The Annals of Applied Probability
2001, Vol. 11, No. 4, 1005–1041

CENTRAL LIMIT THEOREMS FOR SOME GRAPHS IN COMPUTATIONAL GEOMETRY

BY MATHEW D. PENROSE AND J. E. YUKICH[1]

University of Durham and Lehigh University

Let (B_n) be an increasing sequence of regions in d-dimensional space with volume n and with union $\mathbb{R}^d$. We prove a general central limit theorem for functionals of point sets, obtained either by restricting a homogeneous Poisson process to B_n, or by by taking n uniformly distributed points in B_n. The sets B_n could be all cubes but a more general class of regions B_n is considered. Using this general result we obtain central limit theorems for specific functionals such as total edge length and number of components, defined in terms of graphs such as the k-nearest neighbors graph, the sphere of influence graph and the Voronoi graph.

1. Introduction. The purpose of this paper is to develop a general methodology to establish central limit theorems (CLTs) for functionals of graphs in computational geometry. Functionals of interest include total edge length, total number of edges, total number of components and total number of vertices of fixed degree. Graphs of interest include the k-nearest neighbors graph, the Voronoi and Delaunay tessellations, the sphere of influence graph, the Gabriel graph and the relative neighbor graph. These graphs are formally defined later on. In each case, the graph or its dual graph (as with the Voronoi graph) is constructed as follows: given a finite vertex set in $\mathbb{R}^d$, $d \geq 1$, undirected edges are drawn from each vertex to various nearby vertices, the choice of edges to include being determined by the local point configuration according to some specified rule. Sometimes such graphs are called *proximity graphs*; see [3] for a precise definition.

Our graphs are random in the sense that the vertex set is a random point set in $\mathbb{R}^d$, $d \geq 1$. We establish CLTs for two related types of random point sets: the homogeneous Poisson point process on a large region or "window" of $\mathbb{R}^d$ and the point set consisting of a large independent sample of nonrandom sample size from the uniform distribution on such a region. By scaling, these often yield a CLT for Poisson processes of high intensity on a fixed set such as the unit cube $[0, 1]^d$, or for large independent samples of nonrandom size from the uniform distribution on a fixed set. As a by-product, we also prove the convergence of the (scaled) variance of our functionals of interest.

One of our more interesting new results is a CLT for the total number of components of the k-nearest neighbors graph, either on a Poisson process or on a sample of nonrandom size, and likewise for the sphere of influence graph.

Received November 1999; revised June 2000.

[1]Supported in part by NSA Grant MDA904-97-1-0053.

AMS 2000 *subject classifications*. Primary 60F05; secondary 60D05.

Key words and phrases. Central limit theorems, computational geometry, k-nearest neighbors graph, sphere of influence graph, Voronoi graph.

We also establish CLTs for the total edge length of the k-nearest neighbors graph on a sample of nonrandom size, and likewise for the sphere of influence graph and the Voronoi graph. These latter results add to existing results for their Poisson counterparts [1, 8, 10]. We believe that the CLTs established here, particularly those for nonrandom sample sizes, may have uses in the statistical analysis of data and may lead to useful tests for clustering.

All of our CLTs follow from a general CLT which can be viewed as an attempt to capture the essence of the martingale method developed by Kesten and Lee [11], and extended by Lee [12, 13], in their deep study of the random Euclidean minimal spanning tree. The martingale method is developed into a general CLT for functionals of lattice-indexed white noise in [19]. Using the method of [19], we find a general CLT for functionals of graphs over Poisson point sets. To obtain CLTs on nonrandom sample sizes, we de-Poissonize by using a coupling lemma and some key ideas of Kesten and Lee [11].

An important earlier paper developing general CLTs of this type is that of Avram and Bertsimas [1], and a brief comparison is in order. First, the martingale method used here achieves CLTs for some functionals for which it is not apparent how to use the dependency graph method of [1]. Second, our method yields convergence of variances, an issue not addressed in [1]. Third, nonrandom sample sizes are considered here, but not in [1]. Fourth, we prove CLTs for functionals of point sets defined on rather general regions, not just on cubes of volume n. On the other hand, the method of [1], where applicable, can yield error bounds providing useful information about the rate of convergence, which is not given by our method.

Specific graphs of interest are defined in terms of distances between points. We use the Euclidean norm, denoted $|\cdot|$, throughout, but our results should carry through to other norms. Let us now define the graphs of main interest.

k-nearest neighbors graph. The k-nearest neighbors graph on a point set $\mathscr{X} \subset \mathbb{R}^d$ is 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.

If the kth nearest neighbor of x is not well defined (i.e., if there is a "tie" in the ordering of interpoint distances involving x), use the lexicographic ordering as a "tie-breaker" to determine the k nearest neighbors. Such a tie has zero probability for the random point sets under consideration here. The k-nearest neighbors graph is an example of a dependent random connection model in percolation theory (see [16, 7]). This graph is also used in clustering methods in statistics and computer science. See [8, 21, 23] for additional applications and references.

Sphere of influence graph. Given a point set $\mathscr{X} \subset \mathbb{R}^d$, the sphere of influence graph (SIG) is constructed as follows: for each $x \in \mathscr{X}$, let $\mathscr{S}_x$ denote the closed ball centered at x with radius equal to the distance between x and its nearest neighbor in $\mathscr{X}$. This ball is often called the sphere of influence of x. The sphere of influence graph puts an edge between x and y if and only if the balls $\mathscr{S}_x$ and $\mathscr{S}_y$ overlap.

In the language of continuum percolation [16], much as the k-nearest neighbors graph can be viewed as a dependent random connection model, the sphere of influence graph can be viewed as a dependent Boolean model. The sphere of influence graph is used in pattern recognition and computer science and we refer to the survey [17] for details. One of our new results is "uniqueness of the infinite component" for the SIG on a homogeneous Poisson process on $\mathbb{R}^d$, a result which is required for one of our CLTs. This adds to known uniqueness results for other graphs [16, 9].

Voronoi tessellations. Given a point set $\mathcal{X} \subset \mathbb{R}^d$ and $x \in \mathcal{X}$, consider the locus of points closer to x than to any other point. This set of points is the intersection of half planes and is a convex polyhedral cell. The cells partition $\mathbb{R}^d$ into a convex net which is variously called the Voronoi tessellation, Voronoi graph, Voronoi diagram or Dirichlet tessellation of $\mathbb{R}^d$. Voronoi tessellations have numerous applications and are used to model natural phenomena in astrophysics, cell biology, crystallography, geology, metallography and other applied fields. See the encyclopedic work, [18], for details and a thorough treatment of the many applications.

This paper is organized as follows. Sections 2–5 contain the general results and their proofs, and Sections 6–9 contain applications to particular functionals of particular graphs.

Notational conventions: c denotes a generic finite positive constant whose value may change from line to line. For any set $\mathcal{X} \subset \mathbb{R}^d$ and any $y \in \mathbb{R}^d$, we denote by $\mathcal{X} - y$ the translated set $\{x - y\colon x \in \mathcal{X}\}$, and likewise set $\mathcal{X} + y = \{x + y\colon x \in \mathcal{X}\}$. Also, if $a > 0$, we let $a\mathcal{X}$ denote the set $\{ax\colon x \in \mathcal{X}\}$. For $x \in \mathbb{R}^d$ and $r > 0$, let $B_r(x)$ denote the Euclidean ball centered at x and with radius r, and let $Q_r(x)$ denote the corresponding l_∞ ball (a cube); that is, $Q_r(x) = [-r, r]^d + x$. For $F \subset \mathbb{R}^d$ let $|F|$ denote the Lebesgue measure of the set F, let ∂F denote the intersection of the closure of F with that of its complement and for $r > 0$, set $\partial_r F = \cup_{x \in \partial F} Q_r(x)$, the r-neighborhood of the boundary of F. Let $\operatorname{diam}(F) = \sup\{|x - y|\colon x, y \in F\}$, and let $\operatorname{card}(F)$ denote its cardinality (when finite).

Let $\xrightarrow{\mathcal{D}}$ denote convergence in distribution, let $\xrightarrow{P}$ denote convergence in probability and let $\mathcal{N}(\mu, \sigma^2)$ denote a normally distributed random variable with mean μ and variance σ^2.

2. A general central limit theorem.

Let $d \geq 1$. Throughout the rest of this paper, $\lambda > 0$ is a constant and $(B_n)_{n \geq 1}$ denotes a sequence of bounded Borel subsets ("regions" or "windows") of $\mathbb{R}^d$, satisfying the following conditions. First, $|B_n| = n/\lambda$ for all n; second, B_n tends to $\mathbb{R}^d$, by which we mean $\cup_{n \geq 1} \cap_{m \geq n} B_m = \mathbb{R}^d$; third, $\lim_{n \to \infty}(|\partial_r B_n|/n) = 0$ for all $r > 0$ (the *vanishing relative boundary* condition) and fourth, there exists a constant β_1 such that $\operatorname{diam}(B_n) \leq \beta_1 n^{\beta_1}$ for all n (the *polynomial boundedness* condition on B_n). Subject to these conditions, the choice of $(B_n)_{n \geq 1}$ is arbitrary.

Let $U_{1,n}, U_{2,n}, \ldots$ be independent identically distributed uniform variables on B_n. Let $\mathcal{U}_{m,n} = \{U_{1,n}, \ldots, U_{m,n}\}$ (a binomial point process) and let $\mathcal{P}_n$ be a homogeneous Poisson process on B_n of intensity λ. Also, let B_0 be a fixed bounded Borel set in $\mathbb{R}^d$ satisfying $|B_0| = 1$ and $|\partial B_0| = 0$ (for example, the unit cube), and let $\mathscr{X}_n = \{X_1, \dots, X_n\}$, where $X_1, X_2, X_3, \dots$ are independent and uniformly distributed on B_0. All of our results refer to the point processes $\mathscr{P}_n$, $\mathscr{U}_{n,n}$ and $\mathscr{X}_n$, defined in this way.

In the proof of our results, we shall need to consider translates of the regions B_n. With this in mind, let $\mathscr{B}$ be the collection of all regions $A \subset \mathbb{R}^d$ of the form $A = \{B_n + x \colon x \in \mathbb{R}^d, n \geq 1\}$. Likewise, let $\mathscr{B}_0$ be the collection of all regions of the form $A = aB_0 + x$ with $a \geq 1$ and $x \in \mathbb{R}^d$.

Let H be a real-valued functional defined for all finite subsets of $\mathbb{R}^d$. Assume that H is translation-invariant, meaning that $H(\mathscr{X} + y) = H(\mathscr{X})$ for all $\mathscr{X} \subset \mathbb{R}^d$ and all $y \in \mathbb{R}^d$. We derive central limit theorems for $H(\mathscr{P}_n)$, $H(\mathscr{U}_{n,n})$ and $H(\mathscr{X}_n)$.

The conditions on H for our central limit theorems are defined in terms of the "add one cost," by which we mean the increment in H caused by inserting a point at the origin into a finite point set $\mathscr{X} \subset \mathbb{R}^d$, formally given by

$$\Delta(\mathscr{X}) := H(\mathscr{X} \cup \{0\}) - H(\mathscr{X}).$$

Let $\mathscr{P}$ be a homogeneous Poisson process of intensity λ on $\mathbb{R}^d$. Our first condition on H develops a notion of *stabilization* having its origins in [12, 13].

DEFINITION 2.1. The functional H is *strongly stabilizing* if there exist a.s. finite random variables S (a *radius of stabilization* of H) and $\Delta(\infty)$ such that with probability 1, $\Delta((\mathscr{P} \cap B_S(0)) \cup \mathscr{A}) = \Delta(\infty)$ for all finite $\mathscr{A} \subset (\mathbb{R}^d \setminus B_S(0))$.

Thus, S is a radius of stabilization if the add one cost for $\mathscr{P}$ is unaffected by changes in the configuration outside the ball $B_S(0)$.

Given $A \in \mathscr{B}$, let $\mathscr{U}_{m,A}$ be a point process consisting of m independent uniform variables on A. Our second condition on H is a uniform bound on the fourth moments of the add one cost for this point process. Our third condition is a mild uniform bound on the size of H.

DEFINITION 2.2. The functional H satisfies the *uniform bounded moments condition* on $\mathscr{B}$ if

$$\sup_{A \in \mathscr{B} \colon 0 \in A} \ \sup_{m \in [\lambda|A|/2, 3\lambda|A|/2]} \{\mathbb{E}[\Delta(\mathscr{U}_{m,A})^4]\} < \infty.$$

DEFINITION 2.3. The functional H is *polynomially bounded* if there exists a constant β_2 such that for all finite sets $\mathscr{X} \subset \mathbb{R}^d$,

$$|H(\mathscr{X})| \leq \beta_2(\operatorname{diam}(\mathscr{X}) + \operatorname{card}(\mathscr{X}))^{\beta_2}.$$

The following result is basic to this paper.

THEOREM 2.1. *Suppose that H is strongly stabilizing, satisfies the uniform bounded moments condition on $\mathscr{B}$, and is polynomially bounded. Then there* *exist constants* σ^2, τ^2, *with* $0 \le \tau^2 \le \sigma^2$, *such that as* $n \to \infty$,

$$n^{-1}\operatorname{Var}(H(\mathscr{P}_n)) \to \sigma^2 \tag{2.1}$$

and

$$n^{-1/2}(H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2), \tag{2.2}$$

while

$$n^{-1}\operatorname{Var}(H(\mathscr{U}_{n,n})) \to \tau^2 \tag{2.3}$$

and

$$n^{-1/2}(H(\mathscr{U}_{n,n}) - \mathbb{E}H(\mathscr{U}_{n,n})) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{2.4}$$

Also, given λ, σ^2 *and* τ^2 *are independent of the choice of* (B_n).

If the distribution of $\Delta(\infty)$ *is nondegenerate, then* $\tau^2 > 0$, *and hence also* $\sigma^2 > 0$.

The proof of Theorem 2.1 will show that $\tau^2 = \sigma^2 - (\mathbb{E}\Delta(\infty))^2$. In most of our examples $\mathbb{E}\Delta(\infty)$ will be strictly positive because adding a point tends to increase the value of the functional and thus τ^2 will be strictly less than σ^2. In other words, Poissonization contributes extra randomness which shows up in the limiting variance.

In Theorem 2.1 the condition $|B_n| = n/\lambda$ can be relaxed to

$$\limsup_{n\to\infty} n^{-1/2}|(n - \lambda|B_n|)| < \infty.$$

The proof under this weaker condition is essentially unchanged. Also, the polynomial boundedness condition can be weakened to (2.6) and (2.7) below, and the first two limits (2.1) and (2.2) remain true under somewhat weaker forms of the moments and stabilization conditions (Theorem 3.1 below).

To deduce CLTs for functionals H on the point process $\mathscr{X}_n$ of independent uniform points in B_0, we require one further scaling property for H. Given $\gamma \in \mathbb{R}$, we shall say H is *homogeneous of order* γ if for all $\mathscr{X} \subset \mathbb{R}^d$ on which H is defined, and all $a \in \mathbb{R}$,

$$H(a\mathscr{X}) = a^\gamma H(\mathscr{X}).$$

If H satisfies homogeneity, it is easy to deduce from the above theorems a CLT for homogeneous Poisson processes of high intensity, or for a large sample of nonrandom size from the uniform distribution, on B_0. We just state such a result for the sample $\mathscr{X}_n$ of large nonrandom size on B_0.

COROLLARY 2.1. *Suppose* H *is strongly stabilizing, satisfies the uniform bounded moments condition on* $\mathscr{B}_0$, *is polynomially bounded and is homogeneous of order* γ. *Then with* τ^2 *the constant given in the case* $\lambda = 1$ *of Theorem* 2.1, $n^{(2\gamma/d)-1}\operatorname{Var} H(\mathscr{X}_n) \to \tau^2$, *and*

$$n^{(\gamma/d)-1/2}(H(\mathscr{X}_n) - \mathbb{E}H(\mathscr{X}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2).$$

Corollary 2.1 is easily proved by assuming the origin lies in the interior of B_0 (if not, consider a suitable translate), taking $B_n = n^{1/d} B_0$, applying Theorem 2.1 and using homogeneity of H.

