Multivariate spatial central limit theorems with applications to percolation and spatial graphs

MATHEW D. PENROSE
University of Bath

February 2004

Abstract

Suppose $X = (X_x, x \text{ in } Z^d)$ is a family of i.i.d. variables in some measurable space, B_0 is a bounded set in R^d, and for $t > 1$, H_t is a measure on tB_0 determined by the restriction of X to lattice sites in or adjacent to tB_0. We prove convergence to a white noise process for the random measure on B_0 given by $t^{-d/2}(H_t(tA) - EH_t(tA))$ for subsets A of B_0, as t becomes large, subject to H satisfying a "stabilization" condition (whereby the effect of changing X at a single site x is local) but with no assumptions on the rate of decay of correlations. We also give a multivariate central limit theorem for the joint distributions of two or more such measures H_t, and adapt the result to measures based on Poisson and binomial point processes. Applications given include a white noise limit for the measure which counts clusters of critical percolation, a functional central limit theorem for the empirical process of the edge lengths of the minimal spanning tree on random points, and central limit theorems for the on-line nearest neighbour graph.

Short title: **Multivariate spatial CLTs**

Keywords. Central limit theorem, white noise, minimal spanning tree, empirical process, on-line nearest neighbour graph, percolation.

American Mathematical Society 2000 Classifications. Primary-60F05, 60D05; Secondary-05C80, 60K35.

1 Introduction

Several approaches have been developed for proving central limit theorems for random variables which arise as the sum of contributions from points of a Poisson or binomial point process in $\mathbb{R}^d$, when each contribution is locally determined in some sense. These include Stein's method, the method of moments, and a martingale method.

By keeping track of the location of each contribution in $\mathbb{R}^d$, one can often in a natural way create a random measure, i.e. a random field indexed by subsets of $\mathbb{R}^d$ or by test functions on $\mathbb{R}^d$. It is of interest to look for multivariate central limit theorems for such random fields, typically with weak convergence of finite-dimensional distributions to those of white noise. Multivariate central limit theorems of this type were recently derived using the method of moments by Baryshnikov and Yukich [2] and can also be derived using Stein's method [16]. Both of these methods seem to require, in addition to a 'stabilization' condition which formalizes the locally determined contributions, a form of exponential decay of spatial correlations.

It is of interest to extend these results to cases which satisfy stabilization but are not believed to satisfy exponential decay. These include, for example, measures associated with the minimal spanning tree (MST) on a Poisson point process or with critical percolation. The martingale method is especially powerful in giving central limit theorems for these examples (Kesten and Lee [7], Zhang [22]). For an exposition of this method in a general setting, see Penrose [11], Penrose and Yukich [14]. However, these works do not address the convergence to white noise of random measures.

In the present paper we extend the martingale method to give such a convergence to white noise of stabilizing random fields indexed by subsets of $\mathbb{R}^d$, and illustrate the method both for percolation and for the minimal spanning tree. A further multivariate direction in which we extend the existing literature is by considering convergence to a multivariate normal for two or more random fields based on the same underlying spatial process; in particular, we shall show that the finite-dimensional distributions of the empirical process of the lengths of the MST on random points, suitably scaled and centred, converge to those of a certain Gaussian process.

A further aim of this paper is to treat discrete examples (such as percolation) and continuous ones (such as the MST) in a unified manner. In this spirit we shall derive our basic general result for Poissonian continuum systems (Theorem 2.2) by direct application of the basic result for lattice systems (Theorem 2.1), although some extra work is needed to give the limiting covariances for Theorem 2.2 in a more explicit form than was done in previous continuum central limit theorems proved by the martingale method. For this reason, in the continuum we consider only uniform densities of points over a fixed d-dimensional set (denoted B_0 in the sequel), unlike Baryshnikov and Yukich who consider non-uniform densities. It is likely that with extra work, the martingale-based proof of Theorem 2.2 could be extended to give multivariate central limit theorems for point processes with non-uniform densities. For martingale-based proofs of univariate central limit theorems on non-uniform points, see Lee [9] and Penrose ([12], Section 13.7).

In the continuum, stabilizing random fields are often defined in terms of graphs G which are themselves stabilizing, i.e., locally determined in a certain sense. Stabilizing graphs include the MST, k-nearest neighbour, and sphere of influence graphs. Given a stabilizing graph G, the theory presented here applies to random fields (indexed by subsets A of $\mathbb{R}^d$) which count, for example, the number of leaves of G in A, the number of components of G that include vertices in A, or the sum of weighted edge lengths $\phi(|e|)$ with the sum over edges e of G having endpoints in A.

One stabilizing graph which has not been considered in previous discussions of stabilizing graphs is the *on-line nearest neighbour graph*, in which random points in B_0 are randomly ordered, and each point (except the first) is connected to its nearest neighbour amongst its predecessors in the ordering. This graph is of recent interest in connection with the modelling of scale-free networks [3, 5]. Unlike methods based on exponential decay, our methods provide central limit theorems for this graph too; see Section 3.4.

The rest of the paper is laid out as follows. The next section contains statements of the main general results. Section 3 contains applications of these. Sections 4 and 5 contain proofs of the general results.

2 General multivariate central limit theorems

2.1 Notation used throughout

Let $d \geq 1$ be an integer and let $\mathbf{0}$ denote the origin of $\mathbb{R}^d$. For $A \subseteq \mathbb{R}^d$, $t \in \mathbb{R}$, and $y \in \mathbb{R}^d$, let tA denote the scaled set $\{tx : x \in A\}$, and let $\tau_y(A)$ denote the translated set $\{x + y : x \in A\}$. Let $\partial(A)$ denote the boundary of A, that is, the intersection of the closure of A with that of its complement. If A is (Lebesgue) measurable, write $|A|$ for its Lebesgue measure, and if A is finite, write $\mathrm{card}(A)$ for the number of elements of A. Given a sequence of sets $(A_n)_{n\geq 1}$, write $\liminf(A_n)$ for $\cup_{n=1}^{\infty}(\cap_{m=n}^{\infty} A_m)$.

For $z \in \mathbb{Z}^d$, and $\varepsilon > 0$, let Q_z^{ε} denote the cube $\tau_{\varepsilon z}([-\varepsilon, 0)^d)$. For $x \in \mathbb{R}^d$, and $r > 0$, let $B_r(x)$ be the closed Euclidean ball of radius r centred at x. Let ρ be a finite constant, satisfying $\rho \geq \sqrt{d}$ but otherwise arbitrary. For $A \subseteq \mathbb{R}^d$, let $\widetilde{A}$ denote the discretization of A given by

$$\widetilde{A} := \{z \in \mathbb{Z}^d : B_\rho(z) \cap A \neq \emptyset\}. \tag{2.1}$$

The condition $\rho \geq \sqrt{d}$ ensures that $Q_z^{\varepsilon} \subset B_\rho(\varepsilon z)$ for $\varepsilon \leq 1$.

For $\sigma > 0$, let $\mathcal{N}(0, \sigma^2)$ be the normal probability distribution on $\mathbb{R}$ with density $f(x) = (2\pi\sigma^2)^{-1/2}\exp(-x^2/(2\sigma^2))$. Also, let $\mathcal{N}(0, 0)$ represent the degenerate probability distribution on $\mathbb{R}$ consisting of a unit point mass at zero, which we view as a special case of the normal. Given a nonnegative definite $k \times k$ matrix Σ, let

$\mathcal{N}(\mathbf{0}, \Sigma)$ denote the centred multivariate normal distribution with covariance matrix Σ, i.e. the distribution of a random k-vector $\mathbf{X}$ satisfying $\mathbf{a}'\mathbf{X} \sim \mathcal{N}(0, \mathbf{a}'\Sigma\mathbf{a})$ for all deterministic k-vectors $\mathbf{a}$. Denote convergence in probability by $\xrightarrow{P}$, convergence in pth moment by $\xrightarrow{L^p}$, and denote equality in distribution by $\stackrel{\mathcal{D}}{=}$.

Let B_0 be a Riemann measurable subset of $\mathbb{R}^d$, i.e., one with Riemann integrable indicator function, or in other words, a bounded measurable set with Lebesgue-null boundary (see Rudin [17]). Assume that B_0 has non-empty interior. For example, B_0 could be the d-dimensional unit cube. The set B_0 is fixed throughout the sequel. Let $\mathcal{R}(\mathbb{R}^d)$ denote the collection of Riemann measurable subsets of $\mathbb{R}^d$, and let $\mathcal{R}(B_0)$ denote the collection of Riemann measurable subsets of B_0.

