prob.* **11,** 185–214.

[7] Brito, M. R., Quirox, A. J. and Yukich, J. E. (2002). Graph-theoretic procedures for dimension identification. *J. Multivariate Anal.* **81,** 67–84.

[8] Evans, D. and Jones, A. J. (2002). A proof of the Gamma test. *Proc. R. Soc. London A* **458,** 2759–2799.

[9] Evans, D., Jones, A. J. and Schmidt, W. M. (2002). Asymptotic moments of near-neighbour distance distributions. *Proc. R. Soc. London A* **458,** 2839–2849.

[10] Friedman, J. H. and Rafsky, L. C. (1983). Graph-theoretic measures of multivariate association and prediction. *Ann. Statist.* **2,** 377–391.

[11] Gabriel, K. R. and Sokal, R. R. (1969). A new statistical approach to geographic variation analysis. *Systematic Zoology* **18,** 259–278.

[12] Henze, N. (1987). On the fraction of random points with specified nearest-neighbour interrelations and degree of attraction. *Adv. Appl. Prob.* **19,** 873–895.

[13] Henze, N. and Voigt, B. (1992). Almost sure convergence of certain slowly changing symmetric one- and multi-sample statistics. *Ann. Prob.* **20,** 1086–1098.

[14] Huang, K. (1987). *Statistical Mechanics*, 2nd edn. John Wiley, New York.

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

[16] Kolars, J. F. and Nystuen, J. D. (1974). *Human Geography: Spatial Design in World Society.* McGraw-Hill, New York.

[17] McGivney, K. (1997). Probabilistic limit theorems for combinatorial optimization problems. Doctoral Thesis, Lehigh University.

[18] Miles, R. E. (1970). On the homogeneous planar Poisson point process. *Math. Biosci.* **6,** 85–127.

[19] Penrose, M. (2003). *Random Geometric Graphs* (Oxford Studies Prob. **6**). Oxford University Press.

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

[21] Penrose, M. D. and Wade, A. R. (2004). Random minimal directed spanning trees and Dickman-type distributions. *Adv. Appl. Prob.* **36,** 691–714.

[22] Penrose, M. D. and Wade, A. R. (2006). On the total length of the random minimal directed spanning tree. *Adv. Appl. Prob.* **38,** 336–372.
[23] Penrose, M. D. and Wade, A. R. (2007). Limit theory for the random on-line nearest-neighbour graph. To appear in *Random Structures Algorithms.*
[24] Penrose, M. D. and Yukich, J. E. (2001). Central limit theorems for some graphs in computational geometry. *Ann. Appl. Prob.* **11,** 1005–1041.
[25] Penrose, M. D. and Yukich, J. E. (2003). Weak laws of large numbers in geometric probability. *Ann. Appl. Prob.* **13,** 277–303.
[26] Percus, A. G. and Martin, O. C. (1998). Scaling universalities of kth nearest neighbor distances on closed manifolds. *Adv. Appl. Math.* **21**, 424–436.
[27] Pielou, E. C. (1977). *Mathematical Ecology*, 2nd edn. John Wiley, New York.
[28] Smith, W. D. (1989). Studies in computational geometry motivated by mesh generation. Doctoral Thesis, Princeton University.
[29] Steele, J. M. (1997). *Probability Theory and Combinatorial Optimization*. Society for Industrial and Applied Mathematics, Philadelphia, PA.
[30] Toussaint, G. (2005). Geometric proximity graphs for improving nearest neighbour methods in instance-based learning and data mining. *Internat. J. Comput. Geometry Appl.* **15,** 101–150.
[31] Yukich, J. E. (1998). *Probability Theory of Classical Euclidean Optimization Problems* (Springer Lecture Notes Math. **1675**). Springer, Berlin.

https://doi.org/10.1239/aap/1183667613 Published online by Cambridge University Press