If H is homogeneous, the general case of Theorem 2.1 follows from the special case $\lambda = 1$. All of the specific functionals considered in detail here will be homogeneous for some γ, so we consider only the case $\lambda = 1$ for these examples. For examples where H is *not* homogeneous, see the last example in Section 9 and also [20].

Kesten and Lee [11] essentially showed that the total edge length of the power-weighted Euclidean minimal spanning tree satisfies all the conditions of Theorem 2.1 and Corollary 2.1. In this way they proved a nontrivial CLT for one of the archetypical problems of combinatorial optimization. The second half of this paper shows that various functionals of proximity graphs also satisfy the conditions of Theorem 2.1. Some of these applications involve some intricate work. In a related paper [20], we generalize Theorem 2.1 to functionals of *marked* point processes and thus central limit theorems for sphere packing and related problems.

A long-standing open problem is to find convergence of the variance, and a CLT, for the length of the optimal traveling salesman tour on $\mathscr{X}_n$ (or on a Poissonized point process). Other open problems of this kind concern the total length of the minimal matching on random points and the total length of the Steiner tree on random points. Our results show that one possible approach involves showing strong stabilization for these functionals. Although a proof of stabilization remains elusive, we believe that the approach here might be useful in attacking such problems.

In many of our applications, the following condition on $\mathscr{B}$ will be used for checking the bounded moments condition. Let us say $\mathscr{B}$ is *regular* if there exists $\delta > 0$ such that for all $r \in [1, \infty)$, whenever $A \in \mathscr{B}$ and $x, y \in A$ with $|x - y| = r$, we have

$$|B_{r/4}(x) \cap A| \geq \delta r^{\delta}. \tag{2.5}$$

By a *box* we shall mean a set $B \subset \mathbb{R}^d$ of the form $\prod_{i=1}^d [a_i, b_i]$, with $b_i \geq a_i + 1$ for each i. It is not hard to show that the collection of all boxes is regular, and therefore in applications that require regularity, taking the sets B_n to be all boxes is sufficient to ensure that $\mathscr{B}$ is regular. Incidentally, if the sets B_n are all boxes then the vanishing boundary and polynomial boundedness conditions for B_n follow automatically from the assumptions that $B_n \to \mathbb{R}^d$ and $|B_n| = n/\lambda$.

If instead of boxes, $\mathscr{B}$ is a set of balls or ellipsoids, then again it is regular. It can be seen that a sufficient condition on B_0 for $\mathscr{B}_0$ to be regular is that B_0 has a reasonably smooth boundary in the sense that $r^{-d}|B_r(x) \cap B_0|$ is bounded away from zero, uniformly over $x \in B_0$ and $r \in (0, 1]$.

The next three sections are devoted to the proofs of our main results. The proofs make heavy use of the uniform fourth moment condition. It is likely that this can be replaced by a $2 + \varepsilon$ moment condition, but since the weaker condition does not seem to increase the range of applications considered here, we have used the fourth moment condition for technical ease. Our CLTs also hold if the deterministic polynomial boundedness conditions on H and B_n are replaced by the weaker moment bounds

$$\max(\mathbb{E}[H(\mathscr{U}_{n,n})^4], \mathbb{E}[H(\mathscr{P}_n)^4]) \leq \beta_3 n^{\beta_3} \tag{2.6}$$

and for all $A \in \mathscr{B}$,

$$E[\Delta(\mathscr{P} \cap A)^8] \leq \beta_4 |A|^{\beta_4}. \tag{2.7}$$

We will actually use only these weaker conditions in the proofs.

3. Proof of CLT: the Poisson case.

In this section we prove a CLT for $H(\mathscr{P}_n)$ only, under somewhat different conditions than those of Theorem 2.1. These conditions are in fact weaker (see Lemma 4.1) and thus we actually establish (2.1) and (2.2).

DEFINITION 3.1. The functional H is *weakly stabilizing on $\mathscr{B}$* if there is a random variable $\Delta(\infty)$ such that $\Delta(\mathscr{P} \cap A) \xrightarrow{\text{a.s.}} \Delta(\infty)$ as $A \to \mathbb{R}^d$ through $\mathscr{B}$, by which we mean that for any $\mathscr{B}$-valued sequence $(A_n)_{n\geq 1}$ that tends to $\mathbb{R}^d$, $\Delta(\mathscr{P} \cap A_n) \to \Delta(\infty)$ as $n \to \infty$, almost surely.

Observe that strong stabilization implies weak stabilization on $\mathscr{B}$. In fact, for all specific examples considered in this paper, strong stabilization holds. We retain the distinction, first to emphasize which properties are used in the proofs, and second, to allow for the possibility that in some cases not considered here (see, e.g., [19]), it may be possible to prove weak stabilization but not strong stabilization.

As shown in Section 4, the following moments condition is weaker than the uniform moments condition. It is all we need in the Poisson setting.

DEFINITION 3.2. The functional H satisfies the *Poisson bounded moments condition on $\mathscr{B}$* if

$$\sup_{A \in \mathscr{B}:\ 0 \in A} \{\mathbb{E}[\Delta(\mathscr{P} \cap A)^4]\} < \infty.$$

THEOREM 3.1. *Suppose that H is weakly stabilizing on $\mathscr{B}$ and satisfies the Poisson bounded moments condition on $\mathscr{B}$. Then there exists $\sigma^2 \geq 0$ such that as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathscr{P}_n)) \to \sigma^2$ and $n^{-1/2}(H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2)$.*

Theorem 3.1 still holds if in Definition 3.1, almost sure convergence is relaxed to convergence in probability. The proof under this weaker condition is essentially the same as that given, but entails some extra subsequence arguments.

The first step toward a proof of Theorem 3.1 is to show that the conditions in Definitions 3.1 and 3.2 imply alternative stabilization and moment conditions, referring to the modification of the homogeneous Poisson process $\mathscr{P}$ by replacing those Poisson points lying in a unit cube with an independent Poisson process on that unit cube, rather than inserting a single point. Formally, this modification is defined as follows.

Let $\mathscr{P}'$ be an independent copy of the Poisson process $\mathscr{P}$. For $x \in \mathbb{Z}^d$, set

$$\mathscr{P}''(x) = (\mathscr{P} \setminus Q_{1/2}(x)) \cup (\mathscr{P}' \cap Q_{1/2}(x)).$$

Then, given a translation-invariant functional H of point sets in $\mathbb{R}^d$, define

$$\Delta_x(A) := H(\mathscr{P}''(x) \cap A) - H(\mathscr{P} \cap A). \tag{3.1}$$

LEMMA 3.1. *Suppose H is weakly stabilizing on $\mathscr{B}$. Then for all $x \in \mathbb{Z}^d$, there is a random variable $\Delta_x(\infty)$ such that for all $x \in \mathbb{Z}^d$,*

$$\Delta_x(A) \xrightarrow{\text{a.s.}} \Delta_x(\infty) \quad \text{as } A \to \mathbb{R}^d \quad \textit{through } \mathscr{B}. \tag{3.2}$$

Moreover, if H satisfies the Poisson bounded moments condition on $\mathscr{B}$, then

$$\sup_{A \in \mathscr{B},\ x \in \mathbb{Z}^d} \mathbb{E}[(\Delta_x(A))^4] < \infty. \tag{3.3}$$

PROOF. Set $C_0 = Q_{1/2}(0)$. To prove (3.2), by translation-invariance it suffices to consider the case $x = 0$, and therefore suffices to prove that the variables $H(\mathscr{P} \cap A) - H(\mathscr{P} \cap A \setminus C_0)$ converge almost surely as $A \to \mathbb{R}^d$ through $\mathscr{B}$.

The number N of points of $\mathscr{P}$ in C_0 is a Poisson variable with parameter λ. Let $V_1, V_2, \dots, V_N$ be the points of $\mathscr{P} \cap C_0$, taken in an order chosen uniformly at random from the $N!$ possibilities. Then, provided $C_0 \subseteq A$,

$$H(\mathscr{P} \cap A) - H(\mathscr{P} \cap A \setminus C_0) = \sum_{i=0}^{N-1} \delta_i(A),$$

where

$$\delta_i(A) = H((\mathscr{P} \cap A \setminus C_0) \cup \{V_1, \dots, V_{i+1}\}) - H((\mathscr{P} \cap A \setminus C_0) \cup \{V_1, \dots, V_i\}).$$

Since N is a.s. finite, it suffices to prove each $\delta_i(A)$ converges almost surely as $A \to \mathbb{R}^d$ through $\mathscr{B}$. Let U be a uniform variable on C_0, independent of $\mathscr{P}$. The distribution of the translated point process $(\{V_1, \dots, V_i\} \cup (\mathscr{P} \setminus C_0)) - V_{i+1}$ is the same as the conditional distribution of the Poisson process $\mathscr{P}$ given that the number of points of $\mathscr{P}$ in $C_0 - U$ is equal to i, an event of strictly positive probability. By assumption, this satisfies weak stabilization, which proves (3.2).

Next we prove (3.3) under the Poisson bounded moments assertion. If $Q_{1/2}(x) \cap A = \varnothing$ then $\Delta_x(A)$ is zero, a.s. By translation-invariance it suffices to consider the case with $x = 0$, that is, to prove

$$\sup_{A \in \mathscr{B}:\ C_0 \cap A \neq \varnothing} \mathbb{E}[(\Delta_0(A))^4] < \infty. \tag{3.4}$$

Let N' be the number of points of $\mathscr{P}$ in $C_0 \cap A$; in cases with $C_0 \subseteq A$ this is the same as N. In general, N' has a Poisson distribution with mean $\mu := \lambda|C_0 \cap A|$. Let $V_1, \dots, V_{N'}$ be the points of $\mathscr{P} \cap C_0 \cap A$, taken in random order. Then

$$H(\mathscr{P} \cap A) - H(\mathscr{P} \cap A \backslash C_0) = \sum_{i=0}^{N'-1} \delta_i(A),$$

where $\delta_i(A)$ is defined above. Also, $H(\mathscr{P}''(0) \cap A) - H(\mathscr{P}''(0) \cap A \backslash C_0)$ has the same distribution as $H(\mathscr{P} \cap A) - H(\mathscr{P} \cap A \backslash C_0)$. Hence, it suffices to prove

$$\sup_{A \in \mathscr{B}:\ C_0 \cap A \neq \varnothing} \mathbb{E}\left[\left(\sum_{i=0}^{N'-1} \delta_i(A)\right)^4\right] < \infty.$$

Set $p_k = P[N' = k]$. Then, writing simply δ_i for $\delta_i(A)$, we have

$$\begin{aligned} \mathbb{E}\left[\left(\sum_{i=0}^{N'-1} \delta_i\right)^4\right] &= \sum_{k=1}^{\infty} p_k \mathbb{E}\left[\left(\delta_{k-2} + \delta_{k-1} + \sum_{i=0}^{k-3} \delta_i\right)^4\right] \\ &\le \sum_{k=1}^{\infty} p_k 3^3 \mathbb{E}\left(\delta_{k-2}^4 + \delta_{k-1}^4 + (k-2)^3 \sum_{i=0}^{k-3} \delta_i^4\right), \end{aligned} \tag{3.5}$$

where the last line is obtained using Cauchy–Schwarz, where we set $\delta_{-1} = \delta_{-2} = 0$ and where the summation $\sum_{i=0}^{k-3}$ is taken to be zero when $k \le 2$.

Consider the final sum. Rearranging, we have

$$\sum_{k=3}^{\infty} p_k (k-2)^3 \sum_{i=0}^{k-3} \mathbb{E}\delta_i^4 = \sum_{i=0}^{\infty} q_i \mathbb{E}\delta_i^4,$$

with $q_i = \sum_{k \ge i+3} (k-2)^3 p_k$. Then

$$\begin{aligned} \frac{q_i}{p_i} &= \sum_{k \ge i+3} \frac{i!\mu^{k-i}(k-2)^3}{k!} \le \sum_{k \ge i+3} \frac{i!\mu^{k-i}}{(k-3)!} \\ &= \sum_{j \ge 0} \frac{\mu^{j+3} i!}{(i+j)!} \le \mu^3 \sum_{j \ge 0} \frac{\mu^j}{j!} = \mu^3 e^{\mu}. \end{aligned}$$

Hence, $\sum_{i=0}^{\infty} q_i \mathbb{E}\delta_i^4 \le \mu^3 e^{\mu} \sum_{i=0}^{\infty} p_i \mathbb{E}\delta_i^4$.

Similarly, since $p_{i+2}/p_i = \mu^2/((i+2)(i+1)) \le \mu^2$, we have

$$\sum_{k=2}^{\infty} p_k \mathbb{E}\delta_{k-2}^4 = \sum_{i=0}^{\infty} p_{i+2} \mathbb{E}\delta_i^4 \le \mu^2 \sum_{i=0}^{\infty} p_i \mathbb{E}\delta_i^4.$$

One obtains a similar bound (this time with a factor of μ) in the case of $\mathbb{E}\delta_{k-1}^4$. Combining all these estimates in (3.5), and setting $c(\mu) = 27(\mu^3 e^{\mu} + \mu^2 + \mu)$, we have

$$\mathbb{E}\left[\left(\sum_{i=0}^{N'-1}\delta_i\right)^4\right] \le c(\mu)\sum_{i=0}^{\infty} p_i \mathbb{E}\delta_i^4 = c(\mu)\int_{C_0\cap A}\frac{dx}{|C_0\cap A|}\sum_{i=0}^{\infty} p_i \mathbb{E}[\delta_i^4 | V_{i+1} = x]$$

$$= c(\mu)\int_{C_0\cap A}\frac{dx}{|C_0\cap A|}\mathbb{E}\Delta(\mathscr{P}\cap(A-x))^4.$$

By the Poisson bounded moments condition, and the fact that $\mu \le \lambda$, this is uniformly bounded, yielding (3.3). □

PROOF OF THEOREM 3.1. The proof is similar to that of Theorem 2.1 of [19], adapted to the continuum. Note that $\mathscr{P}_n$ has the same distribution as $\mathscr{P}\cap B_n$, so for this proof without loss of generality we assume $\mathscr{P}_n = \mathscr{P}\cap B_n$.

For $x \in \mathbb{Z}^d$, let $\mathscr{F}_x$ denote the σ-field generated by the points of $\mathscr{P}$ in $\cup_{y\preceq x} Q_{1/2}(y)$, where $y \preceq x$ means $y \in \mathbb{Z}^d$ and y precedes or equals x in the lexicographic ordering on $\mathbb{Z}^d$. In other words, $\mathscr{F}_x$ is the smallest σ-field, with respect to which the number of Poisson points in any bounded Borel subset of $\cup_{y\preceq x} Q_{1/2}(y)$ is measurable.

Let B'_n be the set of lattice points $x \in \mathbb{Z}^d$ such that $|Q_{1/2}(x)\cap B_n| \ne \varnothing$. Let $k_n = \mathrm{card}(B'_n)$; then k_n/n tends to λ^{-1} because of the vanishing relative boundary condition and the fact that

$$B_n \subseteq \bigcup_{x\in B'_n} Q_{1/2}(x) \subseteq B_n \cup \partial_1(B_n).$$

Define the filtration $(\mathscr{G}_0, \mathscr{G}_1, \dots, \mathscr{G}_{k_n})$ as follows: let $\mathscr{G}_0$ be the trivial σ-field, label the elements of B'_n in lexicographic order as $x_1, \dots, x_{k_n}$ and let $\mathscr{G}_i = \mathscr{F}_{x_i}$ for $1 \le i \le k_n$. Then $H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n) = \sum_{i=1}^{k_n} D_i$ where we set

$$D_i = \mathbb{E}[H(\mathscr{P}_n)|\mathscr{G}_i] - \mathbb{E}[H(\mathscr{P}_n)|\mathscr{G}_{i-1}] = \mathbb{E}[\Delta_{x_i}(B_n)|\mathscr{F}_{x_i}], \tag{3.6}$$

with $\Delta_{x_i}(B_n)$ defined by (3.1). By orthogonality of martingale differences, $\mathrm{Var}\,[H(\mathscr{P}_n)] = \mathbb{E}\sum_{i=1}^{k_n} D_i^2$. By this fact, along with a CLT for martingale differences (Theorem (2.3) of [15]), it suffices to prove the conditions

$$\sup_{n\ge 1}\mathbb{E}\Big[\max_{1\le i\le k_n}\big(k_n^{-1/2}|D_i|\big)^2\Big] < \infty, \tag{3.7}$$

$$k_n^{-1/2}\max_{1\le i\le k_n}|D_i| \xrightarrow{P} 0, \tag{3.8}$$

and for some $\sigma^2 \ge 0$,

$$k_n^{-1}\sum_{i=1}^{k_n} D_i^2 \xrightarrow{L^1} \lambda\sigma^2. \tag{3.9}$$

The factor of λ is included in (3.9) to make σ^2 consistently defined.