For $y, z \in \mathbb{Z}^d$, write $y \prec z$ if y precedes z in the lexicographic ordering on $\mathbb{Z}^d$, and $y \preccurlyeq z$ if either $y \prec z$ or $y = z$.

2.2 A central limit theorem for lattice systems

Let $(E, \mathcal{E}, P_0)$ be an arbitrary probability space. On a suitable probability space $(\Omega, \mathcal{F}, P)$, let $X = (X_z, z \in \mathbb{Z}^d)$ be be a family of independent identically distributed random elements of E, each X_z having distribution P_0, indexed by the integer lattice, and let X_* be a further E-valued variable with distribution P_0, independent of X (i.e., an independent copy of $X_{\mathbf{0}}$). For existence of such an $(\Omega, \mathcal{F}, P)$ and X, see for example section 8.7 of Williams [19]. For $y \in \mathbb{Z}^d$, let $\tau_y X$ denote the translated family of variables $(X_{z+y}, z \in \mathbb{Z}^d)$.

By a *random set function* we mean a collection $H = (H_t(A) : t \geq 1, A \in \mathcal{R}(B_0))$, where for each $t \geq 1$ and $A \in \mathcal{R}(B_0)$ the $H_t(A)$ is a random variable that is a function of $(X_z, z \in \widetilde{tB_0})$, so that, strictly speaking, $H_t(A)$ is itself a measurable function from $E^{\widetilde{tB_0}}$ to $\mathbb{R}$. If we wish to emphasize the dependence on X of the value of $H_t(A)$ we write $H_t(X, A)$ for $H_t(A)$. In many examples $H_t(\cdot)$ is a (random) measure or outer measure (see, e.g., Durrett [4]) on Borel subsets of tB_0 but we do not need to assume this for the general result (we restrict attention to Riemann measurable A). For $t \geq 1$, $y \in \mathbb{Z}^d$, and $A \in \mathcal{R}(B_0)$, define

$$H_{t,y}(A) = H_{t,y}(X, A) := H_t(\tau_y X, A) \tag{2.1}$$

In [11], a CLT is established under a stabilization condition which says, loosely speaking, that the effect on a given random set function H of resampling the value of X at a single site is local. To extend these to central limit theorems for random fields, we require a modification of the stabilization condition used in [11].

Let $X^{\mathbf{0}}$ be the process X with the value $X_{\mathbf{0}}$ at the origin replaced by the independent copy X_* of $X_{\mathbf{0}}$, but with the values at all other sites the same (i.e., $X^{\mathbf{0}} = (X^{\mathbf{0}}_z, z \in \mathbb{Z}^d)$ is the random field given by $X^{\mathbf{0}}_{\mathbf{0}} = X_*$ and $X^{\mathbf{0}}_z = X_z$ for $z \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$). Given a random set function H, define the increment $\Delta^H_{t,y}(A)$ for $y \in \mathbb{Z}^d$ and $t \geq 1, A \in \mathcal{R}(B_0)$, by

$$\Delta^H_{t,y}(A) = H_{t,y}(X, A) - H_{t,y}(X^{\mathbf{0}}, A). \tag{2.2}$$

We consider random set functions H with the property that there exists a random variable Δ_∞^H such that for all $A \in \mathcal{R}(B_0)$, and all $[1,\infty) \times \mathbb{Z}^d$-valued sequences $(t_n, y_n)_{n\geq 1}$:

$$\Delta_{t_n,y_n}^H(A) \xrightarrow{P} \Delta_\infty^H \quad \text{if} \quad \liminf_{n\to\infty}(\tau_{y_n}(t_n A)) = \mathbb{R}^d \tag{2.3}$$

and also

$$\Delta_{t_n,y_n}^H(A) \xrightarrow{P} 0 \quad \text{if} \quad \liminf_{n\to\infty}(\tau_{y_n}(t_n(B_0 \setminus A))) = \mathbb{R}^d. \tag{2.4}$$

Eqns (2.3) and (2.4) are our stabilization conditions. The second condition (2.4) is a novel feature of this paper; it was not required for the CLTs presented in [11, 14]. The first condition (2.3) is similar to the stabilization condition in Definition 2.3 of [11].

We shall require also that there exist $\gamma > 2$ such that the moments condition

$$\sup\{\mathbb{E}[|\Delta_{t,-y}^H(A)|^\gamma] : A \in \mathcal{R}(B_0), t \geq 1, y \in \widetilde{tB_0}\} < \infty \tag{2.5}$$

is satisfied. Observe that $\Delta_{t,-y}^H(A)$ is identically zero for $y \in \mathbb{Z}^d \setminus \widetilde{tB_0}$, and therefore condition (2.5) is equivalent to

$$\sup\{\mathbb{E}[|\Delta_{t,-y}^H(A)|^\gamma] : A \in \mathcal{R}(B_0), t \geq 1, y \in \mathbb{Z}^d\} < \infty. \tag{2.6}$$

For $y \in \mathbb{Z}^d$, let $\mathcal{F}_y$ be the σ-field generated by $(X_z, z \preccurlyeq y)$ (recall that $\preccurlyeq$ denotes the lexicographic ordering on $\mathbb{Z}^d$). Now we can state our main general result for lattice systems.

Theorem 2.1 *Suppose that $H^1, \ldots, H^k$ are random set functions, each of which satisfies the stabilization conditions* (2.3) *and* (2.4), *along with the moments condition* (2.5) *for some $\gamma > 2$. Let the $k \times k$ matrix $(\sigma_{ij}^*)_{i,j=1}^k$ be given by*

$$\sigma_{ij}^* := \mathbb{E}[\mathbb{E}(\Delta_\infty^{H^i}|\mathcal{F}_{\mathbf{0}})\mathbb{E}(\Delta_\infty^{H^j}|\mathcal{F}_{\mathbf{0}})]. \tag{2.7}$$

Then if $A_, \ldots, A_k$ are Riemann measurable subsets of B, for $1 \leq i \leq j \leq k$ we have

$$\lim_{t\to\infty} t^{-d}\mathrm{Cov}(H_t^i(A_i), H_t^j(A_j)) = \sigma_{ij}^*|A_i \cap A_j|, \tag{2.8}$$

and as $t \to \infty$,

$$(t^{-d/2}(H_t^i(A_i) - \mathbb{E}\, H_t^i(A_i)))_{i=1}^k \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, (\sigma_{ij}^*|A_i \cap A_j|)_{i,j=1}^k). \tag{2.9}$$

In many examples, we consider only the case of a single random set function, that is, the case where each of $H^1, \ldots, H^k$ are all the same random set function H. In this case the result says that all the finite-dimensional joint distributions of $(t^{-d/2}(H_t(A) - \mathbb{E}\, H_t(A)), A \in \mathcal{R}(B_0))$, converge to those of a centred Gaussian process $(W(A), A \in \mathcal{R}(B_0))$ with covariance function

$$E[W(A)W(A')] = |A \cap A'|\mathbb{E}[(\mathbb{E}[\Delta_\infty^H|\mathcal{F}_{\mathbf{0}}])^2],$$

i.e., a white noise process.