Using the representation $D_i = \mathbb{E}[\Delta_{x_i}(B_n)|\mathscr{F}_{x_i}]$ we may easily check conditions (3.7) and (3.8). Indeed, by the conditional Jensen's inequality we have

$$k_n^{-1}\mathbb{E}\Big[\max_{i\le k_n} D_i^2\Big] \le k_n^{-1}\sum_{i=1}^{k_n}\mathbb{E}[D_i^2] \le k_n^{-1}\sum_{i=1}^{k_n}\mathbb{E}[\Delta_{x_i}(B_n)^2],$$

which is uniformly bounded by the alternative bounded moments condition (3.3).

For the second condition (3.8), we use Boole's and Markov's inequalities to obtain

$$P\Big[\max_{1\le i\le k_n}|D_i| \ge k_n^{1/2}\varepsilon\Big] \le \sum_{i=1}^{k_n}\frac{\mathbb{E}[D_i^4]}{k_n^2\varepsilon^4},$$

which tends to zero, again by (3.3).

We now prove (3.9). By the alternative stabilization condition (3.2), for each $x \in \mathbb{Z}^d$ the variables $\Delta_x(A)$ converge almost surely to a limit, denoted $\Delta_x(\infty)$, as $A \to \mathbb{R}^d$ through $\mathscr{B}$. For $x \in \mathbb{Z}^d$ and $A \in \mathscr{B}$, let

$$F_x(A) = \mathbb{E}[\Delta_x(A)|\mathscr{F}_x]; \qquad F_x = \mathbb{E}[\Delta_x(\infty)|\mathscr{F}_x].$$

Then $(F_x, x \in \mathbb{Z}^d)$ is a stationary family of random variables. Set $\sigma^2 = \mathbb{E}[F_0^2]/\lambda$. We claim that the pointwise ergodic theorem (see [4], Chapter 6) implies

$$k_n^{-1}\sum_{x\in B'_n}F_x^2 \xrightarrow{L^1} \lambda\sigma^2. \tag{3.10}$$

To prove this, let $e_1 = (1, 0, \dots, 0) \in \mathbb{Z}^d$. The variables F_{ne_1}, $n \ge 1$, form an ergodic sequence because they take the form $f(T^n(V))$ where T is a shift operator on an i.i.d. sequence $V = \{V_z, z \in \mathbb{Z}\}$. Given $\varepsilon > 0$, by the ergodic theorem we can choose $K > 0$ such that for all $n \ge K$, the average of $F_{e_1}^2, F_{2e_1}^2, \dots, F_{ne_1}^2$ is within an L^1 distance at most ε of $\lambda\sigma^2$. Divide B'_n into one-dimensional *intervals* by which we mean maximal subsets of B_n of the form $(\mathbb{Z}\cap[a,b]) \times \{z_2\} \times \cdots \times \{z_d\}$, with $a, b, z_2, \dots, z_d$ in $\mathbb{Z}$. Let B_n^* be the union of constituent intervals of length at least K. Let $k_n^* = \operatorname{card}(B_n^*)$. Since $(B_n)_{n\ge1}$ has vanishing relative boundary, $\lim_{n\to\infty}(k_n^*/k_n) = 1$. Writing $\|\cdot\|_1$ for the L^1-norm of random variables, we have

$$\begin{aligned}\left\|\left(k_n^{-1}\sum_{x\in B'_n}F_x^2\right) - \lambda\sigma^2\right\|_1 &\le k_n^{-1}\left\|\left(\sum_{x\in B_n^*}F_x^2\right) - k_n^*\lambda\sigma^2\right\|_1\\ &\quad + k_n^{-1}\left\|\left(\sum_{x\in B'_n\setminus B_n^*}F_x^2\right) - (k_n - k_n^*)\lambda\sigma^2\right\|_1.\end{aligned} \tag{3.11}$$

By the choice of K and translation-invariance, for each interval I of length at least K the average of F_z^2, $z \in I$, is within an L^1-distance ε of $\mathbb{E}[F_0^2]$. Therefore the first term on the right-hand side of (3.11) is at most ε, while the second term tends to zero because $(k_n^*/k_n) \to 1$. Therefore the left side of (3.11) is less than 2ε for large n, and (3.10) follows.

We need to show that $F_x(B_n)^2$ approximates to F_x^2. We consider x at the origin 0. For any $A \in \mathscr{B}$, by Cauchy–Schwarz,

$$\mathbb{E}[|F_0(A)^2 - F_0^2|] \le (\mathbb{E}[(F_0(A) + F_0)^2])^{1/2}(\mathbb{E}[(F_0(A) - F_0)^2])^{1/2}.$$

By the definition of F_0 and the conditional Jensen's inequality,

$$\mathbb{E}[(F_0(A) + F_0)^2] = \mathbb{E}[(\mathbb{E}[\Delta_0(A) + \Delta_0(\infty)|\mathscr{F}_0])^2] \\ \le \mathbb{E}[\mathbb{E}[(\Delta_0(A) + \Delta_0(\infty))^2|\mathscr{F}_0]] = \mathbb{E}[(\Delta_0(A) + \Delta_0(\infty))^2]$$

which is uniformly bounded by the alternative stabilization and bounded moments conditions (3.2) and (3.3). Similarly,

$$\mathbb{E}[(F_0(A) - F_0)^2] \le \mathbb{E}[(\Delta_0(A) - \Delta_0(\infty))^2]. \tag{3.12}$$

By (3.2) and (3.3) this is also uniformly bounded. For any $\mathscr{B}$-valued sequence $(A_n)_{n\ge 1}$ tending to $\mathbb{R}^d$, the sequence $(\Delta_0(A_n) - \Delta_0(\infty))^2$ tends to 0 a.s. by (3.2), and is uniformly integrable by (3.3), and therefore (see [4], Chapter 4, Theorem 5.2) the expression (3.12) tends to zero so that $\mathbb{E}[|F_0(A_n)^2 - F_0^2|] \to 0$.

Returning to the given sequence (B_n), let $\varepsilon > 0$. By the vanishing relative boundary condition, we can choose K_n so that $\lim_{n\to\infty} K_n = \infty$ and $|\partial_{K_n}(B_n)| \le \varepsilon n$ for all n. Set $B''_n = B'_n \setminus \partial_{K_n} B_n$. Using the conclusion of the previous paragraph and translation-invariance, it is not hard to deduce that

$$\lim_{n\to\infty} \sup_{x \in B''_n} \mathbb{E}[|F_x(B_n)^2 - F_x^2|] = 0. \tag{3.13}$$

Using (3.13), the uniform boundedness of $\mathbb{E}[|F_x(B_n)^2 - F_x^2|]$ and the fact that ε can be taken arbitrarily small in the above argument, it is routine to deduce that

$$k_n^{-1} \sum_{x \in B_n} (F_x(B_n)^2 - F_x^2) \xrightarrow{L^1} 0,$$

and therefore (3.10) remains true with F_x replaced by $F_x(B_n)$; that is, (3.9) holds and the proof of Theorem 2.1 is complete. □

4. Proof of CLT: the non-Poisson case.

In this section we prove Theorem 2.1, subject to showing that the limiting variance τ^2 is nonzero. The first step is to show that the conditions of Theorem 2.1 imply those of Theorem 3.1, as follows.

LEMMA 4.1. *If H satisfies the uniform bounded moments condition and is polynomially bounded, then H satisfies the Poisson bounded moments condition.*

PROOF. Suppose $A \in \mathscr{B}$. Let N be the number of Poisson points in A. Then

$$\mathbb{E}\Delta(\mathscr{P} \cap A)^4 = \sum_{m=0}^{\infty} P[N = m]\mathbb{E}\Delta(\mathscr{U}_{m,A})^4.$$

Assuming the uniform bounded moments condition, the restriction of the sum to those m with $\lambda|A|/2 \le m \le 3\lambda|A|/2$ is uniformly bounded. By Cauchy–Schwarz, the remainder of the sum is bounded by the square root of $\mathbb{E}[\Delta(\mathscr{P}\cap A)^8]P[|(N-\lambda|A|)| > |A|/2]$. By (2.7), this too is uniformly bounded. □

It follows from Lemma 4.1 that if H satisfies the conditions of Theorem 2.1, then Theorem 3.1 applies and we have (2.1) and (2.2). To de-Poissonize these limits and obtain (2.3) and (2.4), we use a technique related to that used in [11] and [12]. We make the definition

$$R_{m,n} = H(\mathscr{U}_{m+1,n}) - H(\mathscr{U}_{m,n}), \tag{4.1}$$

and use the following coupling lemma.

LEMMA 4.2. *Suppose H is strongly stabilizing. Let $\varepsilon > 0$. Then there exists $\delta > 0$ and $n_0 \ge 1$ such that for all $n \ge n_0$ and all $m, m' \in [(1-\delta)n, (1+\delta)n]$ with $m < m'$, there exists a coupled family of variables D, D', R, R' with the following properties*:

(i) *D and D' each have the same distribution as $\Delta(\infty)$.*
(ii) *D and D' are independent.*
(iii) *(R, R') have the same joint distribution as $(R_{m,n}, R_{m',n})$.*
(iv) *$P[\{D \ne R\} \cup \{D' \ne R'\}] < \varepsilon$.*

PROOF. Suppose we are given n. On a suitable probability space, let $\mathscr{P}$ and $\mathscr{P}'$ be independent homogeneous Poisson processes on $\mathbb{R}^d$ of intensity λ; let $U, U', V_1, V_2, \dots$ be independent variables uniformly distributed over B_n, independent of $\mathscr{P}$ and $\mathscr{P}'$. The variables U and U' will play the role of $U_{m,n}$ and $U_{m',n}$.

Let $\mathscr{P}''$ be the point process consisting of those points of $\mathscr{P}$ which lie closer to U than to U' (in the Euclidean norm), together with those points of $\mathscr{P}'$ which lie closer to U' than to U. Clearly $\mathscr{P}''$ is a Poisson process of rate λ on $\mathbb{R}^d$, and moreover it is independent of U and of U'.

Let N denote the number of points of $\mathscr{P}''$ lying in B_n (a Poisson variable with mean n). Choose an ordering on the points of $\mathscr{P}''$ lying in B_n, uniformly at random from all $N!$ possible such orderings. Use this ordering to list the points of $\mathscr{P}''$ in B_n as $W_1, W_2, \dots, W_N$. Also, set $W_{N+1} = V_1$, $W_{N+2} = V_2$, $W_{N+3} = V_3$ and so on.

Let $R = H(\{W_1, \dots, W_m, U\}) - H(\{W_1, \dots, W_m\})$. Let $R' = H(\{W_1, \dots, W_{m'-1}, U, U'\}) - H(\{W_1, \dots, W_{m'-1}, U\})$. The variables $U, U', W_1, W_2, W_3, \dots$ are independent uniformly distributed variables on B_n, and therefore the pairs (R, R') and $(R_{m,n}, R_{m',n})$ have the same joint distribution as claimed.

Let $\widetilde{\mathscr{P}}$ be the translated point process $\mathscr{P} - U$. Similarly, let $\widetilde{\mathscr{P}}' = \mathscr{P}' - U'$. Clearly, $\widetilde{\mathscr{P}}$ and $\widetilde{\mathscr{P}}'$ are independent Poisson processes of rate λ on $\mathbb{R}^d$. Let S, S'

be radii of stabilization for $\widetilde{\mathscr{P}}$ and $\widetilde{\mathscr{P}}'$, respectively, and define

$$D = \Delta(\widetilde{\mathscr{P}} \cap B_S(0)); \qquad D' = \Delta(\widetilde{\mathscr{P}}' \cap B_{S'}(0)).$$

Then D and D' are independent, and each have the same distribution as $\Delta(\infty)$.

It remains to show that $(D, D') = (R, R')$ with high probability. Choose K such that $P[S > K] < \varepsilon/9$ and $P[S' > K] < \varepsilon/9$. Using the vanishing relative boundary condition, take n to be so large that except on an event (denoted E_0) of probability less than $\varepsilon/9$, the positions of U and U' are sufficiently far from ∂B_n and from each other, that the cubes $Q_K(U)$ and $Q_K(U')$ are contained entirely within B_n and also are such that every point of $Q_K(U)$ lies closer to U than to U' and every point of $Q_K(U')$ lies closer to U' than to U.

Set $\delta = \varepsilon(2K)^{-d}/(18\lambda)$. We assume $|m - n| \le \delta n$ and $|m' - n| \le \delta n$. For n large enough, except on an event (denoted E_1) of probability at most $\varepsilon/9$, we have $|N-m| \le 2\delta n = \varepsilon(2K)^{-d}n/(9\lambda)$, and likewise $|N-m'| \le \varepsilon(2K)^{-d}n/(9\lambda)$.

Let E be the event that the set of points of $\{W_1, \dots, W_m\}$ lying in $Q_K(U)$ is not the same as the set of points of $\mathscr{P}$ lying in $Q_K(U)$. This will happen either if one or more of the $(N - m)^+$ "discarded" points of $\mathscr{P}''$ or one or more of the $(m - N)^+$ "added" points of $\{V_1, V_2, \dots\}$ lies in $Q_K(U)$. For each added or discarded point, the probability of lying in $Q_K(U)$ is at most $(2K)^d\lambda/n$, and so the probability of E, given that E_1 does not occur, is less than $\varepsilon/9$. Similarly, with E' denoting the event that the set of points of $\{W_1, \dots, W_{m'}\}$ lying in $Q_K(U')$ is not the same as the set of points of $\mathscr{P}'$ lying in $Q_K(U')$, we have $P[E'|E_1^c] \le \varepsilon/9$.

Combining all these estimates, using the definition of the radius of (strong) stabilization for $\mathscr{P}$ and $\mathscr{P}'$, and using Boole's inequality, we obtain for large enough n that

$$\begin{aligned} P[(D, D') \ne (R, R')] \le {} & P[E_0] + P[E_1] + P[S > K] + P[S' > K] \\ & + P[E \setminus E_1] + P[E' \setminus E_1] \le \varepsilon. \end{aligned}$$
□

LEMMA 4.3. *Suppose H is strongly stabilizing and satisfies the uniform bounded moments condition. Let $(h(n))_{n\ge 1}$ be a sequence with $h(n)/n \to 0$ as $n \to \infty$. Then*

$$\lim_{n\to\infty} \sup_{n-h(n)\le m\le n+h(n)} |\mathbb{E}R_{m,n} - \mathbb{E}\Delta(\infty)| = 0. \tag{4.2}$$

Also

$$\lim_{n\to\infty} \sup_{n-h(n)\le m< m'\le n+h(n)} |\mathbb{E}R_{m,n}R_{m',n} - (\mathbb{E}\Delta(\infty))^2| = 0 \tag{4.3}$$

and

$$\lim_{n\to\infty} \sup_{n-h(n)\le m\le n+h(n)} |\mathbb{E}R_{m,n}^2| < \infty. \tag{4.4}$$

PROOF. Let m be an arbitrary integer satisfying $n - h(n) \le m \le n + h(n)$. Let $\varepsilon > 0$. Provided n is large enough, by Lemma 4.2 we can find coupled variables D and R, with D having the same distribution as $\Delta(\infty)$, with R having the same distribution as $R_{m,n}$ and with $P[D \neq R] < \varepsilon$. Then

$$\mathbb{E}R_{m,n} = \mathbb{E}R = \mathbb{E}[D] - \mathbb{E}[D\mathbf{1}\{D \neq R\}] + \mathbb{E}[R\mathbf{1}\{D \neq R\}].$$

By Cauchy–Schwarz and the Poisson and uniform moments conditions, there is a constant c, independent of ε, such that $|\mathbb{E}[D\mathbf{1}\{D \neq R\}]| \leq c\varepsilon^{1/2}$ and $|\mathbb{E}[R\mathbf{1}\{D \neq R\}]| \leq c\varepsilon^{1/2}$. Since ε is arbitrarily small, (4.2) follows. Moreover, the proof of (4.4) is very similar and is omitted.

Next we consider m, m' with $n - h(n) \leq m < m' \leq n + h(n)$. By Lemma 4.2 we can find coupled variables D, D', R, R' such that D and D' are independent and each have the same distribution as $\Delta(\infty)$, (R, R') have the same joint distribution as $(R_{m,n}, R_{m',n})$, and $P[(D, D') \neq (R, R')] < \varepsilon$. Then

$$\begin{aligned}\mathbb{E}[RR'] - \mathbb{E}[DD'] = {} & \mathbb{E}[RR'\mathbf{1}\{(D, D') \neq (R, R')\}] \\ & - \mathbb{E}[DD'\mathbf{1}\{(D, D') \neq (R, R')\}],\end{aligned}$$

and by Cauchy–Schwarz and the Poisson and uniform moment conditions, the right-hand side has modulus bounded by a constant multiple of $\varepsilon^{1/2}$. Since ε is arbitrarily small, (4.3) follows. □

PROOF OF THEOREM 2.1. We prove here the limits (2.3) and (2.4). We defer showing the strict positivity of τ^2 until the next section.

Let $H_n = H(\mathscr{U}_{n,n})$ and $H'_n = H(\mathscr{P}_n)$. For this proof, assume $\mathscr{P}_n$ is coupled to $\mathscr{U}_{n,n}$ by setting $\mathscr{P}_n = \{U_{1,n}, U_{2,n}, \ldots, U_{N_n,n}\}$ with N_n an independent Poisson variable with mean n. Let $\alpha = \mathbb{E}\Delta(\infty)$. The first step is to prove that as $n \to \infty$,

$$\mathbb{E}\,[(n^{-1/2}(H'_n - H_n - (N_n - n)\alpha))^2] \to 0. \tag{4.5}$$

To prove this, note that the expectation in the left-hand side is equal to

$$\begin{aligned}&\sum_{m:\ |m-n|\leq n^{3/4}} \mathbb{E}\,[n^{-1}(H(\mathscr{U}_{m,n}) - H(\mathscr{U}_{n,n}) - (m-n)\alpha)^2]P[N_n = m] \\ &\qquad + n^{-1}\mathbb{E}\,[(H'_n - H_n - (N_n - n)\alpha)^2\mathbf{1}\{|N_n - n| > n^{3/4}\}].\end{aligned} \tag{4.6}$$

Let $\varepsilon > 0$. By (4.1) and Lemma 4.3, there exists $c > 0$ such that for large enough n and all m with $n \leq m \leq n + n^{3/4}$,

$$\begin{aligned}&\mathbb{E}[(H(\mathscr{U}_{m,n}) - H(\mathscr{U}_{n,n}) - (m-n)\alpha)^2] \\ &\qquad = \mathbb{E}\left[\left(\sum_{l=n}^{m-1}(R_{l,n} - \alpha)\right)^2\right] \leq \varepsilon(m-n)^2 + c(m-n),\end{aligned}$$

where the bound comes from expanding out the double sum arising from the expectation of the squared sum. A similar argument applies when $n - n^{3/4} \leq m \leq n$, and hence the first term in (4.6) is bounded by

$$n^{-1}\mathbb{E}[\varepsilon(N_n - n)^2 + c|N_n - n|],$$

which is bounded by 2ε for n large enough. By the estimate (2.6) and Cauchy–Schwarz, there is a constant β_5 such that the second term in (4.6) is bounded by $\beta_5 n^{\beta_5}(P[|N_n - n| > n^{3/4}])^{1/2}$, which tends to zero. This completes the proof of (4.5).

We prove convergence of $n^{-1}\operatorname{Var}(H_n)$. This follows from the identity

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

In the right-hand side, the third term has variance tending to zero by (4.5), while the second term has variance α^2 and is independent of the first term. It follows that with σ^2 given by Theorem 3.1,

$$\sigma^2 = \lim_{n\to\infty} n^{-1}\operatorname{Var}(H'_n) = \lim_{n\to\infty}(n^{-1}\operatorname{Var}(H_n)) + \alpha^2,$$

so that $\sigma^2 \geq \alpha^2$ and $n^{-1}\operatorname{Var}(H_n) \to \tau^2$, where we set $\tau^2 = \sigma^2 - \alpha^2$. This gives us (2.3).

Theorem 3.1 tells us that $n^{-1/2}(H'_n - \mathbb{E}H'_n) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2)$. Combined with (4.5) this gives us

$$n^{-1/2}(H_n - \mathbb{E}H'_n + (N_n - n)\alpha) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2),$$

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

$$n^{-1/2}(H_n - \mathbb{E}H'_n) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2 - \alpha^2). \tag{4.7}$$

By (4.5), the expectation of $n^{-1/2}(H'_n - H_n - (N_n - n)\alpha)$ tends to zero, so in (4.7) we can replace $\mathbb{E}H'_n$ by $\mathbb{E}H_n$, which gives us (2.4). Save for showing that τ^2 is strictly positive, this completes the proof of Theorem 2.1.

5. A lower bound for the limiting variance. In this section we complete the proof of Theorem 2.1 by showing that the limiting variance τ^2 of $n^{-1/2}H(\mathscr{U}_{n,n})$ is nonzero whenever the distribution of $\Delta(\infty)$ is nondegenerate. Set $\alpha = \mathbb{E}[\Delta(\infty)]$, and using nondegeneracy, take $\delta > 0$ such that

$$P[\Delta(\infty) > \alpha + 4\delta] > 4\delta.$$

In the following we think of n as "fixed" but large enough for various estimates to hold. We construct a martingale in a manner different from that in Section 3. Let $\mathscr{F}_0$ be the trivial σ-field, let $\mathscr{F}_i = \sigma(U_{1,n}, \dots, U_{i,n})$ and write $\mathbb{E}_i$ for conditional expectation given $\mathscr{F}_i$. Define martingale differences $D_i = \mathbb{E}_i H(\mathscr{U}_{n,n}) - \mathbb{E}_{i-1}H(\mathscr{U}_{n,n})$. Then $H(\mathscr{U}_{n,n}) - \mathbb{E}H(\mathscr{U}_{n,n}) = \sum_{i=1}^n D_i$, and by orthogonality of martingale differences,

$$\operatorname{Var} H(\mathscr{U}_{n,n}) = \sum_{i=1}^n E[D_i^2]. \tag{5.1}$$

We look for lower bounds for $E[D_i^2]$. Given $i \leq m$, let $G_{i,m} = H(\mathscr{U}_{m,n}) - H(\mathscr{U}_{m,n}\setminus\{U_{i,n}\})$, the "contribution of $U_{i,n}$ to $H(\mathscr{U}_{m,n})$". Let $\widetilde{G}_{i,m} = H(\mathscr{U}_{m+1,n}\setminus \{U_{i,n}\}) - H(\mathscr{U}_{m,n}\setminus\{U_{i,n}\})$. Then $D_i = \mathbb{E}_i[G_{i,n} - \widetilde{G}_{i,n}]$.

We start by looking at $G_{i,n}$. We approximate it by $G_{i,i}$ which is a good approximation when i is close to n. By the coupling lemma (Lemma 4.2), we can find $\varepsilon_1 > 0$ such that if $i > (1-\varepsilon_1)n$ and n is sufficiently large, then

$$P[G_{i,i} > \alpha + 3\delta] > 3\delta.$$

Let $\eta > 0$ (to be given later on) and choose $\varepsilon_2 \in (0, \varepsilon_1)$ so that if $i > (1-\varepsilon_2)n$, then $P[G_{i,n} \neq G_{i,i}] < \eta$. Then for $i > n(1-\varepsilon_2)$,

$$\mathbb{E}[|G_{i,n} - G_{i,i}|] \leq \eta^{1/2}\mathbb{E}[(G_{i,n} - G_{i,i})^2]^{1/2} \leq c\eta^{1/2},$$

by the uniform bounded moments assumption. Provided η is small enough, this is less than δ^2. Then by Markov's inequality,

$$P[\mathbb{E}_i[|G_{i,n} - G_{i,i}|] > \delta] \leq \delta^{-1}\mathbb{E}[|G_{i,n} - G_{i,i}|] \leq \delta.$$

Since $\mathbb{E}_i G_{i,n} = G_{i,i} + \mathbb{E}_i[G_{i,n} - G_{i,i}]$, it follows that

$$P[\mathbb{E}_i G_{i,n} \geq \alpha + 2\delta] > 2\delta. \tag{5.2}$$

Define $f\colon \mathbb{R} \to \mathbb{R}$ by $f(x) = 0$ for $x \leq \alpha + \delta$ and $f(x) = 1$ for $x \geq \alpha + 2\delta$, interpolating linearly between $\alpha + \delta$ and $\alpha + 2\delta$. Set $Y_i = f(\mathbb{E}_i G_{i,n})$. Then (5.2) implies that

$$\mathbb{E}[(G_{i,n} - \alpha)Y_i] = \mathbb{E}[Y_i \mathbb{E}_i(G_{i,n} - \alpha)] \geq 4\delta^2. \tag{5.3}$$

Next consider $\mathbb{E}[(\widetilde{G}_{i,n} - \alpha)Y_i]$, writing the second factor as the sum of $f(G_{i,i})$ and $f(\mathbb{E}_i G_{i,n}) - f(G_{i,i})$. We claim that there exists $\varepsilon_3 \in (0, \varepsilon_2)$ such that provided n is sufficiently large and $i > (1-\varepsilon_3)n$,

$$\mathbb{E}[(\widetilde{G}_{i,n} - \alpha)f(G_{i,i})] \leq \delta^2. \tag{5.4}$$

This is because the two factors $\widetilde{G}_{i,n} - \alpha$ and $f(G_{i,i})$ are almost independent and the first of them has mean close to zero. More formally, it is proved as follows. By the proof of our coupling lemma (Lemma 4.2), provided ε_3 is sufficiently small we can take coupled variables R, R', D, D' such that D and D' are independent and each have the same distribution as $\Delta(\infty)$, such that R and R' have the same joint distribution as $\widetilde{G}_{i,n}$ and $G_{i,i}$, and such that $P[(R, R') \neq (D, D')]$ is small. Then

$$\begin{aligned}\mathbb{E}[(\widetilde{G}_{i,n} - \alpha)f(G_{i,i})] &= \mathbb{E}[(R-\alpha)f(R')] \\ &= \mathbb{E}[(D-\alpha)f(D')] - \mathbb{E}[(D-\alpha)f(D')\mathbf{1}\{(R,R')\neq(D,D')\}] \\ &\quad + \mathbb{E}[(R-\alpha)f(R')\mathbf{1}\{(R,R')\neq(D,D')\}];\end{aligned}$$

the first of the three terms in the right side is zero, while by Cauchy–Schwarz and the Poisson and uniform bounded moments conditions, the second and third terms are bounded by a constant times the square root of $P[(R, R') \neq (D, D')]$. This is less than δ^2 for an appropriate choice of ε_3, justifying the claim (5.4).

Since f' is bounded, and $G_{i,i}$ is $\mathscr{F}_i$-measurable, there is a constant c such that

$$
\begin{aligned}
&|\mathbb{E}[(\widetilde{G}_{i,n}-\alpha)(f(\mathbb{E}_i G_{i,n})-f(G_{i,i}))]| \\
&\quad \le \mathbb{E}[(\widetilde{G}_{i,n}-\alpha)^2]^{1/2}\mathbb{E}[(f(\mathbb{E}_i G_{i,n})-f(G_{i,i}))^2]^{1/2} \le c\mathbb{E}[(\mathbb{E}_i G_{i,n}-G_{i,i})^2]^{1/2} \\
&\quad = c\mathbb{E}[(\mathbb{E}_i[G_{i,n}-G_{i,i}])^2]^{1/2} \le c\mathbb{E}[(G_{i,n}-G_{i,i})^2]^{1/2}.
\end{aligned}
$$

However,

$$\mathbb{E}[(G_{i,n}-G_{i,i})^2] \le \mathbb{E}[(G_{i,n}-G_{i,i})^4]^{1/2} P[G_{i,n} \ne G_{i,i}]^{1/2} \le c'\eta^{1/2},$$

and provided η was well chosen it follows that

$$|\mathbb{E}[(\widetilde{G}_{i,n}-\alpha)(f(\mathbb{E}_i G_{i,n})-f(G_{i,i}))]| \le \delta^2.$$

Combining this with (5.4), for n large, we have

$$
\begin{aligned}
\mathbb{E}[(\widetilde{G}_{i,n}-\alpha)Y_i] &= \mathbb{E}[(\widetilde{G}_{i,n}-\alpha)f(G_{i,i})] \\
&\quad + \mathbb{E}[(\widetilde{G}_{i,n}-\alpha)(f(\mathbb{E}_i G_{i,n})-f(G_{i,i}))] \le 2\delta^2.
\end{aligned}
$$

Combined with (5.3) this implies that for large n and $i \ge (1-\varepsilon_3)n$, we have

$$\mathbb{E}[(G_{i,n}-\widetilde{G}_{i,n})Y_i] \ge 2\delta^2.$$

Hence, using the fact that Y_i is $\mathscr{F}_i$-measurable and lies in the range [0, 1], we obtain

$$2\delta^2 \le \mathbb{E}[Y_i \mathbb{E}_i(G_{i,n}-\widetilde{G}_{i,n})] \le \mathbb{E}[|\mathbb{E}_i(G_{i,n}-\widetilde{G}_{i,n})|] = \mathbb{E}[|D_i|],$$

and hence, $\mathbb{E}[D_i^2] \ge \mathbb{E}[|D_i|]^2 \ge 4\delta^4$. Thus, using (5.1), we have $\operatorname{Var} H(\mathscr{U}_{n,n}) \ge (\varepsilon_3 n - 1)(4\delta^4)$, and therefore $\tau^2 > 0$ by (2.3).

6. The k-nearest neighbors graph. Fix $k \in \mathbb{N}$ and $d \ge 1$ and let $\mathrm{NG}(\mathscr{X})$ denote the k-nearest neighbors graph on a point set $\mathscr{X} \subset \mathbb{R}^d$. Here we show that the total edge length and the number of components of the k-nearest neighbors graph on points in $\mathbb{R}^d$ satisfy strong stabilization as well as the bounded moments conditions. Using similar methods one can show that other functionals, such as the number of vertices of a fixed degree in the k-nearest neighbors graph, the number of vertices which are the nearest neighbors of exactly k other points and the number of vertices which are the lth nearest neighbors to their own kth nearest neighbors all satisfy the central limit behavior of Theorem 2.1 and Corollary 2.1. We will leave the details of these other applications to the reader. Laws of large numbers for the latter two functionals have been obtained by Henze [9].

A variant of $\mathrm{NG}(\mathscr{X})$ which has also been considered in the literature is the *directed* graph $\mathrm{NG}'(\mathscr{X})$, formed by inserting a directed edge (x, y) whenever y is one of the k nearest neighbors of x. Thus, for example, the total length of $\mathrm{NG}'(\mathscr{X})$ counts some of the edges of $\mathrm{NG}(\mathscr{X})$ twice. It should be possible to modify proofs of our CLTs for $\mathrm{NG}(\mathscr{X})$ to give analogous CLTs for $\mathrm{NG}'(\mathscr{X})$.

Throughout this section we assume $\lambda = 1$.

6.1. *Total edge length.* In this section, $H(\mathscr{X})$ denotes the total edge length of $\mathrm{NG}(\mathscr{X})$. Bickel and Breiman [2] prove, among other things, a CLT for the total edge length of $\mathrm{NG}'(\mathscr{X}_n)$, in the case $k = 1$; Avram and Bertsimas [1] prove a CLT for the total edge length of $\mathrm{NG}'(\mathscr{P}_n)$, in the case where all the sets B_n are cubes, but do not address the convergence of the variance. The following CLT extends these results.