2.3 Central limit theorems for continuum systems

By a *point process set function* we mean a real-valued functional $h(\mathcal{X}, A)$ defined for all $A \in \mathcal{R}(\mathbb{R}^d)$ and finite subsets $\mathcal{X}$ of $\mathbb{R}^d$, such that $h(\{x_1, \dots, x_k\})$ is a Borel-measurable function of $(x_1, \dots, x_k)$ for all k, and h satisfies the translation-invariance condition

$$h(\tau_y(\mathcal{X}), \tau_y(A)) = h(\mathcal{X}, A) \ \ \forall \ A \in \mathcal{R}(\mathbb{R}^d), \ y \in \mathbb{R}^d, \ \ \text{finite } \mathcal{X} \subset \mathbb{R}^d. \tag{2.10}$$

Let $\lambda > 0$ and let $\mathcal{P}$ denote a homogeneous Poisson point process in $\mathbb{R}^d$ of intensity λ (viewed as a random subset of $\mathbb{R}^d$). With λ and B_0 both fixed throughout the sequel, define the point processes

$$\mathcal{P}_t := \mathcal{P} \cap (tB_0), \qquad t \geq 1.$$

We derive a multivariate central limit theorem for $(h(\mathcal{P}_t, tA), A \in \mathcal{R}(B_0))$ as $t \to \infty$. The conditions on h for our central limit theorem are defined in terms of the "add one cost on A" defined by

$$\delta(A, \mathcal{X}) := h(\mathcal{X} \cup \{\mathbf{0}\}, A) - h(\mathcal{X}, A). \tag{2.11}$$

We shall say the point process set function h is *strongly stabilizing* if there exist almost surely finite random variables $\delta(\infty)$ (the *stabilizing limit* of h) and S (a *radius of stabilization* of h) such that with probability 1, $\mathcal{P}$ is such that for $A \in \mathcal{R}(\mathbb{R}^d)$ and for all finite $\mathcal{A} \subset (\mathbb{R}^d \setminus B_S(\mathbf{0}))$,

$$\delta(A, (\mathcal{P} \cap B_S(\mathbf{0})) \cup \mathcal{A}) = \delta(\infty) \quad \text{if} \quad A \supseteq B_S(\mathbf{0}), \tag{2.12}$$

and

$$\delta(A, (\mathcal{P} \cap B_S(\mathbf{0})) \cup \mathcal{A}) = 0 \quad \text{if} \quad A \cap B_S(\mathbf{0}) = \emptyset. \tag{2.13}$$

Thus, S is a radius of stabilization if the add one cost on A for the restriction of $\mathcal{P}$ to a region containing the ball $B_S(\mathbf{0})$ is unaffected by changes in the configuration outside the ball $B_S(\mathbf{0})$ if $B_S(\mathbf{0}) \subseteq A$ or $B_S(\mathbf{0}) \cap A = \emptyset$, taking the value $\delta(\infty)$ if $B_S(\mathbf{0}) \subseteq A$ and the value zero if $B_S(\mathbf{0}) \cap A = \emptyset$. Our notion of strong stabilization (2.12) is similar to that used in [14], which also has a notion of 'weak stabilization' that will not concern us here. The second stabilization condition (2.13), like its discrete counterpart (2.4), is new to this paper.

Given $t \in [1, \infty)$, and $m \in \mathbb{N}$, let $\mathcal{U}_{m,t}$ be a point process consisting of m independent random d-vectors, each of them uniformly distributed on tB_0. Also let μ_t be the expected number of points of $\mathcal{P}_t$, i.e., let

$$\mu_t := \lambda t^d |B_0| \tag{2.14}$$

We consider functionals h satisfying the moments condition

$$\sup_{t \geq 1, A \in \mathcal{R}(B_0), x \in -tB_0} \ \sup_{m \in [\mu_t/2, 3\mu_t/2]} \{\mathbb{E}[\delta(\tau_x(tA), \tau_x(\mathcal{U}_{m,t}))^4]\} < \infty. \tag{2.15}$$

In the sequel, it is likely that the fourth moments condition (2.15) can be replaced by a $2+\epsilon$ moment condition, but this would not greatly expand the range of applications known to the author.

We also require a mild uniform bound on the size of h, whereby there exists a constant β_2 such that for all finite sets $\mathcal{X} \subset \mathbb{R}^d$, and all $A \in \mathcal{R}(\mathbb{R}^d)$,

$$|h(\mathcal{X}, A)| \leq \beta_2(\mathrm{diam}(\mathcal{X}) + \mathrm{card}(\mathcal{X}))^{\beta_2}. \tag{2.16}$$

We now give a multivariate CLT for $h(\mathcal{P}_t, tA)$. This is the first of our continuum analogues to Theorem 2.1. 2.1

Theorem 2.2 *Suppose that $h^1, \ldots, h^k$ are point process set functions which satisfy the stabilization conditions* (2.12), (2.13), *the moments condition* (2.15), *and the polynomial bound* (2.16). *Define the $k \times k$ matrix $(\sigma_{ij})_{i,j=1}^k$ by*

$$\sigma_{ij} = \mathbb{E}[\mathbb{E}(\delta^i(\infty)|\mathcal{F})\mathbb{E}(\delta^j(\infty)|\mathcal{F})], \tag{2.17}$$

where $\mathcal{F}$ denotes the σ-field generated by the Poisson configuration in the half-space $\{x = (x_1, \ldots, x^d) \in \mathbb{R}^d : x_1 < 0\}$, and δ_i is the stabilizing limit of h^i. If $A_1, \ldots, A_k$ are sets in $\mathcal{R}(B_0)$, then as $t \to \infty$,

$$t^{-d}\mathrm{Cov}(h^i(\mathcal{P}_t, tA_i), h^j(\mathcal{P}_t, tA_j)) \to \lambda\sigma_{ij}|A_i \cap A_j| \tag{2.18}$$

and

$$t^{-d/2}(h^i(\mathcal{P}_t, A_i) - \mathbb{E}h^i(\mathcal{P}_t, A_i))_{i=1}^k \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, (\lambda\sigma_{ij}|A_i \cap A_j|)_{i,j=1}^k). \tag{2.19}$$

The next result is a de-Poissonized version of Theorem 2.2.

Theorem 2.3 *Suppose that $h^1, \ldots, h^k$ are point process set functions, each h^j satisfying the strong stabilization conditions* (2.12) (*with stabilizing limit denoted $\delta^j(\infty)$*) *and* (2.13), *along with the moments condition* (2.15), *and the polynomial bound* (2.16). *Let the matrix $\Sigma = (\sigma_{ij})_{i,j=1}^k$ be given by* (2.17). *Then if $A_1, \ldots, A_k$ are sets in $\mathcal{R}(B_0)$, if we define the matrix $\mathcal{T} := (\tau_{ij})_{i,j=1}^k$ by*

$$\tau_{ij} := \frac{\sigma_{ij}|A_i \cap A_j|}{|B_0|} - \frac{|A_i| \cdot |A_j|}{|B_0|^2}E[\delta^i(\infty)]E[\delta^j(\infty)], \tag{2.20}$$

and if $(t_n)_{n\geq 1}$ is a $[1, \infty)$-valued sequence satisfying $n/(t_n^d|B_0|) \to \lambda$ as $n \to \infty$, then for $i, j \in \{1, \ldots, k\}$,

$$\lim_{n\to\infty} n^{-1}\mathrm{Cov}(h^i(\mathcal{U}_{n,t_n}, t_nA_i), h^j(\mathcal{U}_{n,t_n}, t_nA_j)) = \tau_{ij} \tag{2.21}$$

and as $n \to \infty$,

$$n^{-1/2}(h^i(\mathcal{U}_{n,t_n}, t_nA_i) - \mathbb{E}h^i(\mathcal{U}_{n,t_n}, t_nA_i))_{i=1}^k \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, \mathcal{T}). \tag{2.22}$$

Given $\gamma \in \mathbb{R}$, we shall say h is *homogeneous of order* γ if

$$h(a\mathcal{X}, aA) = a^\gamma h(\mathcal{X}, A), \qquad \forall a \in \mathbb{R},\ A \in \mathcal{R}(\mathbb{R}^d),\ \text{finite } \mathcal{X} \subset \mathbb{R}^d. \tag{2.23}$$

If h satisfies homogeneity, it is easy to deduce from the above theorems a multivariate CLT, either for a homogeneous Poisson processes of intensity λ on B_0 as $\lambda \to \infty$, or for a sample $\mathcal{U}_{n,1}$ of non-random size n from the uniform distribution on B_0 as $n \to \infty$. Here, we just state a result of the second type.

Theorem 2.4 *Suppose $h^1, \ldots, h^k$ are point process set functions, satisfying the strong stabilization conditions* (2.12), (2.13), *the moments condition* (2.15), *the polynomial bound* (2.16), *and homogeneity of order γ* (2.23) *for some $\gamma \in \mathbb{R}$. Suppose that $A_1, \ldots, A_k$ are sets in $\mathcal{R}(B_0)$, and let $\mathcal{T} = (\tau_{ij})_{i,j=1}^k$ be given by* (2.20) *in the case $\lambda = |B_0|^{-1}$ of Theorem* 2.3. *Then for $i, j \in \{1, \ldots, k\}$*

$$\lim_{n\to\infty} n^{(2\gamma/d)-1} \mathrm{Cov}(h^i(\mathcal{U}_{n,1}, A_i), h^j(\mathcal{U}_{n,1}, A_j)) = \tau_{ij} \tag{2.24}$$

and as $n \to \infty$,

$$n^{(\gamma/d)-1/2}(h^i(\mathcal{U}_{n,1}, A_i) - \mathbb{E}\, h^i(\mathcal{U}_{n,1}, A_i))_{i=1}^k \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathcal{T}). \tag{2.25}$$

Theorem 2.4 is easily proved by applying Theorem 2.3 with $t_n = n^{1/d}$, and using homogeneity of h^i to deduce that

$$(n^{\gamma/d} h^i(\mathcal{U}_{n,1}, A_i))_{i=1}^k \stackrel{\mathcal{D}}{=} (h^i(\mathcal{U}_{n,n^{1/d}}, n^{1/d} A_i))_{i=1}^k.$$

Many of the applications that we shall consider are concerned with functionals of graphs of the form $G := G(\mathcal{X})$ defined for each locally finite point set $\mathcal{X} \subset \mathbb{R}^d$ (a locally finite subset of $\mathbb{R}^d$ is one with no limit point), where $G(\mathcal{X})$ has vertex set $\mathcal{X}$. See Sections 3.2 and 3.3 for examples.

We shall say G is *translation invariant* if translation by y is a graph isomorphism from $G(\mathcal{X})$ to $G(\tau_y(\mathcal{X}))$ for all $y \in \mathbb{R}^d$ and all locally finite point sets $\mathcal{X}$. We shall say G is *scale invariant* if $G(a\mathcal{X})$ is isomorphic to $G(\mathcal{X})$ for all $\mathcal{X}$ and all $a > 0$.

We use the following notion of stabilization for these graphs. Given G, and given a vertex $x \in \mathcal{X}$, let $\mathcal{E}^+(x; \mathcal{X})$ be the set of edges of $G(\mathcal{X})$ which are not edges of $G(\mathcal{X} \setminus \{x\})$, and let $\mathcal{E}^-(x; \mathcal{X})$ be the set of edges of $G(\mathcal{X} \setminus \{x\})$ which are not edges of $G(\mathcal{X})$. Let $\mathcal{P}^0 := \mathcal{P} \cup \{\mathbf{0}\}$. Our stabilization condition for graphs is that there exists an almost surely finite random variable R such that

$$\mathcal{E}^+(\mathbf{0}; \mathcal{P}^0) = \mathcal{E}^+(\mathbf{0}; (\mathcal{P}^0 \cap B_R(\mathbf{0})) \cup \mathcal{A}) \tag{2.26}$$

and

$$\mathcal{E}^-(\mathbf{0}; \mathcal{P}^0) = \mathcal{E}^-(\mathbf{0}; (\mathcal{P}^0 \cap B_R(\mathbf{0})) \cup \mathcal{A}) \tag{2.27}$$

for all finite $\mathcal{A} \subset \mathbb{R}^d \setminus B_R(\mathbf{0})$.

The stabilization conditions (2.26), (2.27) say that the local behavior of the graph in a bounded region is unaffected by points beyond a finite (though possibly random) distance from that region. As we shall see, the minimal spanning tree (with the definition suitably extended from finite to locally finite point sets) and the k-nearest neighbours, and sphere of influence graphs all satisfy (2.26) and (2.27).

Another technical condition that turns out to be relevant to stabilization is *uniqueness of the infinite component* for $G(\mathcal{P})$ and for $G(\mathcal{P}^0)$. For a locally finite point set $\mathcal{X}$, we say that uniqueness of the infinite component holds for $G(\mathcal{X})$ if there is almost surely at most a single infinite component of $G(\mathcal{X})$.

Given G, we consider three types of functional based on G. Firstly, we consider the number of components of $G(\mathcal{X})$ with at least one vertex in A, which we denote $K^G(\mathcal{X}, A)$.

Second, functionals such as total length of edges in A, number of edges in A, or number of edges in A of less than some specified length may be interpreted as a total of ϕ-weighted edge lengths in A, i.e., as a sum

$$L_\phi^G(\mathcal{X}, A) := \frac{1}{2} \sum_{x \in \mathcal{X} \cap A} \sum_{e=\{x,y\} \in G(\mathcal{X})} \phi(|e|), \tag{2.28}$$

for some appropriately specified function $\phi : [0, \infty) \to [0, \infty)$.

Third, we consider functionals such as the number of vertices of some specified degree, which are obtained by summing over all vertices some function of the local graph landscape of G (not the edge lengths) at that vertex. To make this precise, let $\mathcal{K}$ denote the set of unlabelled connected rooted graphs (i.e., connected graphs with a single vertex distinguished and denoted the root), and for $\kappa \in \mathbb{N}$ let $\mathcal{K}_\kappa$ denote the set of graphs in $\mathcal{K}$ which have all vertices a graph distance at most κ from the root. In the sequel, let $\kappa \in \mathbb{N}$ and ψ be an arbitrary bounded function from $\mathcal{K}_\kappa$ to $\mathbb{R}$. For any vertex x of any locally finite point set $\mathcal{X} \subset \mathbb{R}^d$, let $G_x(\mathcal{X})$ denote the rooted subgraph of $G(\mathcal{X})$ induced by all vertices a graph distance at most κ from x (with root at x), and let

$$V_\psi^G(\mathcal{X}, A) := \sum_{x \in \mathcal{X} \cap A} \psi(G_x(\mathcal{X})).$$

Lemma 2.1 *Suppose G is translation invariant and satisfies the stabilization conditions* (2.26) *and* (2.27). *Then if we set $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$, the stabilization conditions* (2.12) *and* (2.13) *hold. If instead we set $h(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ then, again, conditions* (2.12) *and* (2.13) *hold.*

Suppose in addition that uniqueness of the infinite component holds for $G(\mathcal{P})$ and for $G(\mathcal{P}^0)$; then if we set $h(\mathcal{X}, A) = K^G(\mathcal{X}, A)$, then the stabilization conditions (2.12) *and* (2.13) *hold.*

As we shall see in examples below, one can use Lemma 2.1 to check the applicability of Theorems 2.2, 2.3 and 2.4 to a variety of point process functionals based on stabilizing graphs.

2.4 Marked point processes

In the application in Section 3.4, we need to consider the extension of the results of the preceding section to functionals of *marked* point processes with marks in the unit interval. A *marked point set* in $\mathbb{R}^d$ is a locally finite subset of $\mathbb{R}^d \times [0,1]$ with no two elements having the same coordinate projection onto $\mathbb{R}^d$.

For a marked point set $\{(x_i, t_i)_{i\geq 1}\}$ in $\mathbb{R}^d$, with each $x_i \in \mathbb{R}^d$ and $t_i \in [0,1]$, we abuse notation slightly and write $\mathcal{X}$ for this marked point set, with $\mathcal{X} = \{x_i\}$ a subset of $\mathbb{R}^d$ but with each element x_i of $\mathcal{X}$ viewed as carrying a mark t_i. Then all the notions and results of the previous section carry through, as we now describe.

For $y \in \mathbb{R}^d$, the translation operator τ_y on marked point sets in $\mathbb{R}^d$ is to be understood to preserve the values of all marks. Then the notion of a (marked) point process set function $h(\mathcal{X}, A)$, defined for finite marked point sets $\mathcal{X}$ in $\mathbb{R}^d$ and for $A \in \mathcal{R}(\mathbb{R}^d)$, is as given at the start of Section 2.3. Also, the notion of translation invariance of a graph $G(\mathcal{X})$ is as defined in Section 2.3. When we consider edge lengths and so on, the vertex set of $G(\mathcal{X})$ is still viewed as a subset of $\mathbb{R}^d$, not $\mathbb{R}^{d+1}$. Also, it is to be understood that scalar multiplication operator $\mathcal{X} \mapsto a\mathcal{X}$ on marked point sets in $\mathbb{R}^d$, seen in the homogeneity condition (2.23) for example, leaves all marks unchanged.

In the marked setting, the points of the d-dimensional point processes $\mathcal{P}_t$ and $\mathcal{U}_{m,t}$ are to be understood to carry marks which are each uniformly distributed on $[0,1]$ and independent. Also, the inserted point at $\mathbf{0}$, when defining add one costs such as at (2.11), is assumed to carry an independent mark which is also uniformly distributed on $[0,1]$. The stabilization conditions (2.12) and (2.13) are to be understood to hold for any choice of values for the marks of points in A. Likewise the uniform bound (2.16) is to be understood to hold for any choice of the marks on $\mathcal{X}$.