THEOREM 6.1 (CLT for total edge length of the k-nearest neighbors graph). *There exists $\sigma^2 > 0$ such that provided $\mathscr{B}$ is regular in the sense of* (2.5), *as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathscr{P}_n)) \to \sigma^2$ and*

$$n^{-1/2}(H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2). \tag{6.1}$$

Additionally, there exists $\tau^2 \in (0, \sigma^2]$ such that as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathscr{U}_{n,n}))$ $\to \tau^2$, and

$$n^{-1/2}(H(\mathscr{U}_{n,n}) - \mathbb{E}H(\mathscr{U}_{n,n})) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{6.2}$$

Moreover, if $\mathscr{B}_0$ is regular then $n^{(2/d)-1}\operatorname{Var}H(\mathscr{X}_n) \to \tau^2$ and

$$n^{(1/d)-1/2}(H(\mathscr{X}_n) - \mathbb{E}H(\mathscr{X}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{6.3}$$

To prove Theorem 6.1 it suffices to verify the conditions of Theorem 2.1 and Corollary 2.1. We do this in the remainder of this section. Note first that H is homogeneous of order 1. Also, H is polynomially bounded because $H(\mathscr{X}) \leq k\operatorname{diam}(\mathscr{X})\operatorname{card}(\mathscr{X})$.

LEMMA 6.1. *H is strongly stabilizing.*

PROOF. For simplicity we prove strong stabilization in dimension two, but the argument is easily extended to higher dimensions by using cones instead of triangles (for $d = 1$, take intervals instead of triangles). For each $t > 0$ construct six disjoint equilateral triangles $T_j(t), 1 \leq j \leq 6$, such that the origin is a vertex of each triangle, such that each triangle has edge length t and such that $T_j(t) \subset T_j(u)$ whenever $t < u$.

Given the homogeneous Poisson point process $\mathscr{P}$ of unit intensity on $\mathbb{R}^2$, let the random variable R be the minimum t such that each triangle $T_j(t), 1 \leq j \leq 6$, contains at least $k + 1$ points from $\mathscr{P}$. Then R is a.s. finite, since $T_j(\infty) := \cup_{t>0} T_j(t)$ contains infinitely many Poisson points a.s.

Inserting the origin $\{0\}$ into a point set can cause the addition of some edges and the removal of others. We claim that given the configuration of $\mathscr{P}$ in $B_{4R}(0)$, the set of added edges or removed edges is insensitive to the addition or removal of points outside $B_{4R}(0)$, and therefore $4R$ is a radius of stabilization for H.

To prove the claim, note first that the k nearest neighbors of the origin all lie in $\cup_{j=1}^{6} T_j(R)$. Moreover, if a point at x has the origin as one of its k nearest neighbors, then x must lie in one of the six triangles $T_j(R)$; if not it would lie in some trapezoid $T_j(t) \backslash T_j(R)$, and there would be k points in

$T_j(R)$ lying closer to x than the origin does. Moreover, if x lies in $T_j(R)$ then its k nearest neighbors lie within a distance R and are unaffected by changes outside $B_{2R}(0)$. This shows that the set of added edges is insensitive to changes outside $B_{2R}(0)$.

Next consider removed edges. All edges removed as a result of the insertion of a point at the origin are of the form $\{x, y\}$ with x having 0 as one of its k nearest neighbors and having y as its $(k+1)$st nearest neighbor. As already seen, x must lie in one of the sets $T_j(R)$, and $|y - x| \leq R$ so that $y \in B_{2R}(0)$. Then the edge $\{x, y\}$ is indeed removed unless x is one of the k nearest neighbors of y. However, y has at least k points within a distance $2R$ of it, and so the decision on whether to remove edge $\{x, y\}$ is unaffected by changes outside $B_{4R}(0)$, points outside $B_{4R}(0)$ all lying at a distance greater than $2R$ from y. This proves the claim. □

LEMMA 6.2. *If $\mathscr{B}$ is regular, then H satisfies the uniform bounded moments condition on $\mathscr{B}$.*

PROOF. Let $A \in \mathscr{B}$ with $0 \in A$, and let $|A|/2 \leq m \leq 3|A|/2$. Let the m independent random points comprising $\mathscr{U}_{m,A}$ be denoted $V_1, \ldots, V_m$. Let $L(0)$ be the total length of the edges incident to 0 in $\mathrm{NG}(\mathscr{U}_{m,A} \cup \{0\})$; this is an upper bound for $\Delta(\mathscr{U}_{m,A})^+$, the positive part of $\Delta(\mathscr{U}_{m,A})$. Let $L_{\max}$ be the maximum of these edge lengths, and let $\mathrm{Deg}(0)$ be the degree of the vertex at the origin. Then

$$L(0)^4 \leq \mathrm{Deg}(0)^4 L_{\max}^4 \leq \sum_{i=1}^{m} W_i^4,$$

where W_i denotes the product of $|V_i|$ with the number of points of $\mathscr{U}_{m,A}$ in $B_{|V_i|}(0)$ (including V_i itself) times the indicator of the event that $\{0, V_i\}$ is an edge. Therefore,

$$\mathbb{E}[L(0)^4] \leq m\mathbb{E}[W_1^4] = m \int_A |u|^4 \mathbb{E}[(N_u + 1)^4 \mathbf{1}\{E_u\}] \frac{du}{|A|},$$

where N_u is the number of points of $\mathscr{U}_{m-1,A}$ in $B_{|u|}(0)$ and where E_u is the event that u is joined to 0 in the k-nearest neighbors graph on $\mathscr{U}_{m-1,A} \cup \{0, u\}$. By Cauchy–Schwarz and the fact that $m \leq 2|A|$ by assumption,

$$\mathbb{E}[L(0)^4] \leq 2 \int_A |u|^4 (\mathbb{E}[(N_u + 1)^8])^{1/2} P[E_u]^{1/2}\, du.$$

The mean of N_u is bounded by a constant times $|u|^d$, so by a standard estimate on the binomial distribution, its eighth moment is bounded by a constant times $|u|^{8d}$. Also, E_u happens only if the ball $B_{|u|}(u)$ has at most $k - 1$ points or $B_{|u|}(0)$ has at most $k - 1$ points. Let θ denote the volume of the unit ball.

For $|A|/2 \le m \le 3|A|/2$, and $|u| \ge 1$, we have by regularity (2.5),

$$P[E_u] \le 2 \sum_{j=0}^{k-1} \binom{m-1}{j} \left(\frac{\theta |u|^d}{|A|} \right)^j \left(1 - \frac{\delta |u|^\delta}{|A|} \right)^{m-1-j} \tag{6.4}$$
$$\le c|u|^{d(k-1)} \exp(-(\delta/4)|u|^\delta).$$

This shows that $E[L(0)^4]$ is uniformly bounded by a constant times

$$\int_{\mathbb{R}^d} |u|^{4+4d+d(k-1)} \exp(-(\delta/4)|u|^\delta)\, du,$$

which is finite. Hence $\Delta(\mathscr{U}_{m,A})^+$ has a uniformly bounded fourth moment.

Now consider the fourth moment of $\Delta(\mathscr{U}_{m,A})^-$. Write $V_i \to 0$ if 0 is one of the k nearest neighbors of V_i in the point process $\mathscr{U}_{m,A} \cup \{0\}$. Also, let L_i be the total length of all edges incident to V_i in $\mathrm{NG}(\mathscr{U}_{m,A})$. Then $\Delta(\mathscr{U}_{m,A})^-$ is bounded by the total length of deleted edges, and so is at most $\sum_{i=1}^m L_i \mathbf{1}\{V_i \to 0\}$. Since the number of nonzero terms in this sum is bounded by a geometric constant $C(d, k)$ (see [23], page 102), it follows that there is a constant c such that

$$(\Delta(\mathscr{U}_{m,A})^-)^4 \le c \sum_{i=1}^m L_i^4 \mathbf{1}\{V_i \to 0\}.$$

Taking expectations and using Cauchy–Schwarz yields

$$\mathbb{E}[(\Delta(\mathscr{U}_{m,A})^-)^4] \le mc \int_A (\mathbb{E}L(x)^8)^{1/2} P[x \to 0]^{1/2} \left(\frac{dx}{|A|} \right).$$

Here $L(x)$ is the total length of edges incident to x in $\mathrm{NG}(\{x\} \cup \mathscr{U}_{m-1,A})$, which has a bounded eighth moment by a similar argument to the above proof that $L(0)$ has bounded fourth moment. Also, by the same argument as for (6.4), there are constants c, δ such that $P[x \to 0] \le c|x|^{d(k-1)} \exp(-\delta |x|^\delta)$ for all $x \in A$. Hence, $E[(\Delta(\mathscr{U}_{m,A})^-)^4]$ is bounded uniformly in A, m. This demonstrates the uniform moments condition.

LEMMA 6.3. *The distribution of* $\Delta(\infty)$ *is nondegenerate.*

PROOF. Let $C_0 = Q_{1/2}(0)$, the unit cube centered at the origin. The annulus $Q_{d+1}(0) \backslash C_0$ will be called the moat. Partition the annulus $Q_{d+2}(0) \backslash Q_{d+1}(0)$ into a finite collection $\mathscr{A}$ of unit cubes. Now define the following events. Let E_2 be the event that there are no points in $\mathscr{P}$ in the moat and there are at least $k+1$ points in each of the unit subcubes in $\mathscr{A}$. Let E_1 be the intersection of E_2 and the event that there are k points in C_0; let E_0 be the intersection of E_2 and the event that there are no points in C_0. Then E_0 and E_1 have strictly positive probability.

Now we notice that if E_0 occurs, then adding the origin creates k new edges and has no other effect. It increases the total edge length by at least kW, where $W = d + 1/2$ is the width of the moat. If E_1 occurs, then before adding the origin there are k edges crossing the moat. Adding the origin destroys these edges and reduces the total edge length by at least $k(W-w)$, where $w=\sqrt{d}$ is the diameter of C_0.

Thus E_0 and E_1 are events with strictly positive probability which give rise to values of $\Delta(\infty)$ which differ by at least $k(W-w)$, a fixed amount. This demonstrates the nondegeneracy of $\Delta(\infty)$. □

Thus, H satisfies the conditions for Theorem 2.1 and Corollary 2.1, so Theorem 6.1 is proved.

6.2. *Number of components.* In this section we let $H(\mathscr{X})$ be the number of components in $\mathrm{NG}(\mathscr{X})$.

THEOREM 6.2 (CLT for the number of components of the k-nearest neighbors graph). *There exists $\sigma^2>0$ such that as $n\to\infty$, $n^{-1}\operatorname{Var}(H(\mathscr{P}_n))\to\sigma^2$ and*

$$n^{-1/2}(H(\mathscr{P}_n)-\mathbb{E}H(\mathscr{P}_n))\xrightarrow{\mathscr{D}}\mathscr{N}(0,\sigma^2). \tag{6.5}$$

Additionally, there exists $\tau^2\in(0,\sigma^2]$ such that as $n\to\infty$, $n^{-1}\operatorname{Var}(H(\mathscr{U}_{n,n}))\to\tau^2$ and

$$n^{-1/2}(H(\mathscr{U}_{n,n})-\mathbb{E}H(\mathscr{U}_{n,n}))\xrightarrow{\mathscr{D}}\mathscr{N}(0,\tau^2). \tag{6.6}$$

Also, $n^{-1}\operatorname{Var}(H(\mathscr{X}_n))\to\tau^2$ and

$$n^{-1/2}(H(\mathscr{X}_n)-\mathbb{E}H(\mathscr{X}_n))\xrightarrow{\mathscr{D}}\mathscr{N}(0,\tau^2). \tag{6.7}$$

Note that in this theorem, there is no requirement that $\mathscr{B}$ or $\mathscr{B}_0$ be regular. We prove it by showing that H satisfies the conditions of Theorem 2.1 and Corollary 2.1. First note that H is homogeneous of order 0, and H is polynomially bounded because $H(\mathscr{X})\le\operatorname{card}(\mathscr{X})$.

It is quite simple to see that H satisfies the uniform bounded moments condition. Recall first that the degree of the vertices of the k-nearest neighbors graph is uniformly bounded by some constant $C(d,k)$ [23]. Thus, inserting an extra point into a given set of points causes the addition of at most $C(d,k)$ edges and cannot decrease the number of components by more than $C(d,k)$. Moreover, the number of edges removed due to the insertion of an extra point is at most one for each point having the inserted point as one of its k nearest neighbors, and therefore is also bounded by $C(d,k)$. Each removed edge increases the number of components by at most one, so inserting a point cannot increase the number of components by more than $C(d,k)$. This demonstrates the uniform moments condition, with no extra requirement on $\mathscr{B}$.

The proof of strong stabilization is more involved than it was for the total length. In particular, it requires a result of [7] on "uniqueness of the infinite cluster" for the k-nearest neighbors graph on the infinite Poisson process $\mathscr{P}$.

LEMMA 6.4. (a) *With probability* 1, $\mathrm{NG}(\mathscr{P})$ *has at most one infinite component*.

(b) *With probability* 1, $\mathrm{NG}(\mathscr{P} \cup \{0\})$ *has at most one infinite component*.

PROOF. Part (a) is Theorem 4.1 of [7]. For part (b), let E_ε be the event that there is a single point of $\mathscr{P}$ in $B_\varepsilon(0)$. Let A be the event that there is at most one infinite component in $\mathrm{NG}(\mathscr{P})$. Then $P[A|E_\varepsilon] = 1$ for all $\varepsilon > 0$, by part (a). Letting ε tend to zero gives the result. □

We now want to show that H satisfies strong stabilization. As before, we prove this only for the case $d = 2$. We prepare for the proof of stabilization with one final lemma. Recall the definitions of the six triangles $T_j(t)$ used in the proof of Lemma 6.1. For each point X of $\mathscr{P}$ let $R(X)$ be the minimum t such that each of the translated triangles $T_j(t) + X$, $1 \le j \le 6$, contains at least $k+1$ points (not counting X itself) of $\mathscr{P}$. The significance of $R(X)$ is that changes in the configuration outside $B_{2R(X)}(X)$ cannot have any effect on the set of edges of the k-nearest neighbors graph incident to X. This is seen by a similar argument to the proof of Lemma 6.1.

LEMMA 6.5. *With probability* 1, *the values of* $R(X)$ *are finite for all points* X *of* $\mathscr{P}$.

PROOF. It suffices to prove the result for all points $X \in \mathscr{P} \cap Q_{1/2}(0)$. By integrating out the respective positions of the points in $Q_{1/2}(0)$, one sees that it suffices to prove that for any finite subset $\{x_1, \dots, x_m\}$ of $Q_{1/2}(0)$, the triangles $T_j(\infty) + x_i$, $1 \le i \le m$, $1 \le j \le 6$, all have infinitely many points of $\mathscr{P} \setminus Q_{1/2}(0)$ a.s. However, this is clearly true. □

PROPOSITION 6.1. *H satisfies strong stabilization and the distribution of* $\Delta(\infty)$ *is nondegenerate*.

PROOF. The insertion of the origin causes the addition of certain edges and the deletion of other edges. As seen in the proof of Lemma 6.1, all such edges lie within $B_{2R}(0)$ and the set of added edges and the set of deleted edges are insensitive to changes outside $B_{4R}(0)$. We need to prove that the effect of these additions and deletions on the number of components is insensitive to changes in the configuration outside a certain range.

Let $\mathscr{C}$ be the set of components of $\mathrm{NG}(\mathscr{P})$ that include one or more vertices lying in $B_{4R}(0)$. At most one of these components is infinite, by Lemma 6.4. Choose a finite $L_1 > 4R$ such that $B_{L_1}(0)$ contains all the *finite* components in the collection $\mathscr{C}$. For every pair of points X, Y of $\mathscr{P} \cap B_{4R}(0)$ which lie in the infinite component of $\mathrm{NG}(\mathscr{P})$ (if there is one), there is a path in $\mathrm{NG}(\mathscr{P})$ connecting X to Y. Since there are a.s. only finitely many points of $\mathscr{P}$ in $B_{4R}(0)$, we can a.s. find a finite $L_2 > L_1$ such that for *every* pair of points X, Y in $B_{4R}(0)$ which are also in the infinite component of $\mathrm{NG}(\mathscr{P})$, there is a path from X to Y staying within the ball $B_{L_2}(0)$.