With these interpretations, all of the results in Section 2.3 remain valid for marked point set functionals on the marked point processes and stabilizing graphs on the marked point processes.

3 Applications of the general results

3.1 Percolation

Let $E = \{0,1\}$, let $\mathcal{E}$ be the power set of E, with $P_0(\{1\}) = p$ and $P_0(\{0\}) = 1-p$, $p \in (0,1)$ a fixed parameter. Let $X = (X_z)_{z\in\mathbb{Z}^d}$ and $X^{\mathbf{0}}$ (the same as X but with X_0 resampled) be as described in Section 2.2 with this choice of $(E, \mathcal{E}, P_0)$. Let the sets $\mathcal{O}$, $\mathcal{O}'$ (the random set of 'occupied sites' induced by X and by $X^{\mathbf{0}}$ respectively) be given by

$$\mathcal{O} := \{z \in \mathbb{Z}^d : X_z = 1\}, \qquad \mathcal{O}' := \{z \in \mathbb{Z}^d : X^{\mathbf{0}}_z = 1\}$$

(so that $\mathcal{O}' \triangle \mathcal{O}$ is either the empty set or the set $\{\mathbf{0}\}$).

For any subset S of $\mathbb{Z}^d$, let $G(S)$ be the graph with vertex set $\{z \in S : X_z = 1\}$, and with edges between each pair vertices at unit Euclidean distance from each other. Then $G(\mathcal{O})$ is a Bernoulli site percolation process with parameter p (bond percolation versions of the results in this section also hold, and are proved by similar means taking $E = \{0,1\}^d$; see [11], page 1517).

For background information on percolation see Grimmett [6]. Let p_c be the critical value of p, i.e., the supremum of the set of p for which the components of $G(\mathcal{O})$ are a.s. all finite. Provided $d \geq 2$, it is known that $0 < p_c < 1$.

By the uniqueness of the infinite cluster in percolation (see, e.g., [6]), there is almost surely at most a single infinite component of $G(\mathcal{O})$. For later use, we denote the vertex set of this infinite component of $G(\mathcal{O})$ by C_∞ (possibly the empty set), and denote the vertex set of this infinite component of $G(\mathcal{O}')$ by C'_∞.

Also for later use, observe that for $y, z \in \mathbb{Z}^d$,

$$(\tau_y X)_z = 1 \Longleftrightarrow y + z \in \mathcal{O} \Longleftrightarrow z \in \tau_{-y}(\mathcal{O}). \tag{3.1}$$

We shall give two applications of Theorem 2.1 to percolation. The next result says that the number of components of $G(\mathcal{O} \cap tB_0)$ in disjoint subregions of a large region tB_0 are asymptotically normal and asymptotically independent of each other. It is of particular interest in the case when $p = p_c$ since in this case, correlations are not believed to decay exponentially.

For percolation and also for some of the other spatial graphs that we consider, there are several ways to count the 'number of components' in a subregion A of $\mathbb{R}^d$, since one has to decide whether to include components that lie only partially in A. In results given here, such components are counted fully, but the same results should hold if they were only counted partially, or not at all.

Theorem 3.1 *For $t \geq 1$, $A \in \mathcal{R}(B_0)$, let $H_t(A)$ be the number of components of $G(\mathcal{O} \cap tB_0)$ which include at least one vertex in tA.*

Define the random variable Δ_∞^H as follows. If $X_0 = 0$ and $X_* = 1$, then let Δ_∞^H be equal to the number of components of $G(\mathcal{O})$ adjacent to the origin, minus 1. If $X_0 = 1$ and $X_* = 0$, then let $-\Delta_\infty^H$ be equal to the number of components of $G(\mathcal{O}')$ adjacent to the origin, minus 1. If $X_0 = X_*$, then set $\Delta_\infty^H = 0$.

Then the conditions (2.3), (2.4) *and* (2.5) *for Theorem* 2.1 *are satisfied, and therefore the conclusions* (2.8) *and* (2.9) *of that result are valid* (*with $H^i = H$ for all i.*)

Proof. Let $t \geq 1$, $y \in \mathbb{Z}^d$, and $A \in \mathcal{R}(B_0)$. By (3.1), $H_{t,y}(A)$ is the number of components of $G((\tau_{-y}\mathcal{O}) \cap tB_0)$ which intersect tA (i.e., contain at least one vertex in tA). Hence, $H_{t,y}(A)$ is the number of components of $G(\mathcal{O} \cap \tau_y(tB_0))$ which intersect $\tau_y(tA)$.

Thus $-\Delta_{t,y}^H(A)$ is the increment in the the number of components of $G(\mathcal{O} \cap \tau_y(tB_0))$ which intersect $\tau_y(tA)$ when we resample $X_{\mathbf{0}}$ (i.e., when we replace the process X by $X^{\mathbf{0}}$).

With Δ^H_∞ defined in the statement of the theorem, we assert that (2.3) and (2.4) hold. To verify (2.3), suppose that $\liminf_{n\to\infty}(\tau_{y_n}(t_nA)) = \mathbb{R}^d$. Suppose first that $X_0 = 0$ and $X_* = 1$. Then there exists a (random) N_1 such that for $n \geq N_1$, every pair of vertices lying adjacent to the origin and in the same component of $G(\mathcal{O})$, is connected by an path in $G(\mathcal{O})$, all of whose vertices lie in $\tau_{y_n}(t_nA)$. Then for all $n \geq N_1$, $\Delta^H_{t_n,y_n}(A) = \Delta^H_\infty$ as described above. A similar argument applies in the case with $X_0 = 1$ and $X_* = 0$, and for other cases clearly $H_{t_n,y_n}(A) = 0$ for all n. Thus (2.3) holds.

Next, suppose $\liminf(\tau_{t_n,y_n}(B_0 \setminus A)) = \mathbb{R}^d$. Suppose $X_0 = 0$. There exists a random N_2 such that for all large enough $n \geq N_2$, the set $\tau_{t_n,y_n}(B_0 \setminus A)$ contains all finite components of $G(\mathcal{O})$ lying adjacent to the origin.

There exists a random N_3 such that for all $n \geq N_3$, each pair of vertices of C_∞ which lie adjacent to the origin is connected by a path in $G(\mathcal{O})$ all of whose vertices lie in the set $\tau_{y_n}(t_n(B_0 \setminus A))$. We assert that if $n \geq \max(N_2, N_3)$, changing of the value of $X_{\mathbf{0}}$ from 0 to 1 does not affect the number of components of $G(\mathcal{O} \cap \tau_{y_n}(t_nB_0))$ that intersect $\tau_{y_n}(t_nA)$. This is because for n this big, any two occupied vertices adjacent to the origin which are both connected by paths in $G(\mathcal{O})$ to vertices in $\tau_{y_n}(t_nA)$, must be part of C_∞ and therefore are connected by a path which avoids the origin, so that they are already part of the same component of $G(\mathcal{O} \cap \tau_{y_n}(t_nB_0))$ even before we add a vertex at the origin to $\mathcal{O}$. In other words, $G(\mathcal{O} \cap \tau_{y_n}(B_0))$ has at most a single component which intersects both the set $\tau_{y_n}(t_nA)$ and and the set of sites adjacent to the origin; see Figure 1. The assertion follows, and one argues similarly for $X_* = 0$. Thus (2.4) holds.

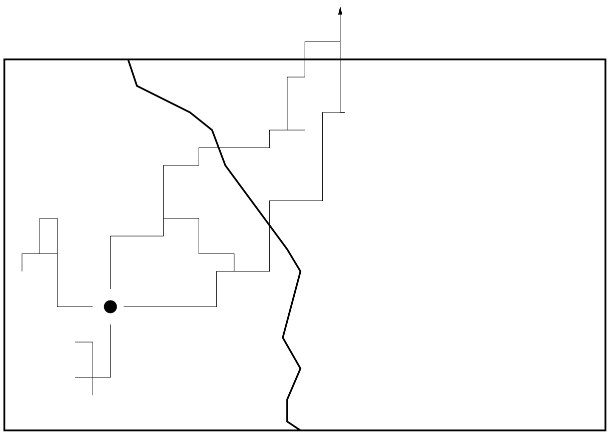

Figure 1: The rectangle represents the region $\tau_{y_n}(t_nB_0)$, and the region to the right of the jagged line is $\tau_{y_n}(t_nA)$. The other rectilinear lines indicate the components of $G(\mathcal{O})$ adjacent to the origin.

Since all vertices in $G(\mathbb{Z}^d)$ have degree $2d$, the absolute value of $\Delta_{t,y}(A)$ is uniformly bounded by $2d - 1$, and therefore the moments condition (2.5) is valid for any finite γ. Therefore Theorem 2.1 is applicable here. ■

We now consider the **largest component** of $G(\mathcal{O} \cap tB_0)$, i.e., the one with the most vertices (there could be more than one largest component). The following result says that the distribution of the vertices of a largest component in tB_0 is asymptotically a white noise distribution. In it, we assume B_0 is rectangular, i.e., that B_0 is a product of intervals. Presumably, the proof can be extended to other shapes of B_0.

Theorem 3.2 *Suppose that $p > p_c$, and that B_0 is rectangular. For $t \geq 1$ and $A \in \mathcal{R}(B_0)$, let $H_t(A)$ be the number of vertices of tA which lie in a largest component of $G(\mathcal{O} \cap tB_0)$. Set*

$$\Delta_\infty^H = \begin{cases} \mathrm{card}(C_\infty \setminus C'_\infty) & \text{if} \quad X_{\mathbf{0}} = 1, X_* = 0, \\ -\mathrm{card}(C'_\infty \setminus C_\infty) & \text{if} \quad X_{\mathbf{0}} = 0, X_* = 1, \\ 0 & \text{if} \quad X_{\mathbf{0}} = X_* \end{cases}$$

Then the conditions (2.3), (2.4) *and* (2.5) *for Theorem* 2.1 *are satisfied, and therefore the conclusions* (2.8) *and* (2.9) *of that result are valid* (*with all $H^i = H$.*)

Proof. Observe first that Δ_∞^H is indeed almost surely finite. For example, if $X_{\mathbf{0}} = 1$ and $X_* = 0$, then $C'_\infty \subseteq C_\infty$ and $C_\infty \setminus C'_\infty$ consists (in the case where $\mathbf{0} \in C_\infty$) of those finite components of $G(\mathcal{O}')$ lying adjacent to the origin, along with the origin itself (with $C_\infty \setminus C'_\infty = \emptyset$ in the case where $\mathbf{0} \notin C_\infty$).

Observe also that for any $t \geq 1$, $y \in \mathbb{Z}^d$ and $A \in \mathcal{R}(\mathcal{B}_0)$, by (3.1), $H_{t,y}(A)$ is the number of vertices in tA in a largest component of $G(\tau_{-y}(\mathcal{O}) \cap tB_0)$, and so is the number of vertices in $\tau_y(tA)$ in a largest component of $G(\mathcal{O} \cap \tau_y(tB_0))$.

In what follows, a few plausible (and actually true) facts are stated without proof. For details of their proofs, see [11].

Suppose that $\liminf_{n\to\infty}(\tau_{y_n}(t_n B_0)) = \mathbb{R}^d$. Then with probability tending to 1, the largest component of $G(\mathcal{O} \cap \tau_{y_n}(t_n B_0))$ is the largest component of $G(C_\infty \cap \tau_{y_n}(t_n B_0))$, and if $\mathbf{0} \in C_\infty$ then the largest component of $G(\mathcal{O} \cap \tau_{y_n}(t_n B_0))$ is the component of $G(C_\infty \cap \tau_{y_n}(t_n B_0))$ containing the origin.

Suppose that $\liminf_{n\to\infty}(\tau_{y_n}(t_n A)) = \mathbb{R}^d$. Then with probability 1, the set $C_\infty \triangle C'_\infty$ is contained in $\tau_{y_n}(t_n A)$ for all large enough n. Hence, the probability that $\Delta_{t_n,y_n}(A)$ is equal to Δ_∞^H defined above tends to 1, and so (2.3) holds.

Suppose that $\liminf_{n\to\infty}(\tau_{y_n}(t_n(B_0 \setminus A))) = \mathbb{R}^d$. Then with probability 1, the set $C_\infty \triangle C'_\infty$ is contained in $\tau_{y_n}(t_n(B_0 \setminus A))$ for all large enough n. If $C_\infty \triangle C'_\infty \subseteq \tau_{y_n}(t_n(B_0 \setminus A))$ and also the largest component of $G(\mathcal{O} \cap \tau_{y_n}(t_n B_0))$ is the component of $G(C_\infty \cap \tau_{y_n}(t_n B_0))$ containing the origin, then changing the value of $X_{\mathbf{0}}$ from 1 to 0 will not remove any vertices of the largest component lying in $\tau_{y_n}(t_n A)$ so that it does not change the value of $H_{t_n,y_n}(A)$. Hence the probability that $\Delta_{t_n,y_n}^H(A)$ is equal to zero tends to 1, and so (2.4) holds.

We need to check the moments condition (2.5). Most of the ingredients in the proof of this are given in the proof of Theorem 3.2 of [11]. The main difference is that we now need to account for a possible *decrease* in the number of elements in $\tau_y(tA)$ of the largest component when we change the status of site $\mathbf{0}$ from 'vacant' to 'occupied' (in [11] we needed only to consider the largest component size, which by contrast really is monotone in $X_{\mathbf{0}}$). Such a decrease could happen if two ore components lying adjacent to $\mathbf{0}$, when merged, form a component larger than and disjoint from the previous largest component. However, the probability of this occurring decays exponentially in t^{d-1} (see, e.g., Theorem 4 in Penrose and Pisztora [13]), and using this we can check (2.5) here. ■

3.2 The minimal spanning tree

Given a locally finite set $\mathcal{X} \subset \mathbb{R}^d$, $d \geq 2$, and given $a > 0$, let $\mathcal{G}_a(\mathcal{X})$ be the graph with vertex set $\mathcal{X}$ and with edge set $\{\{x,y\} : |x-y| < a\}$. Let $\mathrm{MST}(\mathcal{X})$ be the graph with vertex set $\mathcal{X}$ obtained by including each edge $\{x,y\}$ such that x and y lie in different components of $\mathcal{G}_{|x-y|}(\mathcal{X})$ and at least one of these components is finite. If $\mathcal{X}$ is finite with distinct inter-point distances, then $\mathrm{MST}(\mathcal{X})$ is the *minimal spanning tree* on $\mathcal{X}$, i.e. the connected graph with vertex set $\mathcal{X}$ of minimal total edge length; see Aldous and Steele ([1], Lemma 12). Clearly $\mathrm{MST}(\mathcal{X})$ is translation and scale invariant.

The following result says that the totals of ϕ-weighted edges of the minimal spanning tree on random points in disjoint regions, scaled and centred, are asymptotically independent normals, and likewise for the totals of a local graph landscape function summed over points in disjoint regions. In this result, say ϕ is *polynomially bounded* if there exists a constant c such that $\phi(r) \leq c(1+r)^c$ for all $r > 0$.

Theorem 3.1 *Suppose* $G(\mathcal{X})$ *is* $\mathrm{MST}(\mathcal{X})$*.*

For any fixed $\kappa \in \mathbb{N}$*, and any bounded function* ψ *from the set* $\mathcal{K}_\kappa$ *of rooted graphs of radius at most* κ *to* $\mathbb{R}$*, if we set* $h(\mathcal{X}) = V_\psi^G(\mathcal{X}, A)$ *then* h *satisfies all the conditions* (2.12)*,* (2.13)*,* (2.15)*,* (2.16)*, and* (2.23) *(with* $\gamma = 0$*) of Theorems* 2.2*,* 2.3*, and* 2.4 *and therefore satisfies their conclusions* (2.18)*,* (2.19)*,* (2.21) (2.22)*,* (2.24) *and* (2.25) *(with* $h^j(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ *for all* j *and with* $\gamma = 0$*).*

Suppose instead that we set $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$*. Then the stabilization conditions* (2.12) *and* (2.13) *hold. If* ϕ *is uniformly bounded, or if* ϕ *is polynomially bounded and* B_0 *is convex, then the moments condition* (2.15) *holds and so Theorems* 2.2 *and* 2.3 *apply and their conclusions* (2.18)*,* (2.19)*,* (2.21) *and* (2.22) *(with* $h^j = h$ *for all* j*) hold.*

If also $\phi(r) = r^\alpha$ *for some constant* α*, then the homogeneity hypothesis* (2.23) *holds and hence the conclusions* (2.24) *and* (2.25) *of Theorem* 2.4 *are valid with* $\gamma = \alpha$*.*

The proof of this uses the following lemma which we shall use again later.