The upshot of the above argument is that for every pair X, Y of points of $\mathscr{P} \cap B_{4R}(0)$, *either* there is a path from X to Y in $\mathrm{NG}(\mathscr{P})$ that stays within $B_{L_2}(0)$, *or* at least one member of $\{X, Y\}$ lies in a finite component of $\mathrm{NG}(\mathscr{P})$ that is contained within $B_{L_2}(0)$ and does not include the other member of $\{X, Y\}$.

Using part (b) of Lemma 6.4, we can similarly a.s. find a finite $L_3 > L_2$, such that for every pair X, Y of points of $(\mathscr{P} \cup \{0\}) \cap B_{4R}(0)$, *either* there is a path from X to Y in $\mathrm{NG}(\mathscr{P} \cup \{0\})$ that stays within $B_{L_3}(0)$, *or* at least one member of $\{X, Y\}$ lies in a finite component of $\mathrm{NG}(\mathscr{P} \cup \{0\})$ that is contained within $B_{L_3}(0)$ and does not include the other member of $\{X, Y\}$. Thus, whether there is a path from X to Y is determined entirely by edges involving points in $B_{L_3}(0)$.

The number of Poisson points in $B_{L_3}(0)$ is a.s. finite. By Lemma 6.5, the balls $B_{2R(X)}(X)$, $X \in \mathscr{P} \cap B_{L_3}(0)$, all have finite radius. Take a finite $L_4 > L_3$, large enough for $B_{L_4}(0)$ to contain all of these balls. Our claim is that L_4 is a radius of stabilization for H. This is because changes outside $B_{L_4}(0)$ do not create or destroy any edges having at least one endpoint within $B_{L_3}(0)$, and therefore do not affect the question of whether there is a path from X to Y, for all X, Y in $B_{4R}(0)$. Let us now explicitly justify the claim.

Let $e_1, \dots, e_i$ be the set of added edges as a result of the insertion of the origin. Let $\mathscr{C}_1$ be the set of components of $\mathrm{NG}(\mathscr{P})$ in the collection $\mathscr{C}$ which are connected to the origin by one or more of the edges $e_1, \dots, e_i$. Then, regardless of what happens outside $B_{L_4}(0)$, the increment in the number of components, due to the addition of a vertex at 0 and edges $e_1, \dots, e_i$, is equal to $1 - \mathrm{card}(\mathscr{C}_1)$.

Let $f_1, \dots, f_j$ be the set of edges deleted as a result of the insertion of the origin. After deleting these edges, having previously added the edges $e_1, \dots, e_i$, one ends up with the k-nearest neighbors graph on the point set with $\{0\}$ inserted. We consider instead the reverse process, in which we start with the k-nearest neighbors graph on the point set with $\{0\}$ inserted, and then remove edges $f_1, \dots, f_j$.

Let $\mathscr{C}'$ be the set of components of $\mathrm{NG}(\mathscr{P} \cup \{0\})$ that include one or more vertices lying in $B_{4R}(0)$. The edges $f_1, \dots, f_j$ (which are *not* edges of this graph) induce an adjacency relation on the set $\mathscr{C}'$, two elements of $\mathscr{C}'$ being deemed adjacent if one or more of the $f_1, \dots, f_j$ connects them together. If ν_1 denotes the cardinality of $\mathscr{C}'$, and ν_2 denotes the number of components of the graph with vertex set $\mathscr{C}'$ and adjacency as just described, then the increment in the number of components due to adding edges $f_1, \dots, f_j$ is precisely equal to $\nu_2 - \nu_1$.

Combining the above arguments, we see that the increment in the number of components, due to the insertion of $\{0\}$, is equal to $1 - \mathrm{card}(\mathscr{C}) + \nu_1 - \nu_2$, regardless of what happens outside $B_{L_4}(0)$. Thus we have established strong stabilization, and $\Delta(\infty) = 1 - \mathrm{card}(\mathscr{C}) + \nu_1 - \nu_2$.

Finally, let us check the distribution of $\Delta(\infty)$. It suffices to consider the events E_0 and E_1 from Section 6.1 and to note that if E_0 occurs then there is no increase in the number of components, whereas if E_1 occurs then the number of components increases by at least one. This completes the proof of Proposition 6.1. □

The total number of components in the k-nearest neighbors graph thus satisfies the conditions of Theorem 2.1 and Corollary 2.1, and Theorem 6.2 is proved.

7. The sphere of influence graph. Fix $d \geq 1$ and let $\mathrm{SIG}(\mathscr{X})$ denote the SIG on a finite point set $\mathscr{X} \subset \mathbb{R}^d$. We show that the total number of edges, the total edge length, the number of vertices of fixed degree and the total number of components of the SIG on points in $\mathbb{R}^d$ all satisfy the CLT behavior of Theorem 2.1. Along the way we show uniqueness of the infinite component in the SIG on an infinite Poisson process, which is of interest in its own right.

Throughout this section we assume $\lambda = 1$.

7.1. *Total number of edges.* In this section, let $H(\mathscr{X})$ denote the number of edges in $\mathrm{SIG}(\mathscr{X})$. Füredi [5] has shown that $\mathbb{E}H(\mathscr{X}_n)/n$ converges to a limit as $n \to \infty$. We show that H satisfies the following CLT. In this way we recover most of the results of [10], we show a de-Poissonized version of their central limit theorem and we show convergence of the variance of the total number of edges.

THEOREM 7.1 (CLT for the number of edges in the SIG). *Suppose $\mathscr{B}$ is regular. There exists $\sigma^2 > 0$ such that as $n \to \infty$, $n^{-1}\,\mathrm{Var}(H(\mathscr{P}_n)) \to \sigma^2$ and*

$$n^{-1/2}(H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2). \tag{7.1}$$

Additionally, there exists $\tau^2 \in (0, \sigma^2]$ such that as $n \to \infty$, $n^{-1}\,\mathrm{Var}(H(\mathscr{U}_{n,n})) \to \tau^2$ and

$$n^{-1/2}(H(\mathscr{U}_{n,n}) - \mathbb{E}H(\mathscr{U}_{n,n})) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{7.2}$$

Also, if $\mathscr{B}_0$ is regular, $n^{-1}\,\mathrm{Var}(H(\mathscr{X}_n)) \to \tau^2$, and

$$n^{-1/2}(H(\mathscr{X}_n) - \mathbb{E}H(\mathscr{X}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{7.3}$$

We will prove this result by showing that H satisfies the conditions of Theorem 2.1. Since $\mathrm{SIG}(\mathscr{X})$ is a subgraph of the complete graph on $\mathscr{X}$, it follows that H satisfies the growth bound $H(\mathscr{X}) \leq (\mathrm{card}(\mathscr{X}))^2$, and so is polynomially bounded. Also, H is homogeneous of order 0.

LEMMA 7.1. *H is strongly stabilizing.*

PROOF. Inserting a point at the origin creates new edges incident to the origin and may remove edges between points, but cannot create any new edges between two old points. Edges which are removed have a vertex which has the origin as a nearest neighbor. We will show that there is a random ball centered at the origin such that the set of added edges is a.s. determined by what happens inside the ball and likewise for the set of removed edges.

To see this, consider an infinite cone C with its vertex at the origin, subtending an angle of $\pi/6$ radians [for $d = 1$, take C to be the interval $(0, \infty)$]. Let R be the distance from 0 to its closest neighbor in $\mathscr{P} \cap C$, and let Y be the point in $C \backslash B_{6R}(0)$ closest to 0. Then Y exists a.s., because there are a.s. infinitely many Poisson points in C (but only finitely many in any bounded region).

Given $x \in C \backslash B_{|Y|}(0)$ let $x' = (|Y|/|x|)x$. By the triangle inequality and the fact that $\sin(\pi/6) = 1/2$,

$$|x - Y| \le |x - x'| + |x' - Y| \le (|x| - |Y|) + (|Y|/2) \le |x| - 3R. \tag{7.4}$$

Therefore, the sphere of influence of x does not reach the interior of $B_{3R}(0)$. Hence, no points in C at a distance more than $|Y|$ from the origin will be connected to 0. Therefore, the configuration of points outside $B_{3|Y|}(0)$ has no effect on the set of points in C connected to 0.

Now consider a finite number of cones $C_1, \dots, C_m$ congruent to C, each with a vertex at 0 and with union $\mathbb{R}^d$. By the above, there is a.s. a finite random number S_1, which is equal to the maximum of m identically distributed copies of $3|Y|$, such that adding or removing points further than S from 0 does not affect the set of added edges. This gives strong stabilization of H for added edges.

Now we show that H is strongly stabilizing for deleted edges. We will follow the cone argument described above. For $1 \le i \le m$ let x_i be the closest Poisson point in C_i to the origin. The set of vertices having the origin as a nearest neighbor is a subset of $\{x_1, \dots, x_m\}$. Let R'_i be the distance from 0 to its *second* nearest neighbor in the cone C_i and let $R' = \max(R_1, \dots, R_m)$. Note that for $1 \le i \le m$, the sphere of influence of x_i (before the addition of the origin) is contained in $B_{2R'}(0)$.

Let Y'_i be the point in $C_i \backslash B_{6R'}(0)$ closest to 0. By a similar argument to (7.4), for any point in C_i further out than Y'_i the sphere of influence does not reach $B_{3R'}(0)$, and so does not meet any of the spheres of influence of the points x_j, $1 \le j \le m$. Thus, if we set $S_2 = \max_{1 \le i \le m} |Y'_i|$, only edges involving points inside $B_{S_2}(0)$ are possibly deleted as a result of inserting $\{0\}$, and moreover these points all lie within a distance at most $S_2 + R'$ of the points x_i, so their spheres of influence are unaffected by changes outside $B_{3S_2}(0)$. Therefore $3S_2$ serves as a radius of stabilization for deleted edges. Combined with the earlier argument for added edges, this shows that H is strongly stabilizing. $\square$

LEMMA 7.2. *If $\mathscr{B}$ is regular, then H satisfies the uniform bounded moments condition on $\mathscr{B}$.*

PROOF. The proof is quite similar to that of Lemma 6.2. Let $A \in \mathscr{B}$ with $0 \in A$, and assume $|A|/2 \le m \le 3|A|/2$. Let the independent random points comprising $\mathscr{U}_{m,A}$ be denoted $V_1, \dots, V_m$.

Consider first the positive part $\Delta(\mathscr{U}_{m,A})^+$ of $\Delta(\mathscr{U}_{m,A})$. Let $D(0)$ denote the degree of zero in $\mathrm{SIG}(\mathscr{U}_{m,A} \cup \{0\})$. Inserting a point causes some of the existing spheres of influence to shrink and leaves the others unchanged and so does not create any new edges of the SIG except for those incident to the inserted point itself. Therefore $\Delta(\mathscr{U}_{m,A})^+ \le D(0)$.

Since one of the points V_i must be the furthest out from the origin among those adjacent to it in the SIG, we have $D(0)^4 \le \sum_{i=1}^m W_i^4$, where W_i denotes the product of the number of points of $\mathscr{U}_{m,A}$ in $B_{|V_i|}(0)$ (including V_i itself), and the indicator of the event that $\{0, V_i\}$ is an edge. Therefore,

$$\mathbb{E}[D(0)^4] \le m\mathbb{E}[W_1^4] = m\int_A \mathbb{E}[(N_u+1)^4 \mathbf{1}\{E_u\}]\frac{du}{|A|}, \tag{7.5}$$

where N_u is the number of points of $\mathscr{U}_{m-1,A}$ in $B_{|u|}(0)$ and where E_u is the event that u is joined to 0 in the SIG on $\mathscr{U}_{m-1,A} \cup \{0, u\}$. By Cauchy–Schwarz and the fact that $m \le 2|A|$ by assumption,

$$\mathbb{E}[D(0)^4] \le 2\int_A (\mathbb{E}[(N_u+1)^8])^{1/2} P[E_u]^{1/2}\, du.$$

The eighth moment of N_u is bounded by a constant times $|u|^{8d}$. Also, E_u happens only if the ball $B_{|u|/4}(u)$ contains no points or $B_{|u|/4}(0)$ contains no points. For $|u| \ge 1$, regularity (2.5) yields

$$P[E_u] \le 2\left(1 - \frac{(\delta|u|/4)^\delta}{|A|}\right)^{m-2} \le c\exp(-(\delta/4^{1+\delta})|u|^\delta).$$

Combining these estimates gives us a uniform bound for the fourth moment of $E[D(0)^4]$ and hence for that of $\Delta(\mathscr{U}_{m,A})^+$.

Now consider the fourth moment of $\Delta(\mathscr{U}_{m,A})^-$. Write $V_i \to 0$ if 0 is the nearest neighbor of V_i in the point process $\mathscr{U}_{m,A} \cup \{0\}$, and let D_i denote the degree of V_i in $\mathrm{SIG}(\mathscr{U}_{m,A})$.

Inserting a point at the origin causes the sphere of inference of V_i to shrink only if $V_i \to 0$, and therefore causes the possible deletion of an existing edge $\{V_i, V_j\}$ of $\mathrm{SIG}(\mathscr{U}_{m,A})$ only if either $V_i \to 0$ or $V_j \to 0$ (or both). Therefore $\Delta(\mathscr{U}_{m,A})^- \le \sum_{i=1}^m D_i \mathbf{1}\{V_i \to 0\}$. Since the number of nonzero terms in this sum is bounded by a geometric constant $C(d,k)$ ([23], page 102), it follows that there is a constant c such that $(\Delta(\mathscr{U}_{m,A})^-)^4 \le c\sum_{i=1}^m D_i^4 \mathbf{1}\{V_i \to 0\}$. Taking expectations and using Cauchy–Schwarz yields

$$\mathbb{E}[(\Delta(\mathscr{U}_{m,A})^-)^4] \le mc\int_A (\mathbb{E}D(x)^8)^{1/2} P[x \to 0]^{1/2}\left(\frac{dx}{|A|}\right).$$

Here $D(x)$ is degree of x in $\mathrm{SIG}(\{x\} \cup \mathscr{U}_{m-1,A})$, which has a bounded eighth moment by a similar argument to the above proof that $D(0)$ has bounded fourth moment. Also, by regularity, there are constants c, δ such that $P[x \to 0] \le c\exp(-\delta|x|^\delta)$ for all $x \in A$. Hence, $E[(\Delta(\mathscr{U}_{m,A})^-)^4]$ is bounded uniformly in A. This shows the uniform moments condition. □

LEMMA 7.3. *The distribution of* $\Delta(\infty)$ *is nondegenerate.*

PROOF. We will use a construction similar to that used for the k-nearest neighbors graph. Let E_2 be the event that the moat is empty and that there are two points of $\mathscr{P}$ in each of the unit subcubes in $\mathscr{A}$. Let E_0 be the intersection of E_2 and the event that there is no point of $\mathscr{P}$ in C_0. Let E_1 be the intersection of E_2 and the event that there is one point of $\mathscr{P}$ in the ball $B_{1/10}(1/4, 0, \dots, 0)$, and there are no other points in C_0.

Each of E_0 and E_1 have positive probability. Inserting the origin when E_0 happens creates at least one additional edge across the moat and does not destroy any edges. Inserting the origin when E_1 happens creates a single new edge inside C_0 and destroys at least one edge across the moat. So the events E_0 and E_1 have strictly positive probability and give rise to values of $\Delta(\infty)$ which differ by at least 1. This shows nondegeneracy of $\Delta(\infty)$. □

We have shown that H satisfies the conditions of Theorem 2.1 and Corollary 2.1. Together, these results give us Theorem 7.1.

7.2. *Total edge length.* In this section, $H(\mathscr{X})$ denotes the total edge length of $\mathrm{SIG}(\mathscr{X})$.