Lemma 3.1 *If* B_0 *is convex then*

$$\inf_{x \in B_0, r \in (0,1]} r^{-d} |B_r(x) \cap B_0| > 0. \tag{3.1}$$

Proof. Since B_0 is assumed to have non-empty interior, there exists a ball B contained in B_0. For any $x \in B$, the convex hull of $\{x\} \cup B$ is contained in B, and since B_0 is bounded the angle subtended by this cone-like set at x is bounded away from zero, so the result follows. ■

Proof of Theorem 3.1. Condition (2.26) follows from Lemma 2.1 of [15]. Condition (2.27) is more complicated but follows from the proof of Proposition 1 of Lee [9]. Therefore, Lemma 2.1 applies to give us the conditions (2.12), (2.13) in the case where either $h(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ or $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$.

Given a finite set $\mathcal{X} \subset \mathbb{R}^d$, consider the effect on the minimal spanning tree $\mathrm{MST}(\mathcal{X})$ of adding a point at the origin $\mathbf{0}$. Let edges of $\mathrm{MST}(\mathcal{X} \cup \{\mathbf{0}\})$ that are not in $\mathrm{MST}(\mathcal{X})$ be denoted *added edges*, and let edges of $\mathrm{MST}(\mathcal{X})$ that are not in $\mathrm{MST}(\mathcal{X} \cup \{\mathbf{0}\})$ be denoted *deleted edges*.

By the revised add and delete algorithm of Lee [8], the added edges are precisely those incident to $\mathbf{0}$ in $\mathrm{MST}(\mathcal{X} \cup \{\mathbf{0}\})$, and there are fewer deleted edges than added edges. Moreover, there is a uniform non-random bound on vertex degrees in the minimal spanning tree (see [1]), and hence there is a uniform bound both on the number of added edges and on the number of deleted edges. The moments condition (2.15) for $h(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ is immediate from these remarks. Moreover, if ϕ is bounded, then (2.15) for $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$ also follows from these remarks.

Suppose that ϕ is polynomially bounded and B_0 is regular. By the preceding remarks, to prove (2.15) in this case, it suffices to show that for any $K > 0$, there is a deterministic uniform bound on the Kth moment of the length of the longest added edge when a point at $\mathbf{0}$ is inserted into $\mathcal{U}_{m,t}$ with $t \geq 1$, $\mathbf{0} \in \tau_x(tB_0)$ and $m/(\lambda t^d |B_0|)$ in the range $[1/2, 3/2]$, and likewise for the the longest deleted edge.

We assert that the longest deleted edge in a finite set $\mathcal{X}$ is at most twice as long as the longest added edge. To see this, suppose that $\{X, Y\}$ is a deleted edge. Then, since all added edges are incident to the added point at $\mathbf{0}$, and there must be a path from X to Y in $\mathrm{MST}(\mathcal{X} \cup \{\mathbf{0}\})$, there exist points X', Y' in $\mathcal{X}$ such that X', Y' are both adjacent to $\mathbf{0}$ in $G(\mathcal{X} \cup \{\mathbf{0}\})$, and such that there a path in $\mathrm{MST}(\mathcal{X})$ from X to X', and a path in $\mathrm{MST}(\mathcal{X})$ from Y to Y'. By the triangle inequality, $|X' - Y'|$ is at most twice the length of the longest added edge, and also $|X - Y| \leq |X' - Y'|$ since otherwise we could start with $\mathrm{MST}(\mathcal{X})$, then replace edge $\{X, Y\}$ by $\{X', Y'\}$ to obtain a spanning tree on $\mathcal{X}$ of smaller total length, a contradiction. This completes the proof of the assertion.

Thus, to check (2.15) for $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$ when ϕ is polynomially bounded and B_0 is regular, it suffices to prove the Kth moment of the longest added edge is uniformly bounded. This can be proved by an argument along the lines of Lemma 2.1 of Yukich [21] (using convexity of B_0, and (3.1)).

The homogeneity condition (2.23), for V_ψ^G or L_ϕ^G with $\phi(r) = r^\alpha$, follows from the fact that the graph $\mathrm{MST}(\mathcal{X})$ is scale invariant. ■

Remark. In the univariate case Kesten and Lee ([7], Theorem 2) give a CLT for

$L^G_\phi(B_0)$ when ϕ is monotonically increasing. Our result gives central limit theorems, for this univariate case, without this restriction on ϕ.

Remark. Consider the empirical distribution of MST edge lengths. That is, let $N_s(\mathcal{X})$ be the number of edges of $\mathrm{MST}(\mathcal{X})$ of length less than s. For arbitrary B_0, the finite-dimensional distributions of the scaled, centred empirical process

$$(t^{-d/2}(N_s(\mathcal{P}_t) - EN_s(\mathcal{P}_t)), \quad s > 0) \tag{3.2}$$

converge as $t \to \infty$ to those of a Gaussian process. Moreover, the finite-dimensional distributions of the corresponding scaled, centred empirical process for a binomial sample, namely

$$(n^{-1/2}(N_{n^{-1/d}s}(\mathcal{U}_{n,1}) - EN_{n^{-1/d}s}(\mathcal{U}_{n,1})), \quad s > 0) \tag{3.3}$$

converge as $t \to \infty$ to those of another Gaussian process.

In the case of the first empirical process (3.2), this follows by taking an arbitrary set of positive 'times' $s_1, \ldots, s_k$ and applying Theorem 2.2 with $h^i(\mathcal{X}, A) = L^G_{\phi^i}(\mathcal{X}, A)$ and $\phi^i(r) = \mathbf{1}_{\{r \leq s_i\}}$ for $(1 \leq i \leq k)$. The limiting Gaussian process in this case has covariance function given by the function $E[Y_s Y_{s'}], s, s' > 0$, where we set

$$Y_s = E[\delta_s(\infty)|\mathcal{F}], \tag{3.4}$$

with $\delta_s(\infty)$ denoting the stabilizing limit for the functional $h(\mathcal{X}, A) = L^G_{\phi_s}(\mathcal{X}, A)$ with $\phi_s(r) := \mathbf{1}_{\{r \leq s\}}$, and with $\mathcal{F}$ denoting the σ-field generated by the Poisson configuration in the half-space $\{x = (x^1, \ldots, x^d) \in \mathbb{R}^d : x_1 < 0\}$.

In the case of the second empirical process (3.3), the convergence of finite-dimensional distributions follows similarly but this time using Theorem 2.3 and the observation that $N_s(\mathcal{U}_{n,n^{1/d}}) = N_{n^{-1/d}s}(\mathcal{U}_{n,1})$. The limiting covariance function is this time given by $\mathrm{Cov}(Y_s, Y_{s'}), s, s' > 0$, with Y_s given once more by (3.4), but now with $\lambda = |B_0|^{-1}$.

3.3 Nearest-neighbour type graphs

Let $k \in \mathbb{N}$. The k-nearest neighbour graph (k-NNG) on a locally finite set $\mathcal{X} \subset \mathbb{R}^d$ is obtained by including an undirected edge connecting each vertex $x \in \mathcal{X}$ to each of its k nearest neighbours (using the lexicographic ordering as a tie-breaker in the event of a tie). We also consider the sphere of influence graph (SIG), in which, denoting the distance from $x \in \mathcal{X}$ to its nearest neighbour by R_x, we connect vertices x, y of $\mathcal{X}$ by an edge if and only if $B_{R_x}(x) \cap B_{R_y}(y) \neq \emptyset$.

White noise limits for functionals such as L^G_ϕ and V^G_ψ (defined in Section 2.3) can be derived using either the results in this paper or by other methods based on exponential decay, as in [2]. We concentrate here on the component count $K^G(\mathcal{X}, A)$, for which exponential decay is not so clear.

Theorem 3.2 *Suppose $G(\mathcal{X})$ is defined to be the k-NNG on $\mathcal{X}$. Then $h(\mathcal{X}, A) = K^G(\mathcal{X}, A)$ satisfies all the conditions* (2.12)*,* (2.13)*,* (2.15)*,* (2.16)*, and* (2.23) (*with $\gamma = 0$*) *for Theorems* 2.2*,* 2.3*, and* 2.4*. Therefore, $h(\mathcal{X}, A) = K^G(\mathcal{X}, A)$ satisfies the conclusions* (2.18)*,* (2.19)*,* (2.21)*,* (2.22)*,* (2.24)*, and* (2.25) *in those results* (*with $h^i = h$ for all i, and with $\gamma = 0$*)*.*

Proof. There is a deterministic uniform upper bound on the degree of vertices in $G(\mathcal{X})$ (see, e.g., Lemma 8.4 of Yukich [20]), and hence a uniform deterministic bound on the change in the number of components of $G(\mathcal{X})$ caused by inserting a single point; the moments condition (2.15) follows.