THEOREM 7.2 (CLT for the total edge length in the SIG). *Suppose that $\mathscr{B}$ is regular. Then there exists $\sigma^2 > 0$ such that as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathscr{P}_n)) \to \sigma^2$ and*

$$n^{-1/2}(H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2). \tag{7.6}$$

Additionally, there exists $\tau^2 \in (0, \sigma^2]$ such that as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathscr{U}_{n,n})) \to \tau^2$, and

$$n^{-1/2}(H(\mathscr{U}_{n,n}) - \mathbb{E}H(\mathscr{U}_{n,n})) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{7.7}$$

Also, if $\mathscr{B}_0$ is regular, $n^{(2/d)-1}\operatorname{Var}(H(\mathscr{X}_n)) \to \tau^2$, and

$$n^{(1/d)-1/2}(H(\mathscr{X}_n) - \mathbb{E}H(\mathscr{X}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{7.8}$$

PROOF. We verify the conditions of Theorem 2.1 and Corollary 2.1. Notice first that H is homogeneous of order 1. Also, H is polynomially bounded since $\mathrm{SIG}(\mathscr{X})$ is a subgraph of the complete graph on $\mathscr{X}$. Moreover, strong stabilization follows exactly as in the proof of Lemma 7.1.

If $\mathscr{B}$ is regular, then H satisfies the uniform bounded moments condition on $\mathscr{B}$. The proof of this is virtually identical to that of Lemma 7.2, except that the degree $D(0)$ in that argument should be replaced by the total length of edges incident to the origin, and similarly for the degrees D_i and $D(x)$ appearing later on in that proof. A factor of $|V_i|$ needs to be introduced into the definition of the variable W_i, and consequently a factor of $|u|^4$ comes into the integral in (7.5), but this does not invalidate the subsequent argument.

Finally, by using the same arguments as in Section 7.1 and noting that the event E_0 produces an increase of at least 1 in the total edge length of the SIG when the origin is added, while E_1 produces a reduction in the total edge length when the origin is added, we see that the distribution of $\Delta(\infty)$ is nondegenerate. □

7.3. *Number of vertices of fixed degree.* Let $H_k(\mathscr{X})$ denote the number of vertices of fixed degree k in $\mathrm{SIG}(\mathscr{X})$. We may modify the above methods to see that the number of vertices of a fixed degree is asymptotically normal. If $d = 1$, then all vertices have degree $1, 2$ or 3, a.s. If $d \geq 2$ then all degrees are possible.

We assert that for each $k \in \mathbb{N}$ (for $d \geq 2$) and for $k \in \{1, 2, 3\}$ (for $d = 1$), the distribution of $\Delta(\infty)$ is nondegenerate. For simplicity we just consider $d = 2$. All arguments below may be easily extended to the case of general $d \geq 2$. We leave the case $d = 1$ to the reader.

For $k = 1, 2$, we may argue as follows. Let E_2 be the event of the proof of Lemma 6.3. Choose points x_1, x_2 on the unit circle $|x| = 1$ such that $|x_1 - x_2| = 1/2$. Let E_0 be the intersection of E_2 and the event that there is exactly one point in each of the balls $B_{1/100}(x_i), 1 \leq i \leq 2$, and no other point inside $Q_{d+1}(0)$. Notice that on E_0 the SIG does not put edges across the moat. Let E_1 be the intersection of E_2 and the event that there is exactly one point in each of the three balls $B_{1/100}(x_i), 1 \leq i \leq 2$ and $B_{1/100}(0)$ and no other point inside $Q_{d+1}(0)$. Notice that on E_0 the SIG does not put edges across the moat. When E_0 happens, notice that H_1 decreases by 2, and H_2 increases by 3. However, when E_1 happens, H_1 increases by 4 and H_2 decreases by 3. Since E_0 and E_1 each have positive probability, we see that $\Delta(\infty)$ is nondegenerate for $k = 1, 2$.

The arguments are similar for higher values of k. For example if k is even, then consider a variation of the above, where now we place k points $x_1, \ldots, x_k$ on $|x| = 1$ in such a way that there are $k/2$ pairs such that points within each pair are within ε of each other but at least 100ε away from other points. Then consider balls $B_{\varepsilon/10}(x_i), 1 \leq i \leq k$, and let E_0 be the intersection of E_2 and the event that there is exactly one point from the Poisson process in each ball. Let E_1 be the intersection of E_2 and the event that there is exactly one point from the Poisson process in each ball and one point in $B_{\varepsilon/10}(0)$. Then on E_0, H_k increases by 1 if the origin is added whereas on E_1, H_k decreases by 1 if the origin is added. When k is odd, then modify the above in the following way. Put all points $x_i, 1 \leq i \leq k-1$, on one hemisphere and put the point x_k at the pole of the other hemisphere. Consider the analogues of E_0 and E_1: E_0 is the intersection of E_2 and the event that there is one point in each ball $B_{\varepsilon/10}(x_i), 1 \leq i \leq k$ and no point in the ball $B_{\varepsilon/10}(0)$ whereas E_1 is the intersection of E_2 and the event that there is one point in each ball $B_{\varepsilon/10}(x_i), 1 \leq i \leq k$ and one point in the ball $B_{\varepsilon/10}(0)$. Then inserting the origin on E_0 means that H_k increases by 1 (no vertices had degree k prior to the insertion of the origin), whereas inserting the origin on E_1 means that H_k

decreases by 1 (since now no points have degree k). So $\Delta(\infty)$ is nondegenerate for all values of k.

7.4. *Uniqueness of the infinite component.* To show strong stabilization for the number of components of the sphere of influence graph, we need a result on uniqueness of the infinite component, analogous to Lemma 6.4 in the case of the k-nearest neighbors graph.

THEOREM 7.3. $\mathrm{SIG}(\mathscr{P})$ *has a.s. at most one infinite component.*

For $d = 1$, a simple renewals argument shows that there is no infinite component. So we need only a proof for $d \geq 2$. This runs mostly along the lines of the proof of Theorem 4.1 of [7]. The first step is analogous to Lemma 4.2 of [7].

LEMMA 7.4. *Let $r > 0$ and let $E(r)$ be the event that $B_r(0)$ is intersected by an infinite component C of* $\mathrm{SIG}(\mathscr{P})$, *such that if all edges intersecting $B_r(0)$ are removed from this component, three of the resulting components formed from C are infinite. Then $P[E(r)] = 0$.*

This can be proved by the same sort of standard argument used in the proof of Lemma 4.2 of [7]; see, for example, [16], page 67. Therefore we omit the argument here.

LEMMA 7.5. $\mathrm{SIG}(\mathscr{P})$ *has a.s. at most two infinite components.*

PROOF. This lemma is analogous to Lemma 4.3 of [7]. Define the point process $\mathscr{P}'_{2r}$ as follows. Take $\mathscr{P}'$ to be an independent copy of $\mathscr{P}$, and let $\mathscr{P}'_{2r}$ be the union of the point process $(\mathscr{P}' \cup \mathscr{P}) \cap B_{2r}(0)$, thinned by selecting each point randomly with probability 1/2, and the point process $\mathscr{P} \setminus B_{2r}(0)$. Then $\mathscr{P}'_{2r}$ is a homogeneous Poisson process on $\mathbb{R}^d$ of unit intensity. A similarly constructed point process, there denoted X'_{3r}, is used in [16].

Define the event

$$E^*(r) = \{\text{three infinite components of SIG}\,(\mathscr{P}) \text{ intersect } B_r(0)\}.$$

As in [7], the aim is to prove $P[E^*(r)] = 0$ by showing that given $E^*(r)$ occurs, the conditional probability that the event $E(r)$ occurs for $\mathscr{P}'_{2r}$ is nonzero, and then to appeal to Lemma 7.4.

Suppose $E^*(r)$ occurs. Then there are three infinite components of $\mathrm{SIG}\,(\mathscr{P})$ which intersect $B_r(0)$; call them C_1, C_2, C_3. In proving Lemma 4.3 of [7], Häggström and Meester adopt a strategy of removing vertices from C_1, C_2, C_3 until they become connected for the k-nearest neighbors graph, $k \geq 2$. In our setting, it is not clear that such a removal strategy works. Instead we adopt a strategy of *adding* points in $B_r(0)$ to connect together C_1, C_2, C_3.

For $i = 1, 2, 3$ let C_i^* be the union of the spheres of influence of the vertices of C_i. Then by definition, C_1^*, C_2^*, C_3^* are disjoint connected subsets of $\mathbb{R}^d$, all of them entering the set $B_r(0)$. There may be other Poisson points too in $B_r(0)$; unlike in [7] we do not remove these other points.

Given $\varepsilon > 0$, let the *ε-grid* be the set $\varepsilon\mathbb{Z}^d = \{\varepsilon z\colon z \in \mathbb{Z}^d\}$. By a *path in the ε-grid* connecting C_1 and C_2 we mean a nonempty finite sequence $\gamma = (z_1, z_2, \dots, z_m)$ of elements of $\varepsilon\mathbb{Z}^d$, together with *endpoints* z_0, z_{m+1} also in $\varepsilon\mathbb{Z}^d$, such that $\{z_1, \dots, z_m\} \subset B_r(0)\setminus \cup_{i=1}^3 C_i^*$ while $z_0 \in C_1$ and $z_{m+1} \in C_2$, such that for $1 \le i \le m+1$, $\varepsilon^{-1}z_i$ and $\varepsilon^{-1}z_{i-1}$ are nearest neighbors in the integer lattice $\mathbb{Z}^d$, and such that, moreover, none of the points z_i in the path is within a distance less than 2ε from any of the points of $\mathscr{P}$. Define paths connecting C_1 and C_3 or connecting C_2 and C_3, similarly.

Given that $E^*(r)$ occurs, if ε is sufficiently small there will be two paths γ, γ', in the ε-grid, not necessarily disjoint, which together connect up C_1, C_2, C_3. We now show that γ, γ' each induce a path in $\mathrm{SIG}(\mathscr{P}'_{2r})$ which together connect up the clusters C_1, C_2, C_3.

Let $F(r, \varepsilon, \delta)$ be the event that (i) $E^*(r)$ occurs; (ii) there exist two paths γ, γ' in the ε-grid which together connect up C_1, C_2, C_3; (iii) the union of the balls $B_{\delta\varepsilon}(z)$, $z \in \gamma\cup\gamma'$ is contained in $B_r(0)\setminus\cup_{i=1}^3 C_i^*$ and (iv) the balls of radius $4\delta\varepsilon$ centered at the endpoints of the paths γ, γ' are each entirely contained in one of the sets C_i^*.

If $P[E^*(r)] > 0$, there exists $\varepsilon > 0$ and $\delta \in (0, 1/6)$ such that $P[F(r, \varepsilon, \delta)] > 0$. Choose such an ε and δ. If $F(r, \varepsilon, \delta)$ occurs, then there is a positive probability that (i) no points of $(\mathscr{P}\cup\mathscr{P}')\cap B_r(0)$ are discarded in the thinning process, (ii) a single point of $\mathscr{P}'$ is placed in each of the balls $B_{\delta\varepsilon}(z)$ for each $z \in \gamma \cup \gamma'$ and (iii) no points are placed anywhere else by $\mathscr{P}'$. If this happens, then the added points have no effect on previous spheres of influence in C_1^*, C_2^*, C_3^*, since they lie outside the old spheres of influence. On the other hand, given neighboring points z, z' in one of the paths, for any added points Y, Y' in $B_{\delta\varepsilon}(z)$ or $B_{\delta\varepsilon}(z')$, the sphere of influence of Y has radius at least $\varepsilon(1-2\delta)$, and likewise for Y', while $|Y - Y'| \le \varepsilon(1+2\delta)$; therefore the spheres of influence of Y and Y' overlap, and thus the paths γ and γ' each induce a corresponding path in $\mathrm{SIG}(\mathscr{P}'_{2r})$. Finally, if z is in a path and w is an endpoint of the path adjacent to z, and if $Y \in B_{\delta\varepsilon}(Z)$, then $|Y - w| \le \varepsilon(1+\delta)$ so the sphere of influence of Y goes within a distance $3\varepsilon\delta$ of w, and therefore overlaps the cluster C_i^* containing w. Thus the paths in $\mathrm{SIG}(\mathscr{P}'_{2r})$, created by the added points, actually connect up the clusters C_1, C_2, C_3 as desired. Hence, if $P[E^*(r)] > 0$, then $P[E'(r)] > 0$, where $E'(r)$ denotes the event that $E(r)$ occurs for the point process $\mathscr{P}'_{2r}$. This contradicts Lemma 7.4, so we must have $P[E^*(r)] = 0$. □

The proof of Theorem 7.3 is completed by similar results to Lemmas 4.4 and 4.5 of [7], except that where [7] uses a technique of removal of vertices, we use a method of adding vertices as in the proof of the preceding lemma, in such a way as to create paths connecting these spheres of influence. Since we always make sure we add vertices lying outside the spheres of influence of existing infinite components, the added vertices do not affect these existing infinite components, except to connect them together.

7.5. *Number of components.* In this section, we let $H(\mathcal{X})$ denote the number of components in SIG$(\mathcal{X})$.

THEOREM 7.4 (CLT for the number of components of the SIG). *Suppose $\mathcal{B}$ is regular. There exists $\sigma^2 > 0$ such that as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathcal{P}_n)) \to \sigma^2$ and*

$$n^{-1/2}(H(\mathcal{P}_n) - \mathbb{E}H(\mathcal{P}_n)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma^2). \tag{7.9}$$

Additionally, there exists $\tau^2 \in (0, \sigma^2]$ such that as $n \to \infty$, $n^{-1}\operatorname{Var}(H(\mathcal{U}_{n,n})) \to \tau^2$, and

$$n^{-1/2}(H(\mathcal{U}_{n,n}) - \mathbb{E}H(\mathcal{U}_{n,n})) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2). \tag{7.10}$$

Finally, if $\mathcal{B}_0$ is regular then $n^{-1}\operatorname{Var}(H(\mathcal{X}_n)) \to \tau^2$, and

$$n^{-1/2}(H(\mathcal{X}_n) - \mathbb{E}H(\mathcal{X}_n)) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2). \tag{7.11}$$

PROOF. It is clear that $H(\mathcal{X}) \le \operatorname{card}(\mathcal{X})$, so H is polynomially bounded. Also, H is homogeneous of order 0.

Let us prove the uniform bounded moments condition. Given a finite set $\mathcal{X}$, let $M^+(\mathcal{X})$, respectively $M^-(\mathcal{X})$, be the number of edges added to the SIG, respectively removed from the SIG, when $\{0\}$ is added to the set $\mathcal{X}$. Since adding or removing an edge to a graph changes the number of components by at most 1,

$$|\Delta(\mathcal{X})| \le M^+(\mathcal{X}) + M^-(\mathcal{X}) + 1. \tag{7.12}$$

By the proof of Lemma 7.2, both $M^+(\mathcal{U}_{m,A})$ and $M^-(\mathcal{U}_{m,A})$ have fourth moments bounded uniformly in $A \in \mathcal{B}$ and in $m \in [|A|/2, 3|A|/2]$. Thus by (7.12) we obtain the uniform bounded moments property.

The proof of strong stabilization proceeds in the same way as in the case of the number of components of the k-nearest neighbors graph (Proposition 6.1), this time using Theorem 7.3. Since the argument is almost the same as for Proposition 6.1, we omit it.

Together, the above remarks show that H satisfies the conditions for Theorem 2.1 and Corollary 2.1, so the result is proved. □

8. The Voronoi graph.

In this section we assume throughout that $d = 2$, $\lambda = 1$, *the sets B_n are all boxes* and that $B_0 = Q_{1/2}(0)$. We let $\operatorname{Vor}(\mathcal{X})$ denote the Voronoi graph on a point set $\mathcal{X} \subset \mathbb{R}^2$. We show that the total edge length of the Voronoi tessellation satisfies the central limit behavior of Theorem 2.1.

8.1. *Total edge length.* In this section, we let $H(\mathcal{X})$ denote the total edge length of all of the *finite* edges in $\operatorname{Vor}(\mathcal{X})$. The following CLT extends the results of [1] and [8] which restrict attention to Voronoi tessellations over Poisson samples. We also establish the convergence of the variance of H.

THEOREM 8.1 (CLT for the total edge length in the Voronoi graph). *Suppose the sets B_n are all boxes. There exists $\sigma^2 > 0$ such that as $n \to \infty$, $n^{-1}\mathrm{Var}(H(\mathscr{P}_n)) \to \sigma^2$ and*

$$n^{-1/2}(H(\mathscr{P}_n) - \mathbb{E}H(\mathscr{P}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \sigma^2). \tag{8.1}$$

Additionally, there exists $\tau^2 \in (0, \sigma^2]$ such that as $n \to \infty$, $n^{-1}\mathrm{Var}(H(\mathscr{U}_{n,n})) \to \tau^2$ and

$$n^{-1/2}(H(\mathscr{U}_{n,n}) - \mathbb{E}H(\mathscr{U}_{n,n})) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{8.2}$$

Also, if $B_0 = Q_{1/2}(0)$, $n^{(2/d)-1}\,\mathrm{Var}(H(\mathscr{X}_n)) \to \tau^2$ and

$$n^{(1/d)-1/2}(H(\mathscr{X}_n) - \mathbb{E}H(\mathscr{X}_n)) \xrightarrow{\mathscr{D}} \mathscr{N}(0, \tau^2). \tag{8.3}$$

We prove Theorem 8.1 by showing that H satisfies the conditions of Theorem 2.1. Note that by Euler's formula, for any finite $\mathscr{X}$ we have $H(\mathscr{X}) \le 3\,\mathrm{diam}(\mathscr{X})\mathrm{card}(\mathscr{X})$ and thus the functional H is polynomially bounded. Also, H is clearly homogeneous of order 1.

It is easy to see that H is strongly stabilizing. We follow [14] closely and use a construction similar to that used for the k-nearest neighbors graph. Instead of constructing six equilateral triangles, we now construct twelve disjoint congruent isosceles triangles $T_j(t)$, $1 \le j \le 12$, where the origin is a vertex of each triangle, where each triangle has two edges of length t, where $T_j(t) \subset T_j(u)$ whenever $t < u$, and where $\cup_{t>0} \cup_{j=1}^{12} T_j(t) = \mathbb{R}^2$. Let S denote the minimum t such that each triangle $T_j(t)$, $1 \le j \le 12$, contains at least one point from the Poisson point process.

Then the insertion of the origin into the Poisson point process does not affect the structure of the Voronoi diagram at distances farther than $3S$ from the origin (see Section 4 of [14]). (The same is true for the Delaunay, relative neighbor and Gabriel graphs, which are defined in the next section.) As in Lemma 6.1, it is easy to see that the random variable S is a.s. finite, so H is strongly stabilizing.

LEMMA 8.1. *H satisfies the uniform bounded moments condition.*

PROOF. We will make heavy use of the estimates in [14]. From inequalities (4.7) and (4.9) of [14] we know that

$$|\Delta H(\mathscr{X})| \le E(0, \mathscr{X}) + F(0, \mathscr{X}),$$

where $E(0, \mathscr{X})$ denotes the combined lengths of the bounded edges of the cell consisting of points closer to 0 than to any point of $\mathscr{X}$, and where $F(0, \mathscr{X})$ denotes the combined lengths of the intersections of the bounded edges in $\mathrm{Vor}(\mathscr{X})$ with the interior of the Voronoi cell around 0 in $\mathrm{Vor}(\{0\} \cup \mathscr{X})$. Thus it is enough to show that there exists a constant c such that for all boxes B and $m \in [|B|/2, 3|B|/2]$,

$$\max(\mathbb{E}[E(0, \mathscr{U}_{m,B})^4], \mathbb{E}[F(0, \mathscr{U}_{m,B})^4]) \le c.$$

If we follow the arguments of [14] then we see that we need only show

$$\mathbb{E}[D(0, \mathscr{U}_{m,B})^4 K(0, \mathscr{U}_{m,B})^4] \le c, \tag{8.4}$$

where $D := D(0, \mathscr{U}_{m,B})$ denotes the diameter of the intersection of B with the Voronoi cell around 0 in the Voronoi diagram on $\{0\} \cup \mathscr{U}_{m,B}$, and where $K := K(0, \mathscr{U}_{m,B})$ denotes the number of sides of the Voronoi cell around 0 in the Voronoi diagram on $\{0\} \cup \mathscr{U}_{m,B}$. For each $t > 0$, construct twelve disjoint congruent isosceles triangles $T_j(t)$, $1 \le j \le 12$, having union $\mathbb{R}^2$, where the point 0 is a vertex of each triangle, where each triangle has two edges of length t and where $T_j(t) \subset T_j(u)$ whenever $0 < t < u$. For all $1 \le j \le 12$, let S_j be the minimum t such that the triangle $T_j(t)$ contains at least one point from $\mathscr{U}_{m,B}$, if such a t exists, or to be the diameter of $T_j(t) \cap B$, if not. Let $S = \max(S_1, \dots, S_{12})$. As in [14], simple geometric considerations show that for all boxes B including those with 0 near the boundary of B, we have $D \le 2S$. Note that there is a constant c such that for all B and m of interest,

$$P[S_j > t] \le (1 - ct/|B|)^m \le \exp(-ct/2),$$

and therefore the tail of the distribution of S decays exponentially, uniformly in B and m. Then using the arguments of Section 4 of [14], we can obtain (8.4). □

LEMMA 8.2. *The distribution of* $\Delta(\infty)$ *is nondegenerate.*

PROOF. Consider the construction used in the proof of Lemma 6.3. Let E_2 be the event that there are no points of $\mathscr{P}$ in the moat and there is at least one point in each of the subcubes in $\mathscr{A}$. Fix ε small ($<1/100$). Choose points $x_1, x_2, x_3 \in \mathbb{R}^2$ forming an equilateral triangle of side length 1/2, centered at the origin. Consider the balls of radius ε centered at the points x_1, x_2, x_3. Let A_0 be the intersection of E_2 and the event that there is exactly one point in each of the three balls and no other point in the central "island" C_0. Let A_1 be the intersection of E_2 and the event that there is exactly one point in each of the balls of radius $\varepsilon\delta$ centered at the points $\delta x_1, \delta x_2, \delta x_3$, where $\delta \in (0, 1)$ will be chosen shortly, and no other point in the central island.

On the event A_0, the insertion of the origin leads to three new edges, namely the edges of a (nearly equilateral) triangular cell T around the origin. It removes the parts of the three edges of the original Voronoi graph which intersect T. The difference between the sum of the lengths of the added edges and the sum of the lengths of the three removed edges exceeds some fixed positive number α [the reason is this: given an equilateral triangle T, and a point P inside it, the sum of the lengths of the three edges joining P to the vertices of T is strictly less than the perimeter of T since the length of each of the three edges is less than the common length of the side of T. If T is nearly equilateral (our case) this is still true].

On the other hand, on the event A_1, the insertion of the origin cannot increase the total edge length by more than the total edge length of triangular cell around the origin, and this increase is bounded by a constant multiple of δ, which is less than α if δ is small enough. Thus if δ is small enough, the events A_0 and A_1 give rise to values of $\Delta(\infty)$ which differ by at least some fixed amount. This shows the nondegeneracy of $\Delta(\infty)$. □

Thus H satisfies all of the conditions of Theorem 2.1 and thus Theorem 8.1 is proved.

8.2. *Total number of edges and vertices.* The above discussion also applies in part to other functionals of the Voronoi tessellation. For example, if $H(\mathscr{X})$ counts the number of edges in $\mathrm{Vor}(\mathscr{X})$, then it can be checked that H is strongly stabilizing, satisfies the uniform moment condition, is homogeneous of order 0 and therefore satisfies the conclusions of Corollary 2.1 with $\gamma = 0$. However, this is one instance where the limiting variance τ^2 is zero, and therefore the "correct" scaling of the variance is not by n^{-1}, so that our results are not so relevant in this case.

The reason for this degeneracy is as follows. All vertices of $\mathrm{Vor}(\mathscr{X}_n)$ a.s. have degree 3 (see [21], Theorem 5.7). Therefore, if V_n denotes the number of vertices, E_n the number of edges and I_n the number of infinite edges, we have $3V_n = 2E_n - I_n$. On the other hand, by Euler's formula, since the number of faces is n, we have $V_n - E_n + n = 1$ (it is not the usual Euler's formula because we are not counting the "vertex at infinity"). Combining these two simultaneous equations, we obtain $E_n = 3n - 3 - I_n$. Therefore $\mathrm{Var}(E_n) = \mathrm{Var}(I_n)$.

The value of I_n is equal to the number of points of $\mathscr{X}_n$ lying on the boundary of the convex hull of $\mathscr{X}_n$, and therefore its variance is asymptotic to a constant times $\log n$ [6]. This shows that $\tau^2 = 0$. Moreover, a similar discussion applies when H is the number of *vertices* of the Voronoi graph.

It would be interesting to know if there is a way of adapting our method to get nondegenerate CLTs for quantities such as the number of vertices in the convex hull, whose variances do not grow in proportion to n. See [1] for an adaptation of a method to the convex hull problem.

9. Other proximity graphs. The discussion in the preceding sections applies to other graphs in computational geometry and, at a minimum, covers the case when the sets B_n are all boxes and $B_0 = Q_{1/2}(0)$. In the examples which follow, we see that functionals of such graphs (such as total edge length and total number of edges) satisfy strong stabilization, the uniform bounded moments condition and the nondegeneracy of $\Delta(\infty)$. Moreover, the proofs of these facts are nearly exact replicas of the proofs above and we leave the details to the reader. We now describe some of the graphs covered by the above discussion. See [3, 22] for more details on these and related proximity graphs.

Delaunay triangulation. The Delaunay triangulation of a point set $\mathscr{X} \subset \mathbb{R}^d$ is the graph which is dual to the Voronoi tessellation; it puts an edge between two points of $\mathscr{X}$ if and only if these points are centers of adjacent Voronoi cells. The total edge length of the Delaunay triangulation satisfies a CLT analogous to Theorem 8.1. The Delaunay triangulation on n points has at most $n(n-1)/2$

edges and so the total edge length is polynomially bounded. The radius of stabilization for the total edge length of the Delaunay triangulation is the same as that for the Voronoi tessellation. By modifying Lemma 6.2 in a straightforward fashion, one can show that the total edge length of the Delaunay triangulation satisfies the fourth moment condition and in this way avoid the complications present in Lemma 8.1.

When $d = 2$, the number of edges of the Delaunay triangulation is the same as the number of edges of the Voronoi tessellation, so the discussion in Section 8.2 applies.

Gabriel graph. The Gabriel graph on a point set $\mathscr{X}$ puts an edge between two points x, y of $\mathscr{X}$ if the ball centered at $(x+y)/2$ with x and y at opposite poles does not contain any other points in $\mathscr{X}$. The Gabriel graph is a subgraph of the Delaunay triangulation. Strong stabilization for the Gabriel graph can be established by an argument using cones or triangles in much the same way as for the k-nearest neighbors graph or the Voronoi graph. Fourth moments are handled as in Lemma 6.2.

Relative neighborhood graph. The relative neighborhood graph on $\mathscr{X}$ is formed by joining all pairs of points whose loon is empty, where the loon defined by a pair is the intersection of two spheres of equal radius, each having one point as center and the other point on its surface. The relative neighborhood graph is a subgraph of the Gabriel graph. Strong stabilization for the relative neighborhood graph can be established in much the same way as for the k-nearest neighbors graph or the Voronoi graph. Fourth moments are handled as in Lemma 6.2.

Power weighted edges. Kesten and Lee [11] prove a central limit theorem for the total edge length of the minimal spanning tree on a random sample when the edges are power-weighted. All of the graphs described here admit versions with power weighted edges and it is trivial to show that the total edge length of such graphs satisfies all the conditions of our main theorems.

Percolation. One simple way to obtain a graph on $\mathscr{X}$ is to connect all pairs of points which are at most unit distance apart. One can obtain a (nonhomogeneous) functional $H_{\mathrm{occ}}(\mathscr{X})$ by counting the components of the resulting graph. This is equivalent to a basic model of continuum percolation [16], in which one takes balls of unit diameter around each point and counts the connected components of the union of the balls (*occupied clusters*). One can also consider the number $H_{\mathrm{vac}}(\mathscr{X})$ of *vacant clusters*, by which we mean components of the complement of the union of balls. Using results in [16] on uniqueness of the infinite cluster, both H_{occ} and H_{vac} can be shown to satisfy strong stabilization, and hence a CLT for the uniform sample $\mathscr{U}_{n,n}$. This adds to a result in [19] on CLTs for occupied and vacant cluster counts (in a more general setting) for Poisson samples.

Acknowledgment. MDP's work on this research started during a visit in April–June 1999 to the Fields Institute in Toronto, whose hospitality and support is much appreciated.

REFERENCES

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

[2] BICKEL, P. J. and BREIMAN, L. (1983). Sums of functions of nearest neighbor distances, moment bounds, limit theorems and a goodness of fit test. *Ann. Probab.* **11** 185–214.

[3] DEVROYE, L. (1988). The expected size of some graphs in computational geometry. *Comput. Math. Appl.* **15** 53–64.

[4] DURRETT, R. (1991). *Probability: Theory and Examples.* Wadsworth & Brooks/Cole, Pacific Grove, CA.

[5] FÜREDI, Z. (1997). The expected size of the sphere of influence graph. *Bolyai Soc. Math. Stud.* **6** ("Intuitive Geometry Budapest 1995") 319–326.

[6] GROENEBOOM, P. (1988). Limit theorems for convex hulls. *Probab. Theory Related Fields* **79** 327–368.

[7] HÄGGSTRÖM, O. and MEESTER, R. (1996). Nearest neighbor and hard sphere models in continuum percolation. *Random Structures Algorithms* **9** 295–315.

[8] HEINRICH, L. (1994). Normal approximation for some mean value estimates of absolutely regular tessellations. *Math. Methods Statist.* **3** 1–24.

[9] HENZE, N. (1987). On the fraction of random points with specified nearest-neighbour interactions and degree of attraction. *Adv. in Appl. Probab.* **19** 873–895.

[10] HITCZENKO, P., JANSON, S. and YUKICH, J. E. (1999). On the variance of the random sphere of influence graph. *Random Structures Algorithms* **14** 139–152.

[11] KESTEN, H. and LEE, S. (1996). The central limit theorem for weighted minimal spanning trees on random points. *Ann. Appl. Probab.* **6** 495–527.

[12] LEE, S. (1997). The central limit theorem for Euclidean minimal spanning trees I. *Ann. Appl. Probab.* **7** 996–1020.

[13] LEE, S. (1999). The central limit theorem for Euclidean minimal spanning trees II. *Adv. in Appl. Probab.* **31** 969–984.

[14] MCGIVNEY, K. and YUKICH, J. E. (1999). Asymptotics for Voronoi tessellations on random samples. *Stochastic Process. Appl.* **83** 273–288.

[15] MCLEISH, D. L. (1974). Dependent central limit theorems and invariance principles. *Ann. Probab.* **2** 620–628.

[16] MEESTER, R. and ROY, R. (1996). *Continuum Percolation*. Cambridge Univ. Press.

[17] MICHAEL, T. S. and QUINT, T. (1994). Sphere of influence graphs: a survey. *Congr. Numer.* **105** 153–160.

[18] OKABE, A., BOOTS, B. and SUGIHARA, K. (1992). *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley, Chichester.

[19] PENROSE, M. D. (2001). A spatial central limit theorem with applications to percolation, epidemics and Boolean models. *Ann. Probab.* **29**.

[20] PENROSE, M. D. and YUKICH, J. E. Limit theory for random sequential packing and deposition. *Ann. Appl. Probab.* **12**.

[21] PREPARATA, F. P. and SHAMOS, M. I. (1985). *Computational Geometry: an Introduction.* Springer, New York.

[22] TOUSSAINT, G. T. (1982). Computational geometric problems in pattern recognition. In *Pattern Recognition Theory and Applications* (J. Kittler, K. S. Fu and L. F. Pau, eds.), Reidel, Dordrecht. 73–91.

[23] YUKICH, J. E. (1998). *Probability Theory of Classical Euclidean Optimization Problems. Lecture Notes in Math.* **1675.** Springer, Berlin.

DEPARTMENT OF MATHEMATICAL SCIENCES
UNIVERSITY OF DURHAM
SOUTH ROAD, DURHAM DH1 3LE
ENGLAND
E-MAIL: mathew.penrose@durham.ac.uk

DEPARTMENT OF MATHEMATICS
LEHIGH UNIVERSITY
BETHLEHEM, PENNSYLVANIA 18015
E-MAIL: joseph.yukich@lehigh.edu