We can obtain the stabilization conditions (2.12) and (2.13) by using Lemma 2.1. This result is applicable because conditions (2.26) and (2.27) hold by the proof of Lemma 6.1 of [14], while uniqueness of the infinite component in $G(\mathcal{P})$ and $G(\mathcal{P}^0)$ holds by Lemma 6.4 of [14]. The uniform polynomial bound (2.16) and the homogeneity (2.23) are obvious. ∎

Theorem 3.3 *Suppose B_0 is regular. Suppose $G(\mathcal{X})$ is defined to be the SIG on $\mathcal{X}$. Then $h(\mathcal{X}, A) = K^G(\mathcal{X}, A)$ satisfies all the conditions* (2.12)*,* (2.13)*,* (2.15)*,* (2.16)*, and* (2.23) (*with $\gamma = 0$*) *for Theorems* 2.2*,* 2.3*, and* 2.4*. Therefore, $h(\mathcal{X}, A) = K^G(\mathcal{X}, A)$ satisfies the conclusions* (2.18)*,* (2.19)*,* (2.21)*,* (2.22)*,* (2.24)*, and* (2.25) *in those results* (*with $h^i = h$ for all i, and with $\gamma = 0$*)*.*

Proof. We need to check the moments condition (2.15). This can be done as in the proof of Theorem 7.4 of [14]. The 'regularity' condition in that result is implied by the condition (3.1) here; see the remarks on page 1010 of [14].

The rest of the argument is similar to that for Theorem 3.2. For conditions (2.26) and (2.27), see the proof of Lemma 7.1 of [14]. For uniqueness of the infinite component of $G(\mathcal{P})$, see Theorem 7.3 of [14]. The infinite component of $G(\mathcal{P}^0)$ is also unique, by the proof of Lemma 6.4 (b) of [14]. ∎

3.4 The on-line nearest neighbour graph

Suppose $X_1, X_2, \ldots, X_n$ are points in $\mathbb{R}^d$. In the *on-line nearest neighbour graph* (or on-line NNG for short), the points are assumed to arrive sequentially and each point X_i, $i \geq 2$, is connected by an undirected edge to its nearest neighbour in the set of preceding points in the sequence $\{X_1, \ldots, X_{i-1}\}$ (using the lexicographic ordering on $\mathbb{R}^d$ to break any ties). The resulting graph is a tree, which we will denote the *on-line nearest neighbour graph* on the sequence of points $(X_1, X_2, \ldots, X_n)$. One could also consider the on-line k-nearest neighbour graph defined analogously. In our case, the points in the sequence will be random, being independent and uniformly distributed over B_0 or over tB_0.

The on-line nearest neighbour graph on random points is a natural growth model for spatial random graphs, although it was apparently introduced only recently, by

Berger et al. ([3], Section 3). There, the motivation comes from the search for a simple model of scale-free networks, the graph being itself a simplification of a model of Fabrikant et al. [5].

To fit this graph into our present setup, consider a *marked* random finite point set $\mathcal{X}$ in $\mathbb{R}^d$, where each point X of $\mathcal{X}$ carries a random mark T_X which is uniformly distributed on $[0,1]$, independent of the other marks and of the point process $\mathcal{X}$. The points are listed in increasing order of mark, i.e., the marks represent time of arrival. With this ordering, we connect each point of $\mathcal{X}$ to the nearest point that precedes it in the ordering, to obtain a graph which we also call the on-line nearest neighbour graph on the marked point set $\mathcal{X}$. This definition extends to infinite but locally finite point sets.

Clearly the on-line NNG on $\mathcal{U}_{m,t}$ (defined via marked point processes) has the same distribution as the on-line NNG (with the first definition) on a sequence $X_1, \ldots, X_m$ of independent uniform points on tB_0. Likewise, the on-line NNG for $\mathcal{P}_t$ (defined via marked point processes) has the same distribution as the on-line NNG on $(X_1, \ldots, X_N)$ with $\{X_i\}$ independent uniform points on tB_0 and N an independent Poisson variable with parameter $\lambda t^d |B_0|$.

As it turns out, the on-line nearest neighbour graph is a nice example of our methods because it is stabilizing but has only polynomially decaying correlations. In the following discussion, although we think of the graph as undirected, we shall refer to an edge connecting marked points X and Y with $T_X > T_Y$, as an *outgoing edge* from X and an *incoming edge* to Y. Each vertex (except one if $\mathcal{X}$ is finite) has a single outgoing edge.

The existence of an almost surely finite radius of stabilization satisfying (2.26) and (2.27) will be shown later on. To see that its distribution does not have an exponentially decaying tail, let L be the length of the outgoing edge from the origin in the on-line nearest neighbour graph on $\mathcal{P}^{\mathbf{0}}$. With π_d denoting the volume of the unit ball in $\mathbb{R}^d$, we have

$$P[L \geq r] = \int_0^1 \exp(-\lambda \pi_d r^d t) dt = \lambda^{-1} \pi_d^{-1} r^{-d} \int_0^{\lambda \pi_d r^d} e^{-u} du,$$

which shows that the tail of the distribution of L decays only polynomially, and L is clearly a lower bound for any radius of stabilization.

The following result says that for certain ϕ the totals of ϕ-weighted edges of the minimal spanning tree on random points in disjoint regions, scaled and centred, are asymptotically independent normals, and likewise for the totals of any bounded function of vertex degrees summed over vertices in disjoint regions.

Theorem 3.4 *Suppose $G(\mathcal{X})$ is the on-line NNG on $\mathcal{X}$. Assume B_0 is convex. For any bounded function ψ from the set $\mathcal{K}_1$ of rooted graphs of radius at most 1 to $\mathbb{R}$, if we set $h(\mathcal{X}) = V_\psi^G(\mathcal{X}, A)$ then h satisfies all the conditions* (2.12), (2.13), (2.15), (2.16), *and* (2.23) (*with $\gamma = 0$*) *of Theorems* 2.2, 2.3, *and* 2.4 *and therefore satisfies* *their conclusions* (2.18), (2.19), (2.21) (2.22), (2.24) *and* (2.25) (*with* $h^j(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ *for all* j *and with* $\gamma = 0$).

Suppose instead that we set $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$. *Then the stabilization conditions* (2.12) *and* (2.13) *hold. If* ϕ *satisfies the growth bound*

$$\sup_{r>0} \left((1+r)^{-\alpha}\phi(r)\right) < \infty, \qquad \text{some } \alpha < d/4, \tag{3.5}$$

then the moments condition (2.15) *holds and so Theorems* 2.2 *and* 2.3 *apply and their conclusions* (2.18), (2.19), (2.21) *and* (2.22) (*with* $h^j = h$ *for all* j) *hold.*

If also $\phi(r) = r^\alpha$ *for some constant* $\alpha < d/4$, *then the homogeneity hypothesis* (2.23) *holds and hence the conclusions* (2.24) *and* (2.25) *of Theorem* 2.4 *are valid with* $\gamma = \alpha$.

Remarks. Provided B_0 is convex, Theorem 3.4 gives, among other things, a central limit theorem for the number of vertices of any fixed degree in the on-line NNG on $\mathcal{U}_{n,1}$ or on $\mathcal{P}_t$. Since any bounded function of the vertex lengths satisfies the growth bound (3.5), Theorem 3.4 also enables us to obtain similar functional central limit theorems results on the empirical distributions of edge lengths in the on-line nearest neighbour graph, to those described in the preceding section for the minimal spanning tree.

Provided $d > 4$, Theorem 3.4 gives us a central limit theorem for the total length of the on-line NNG on $\mathcal{U}_{n,1}$ or on $\mathcal{P}_t$ (since the function $\phi(r) = r$ satisfies the growth bound (3.5)).

This leaves open the question of the asymptotic behaviour of the total length of the on-line NNG on $\mathcal{U}_{n,1}$, in dimensions $d \leq 4$. As mentioned earlier, it is likely that 4th moments condition (2.15) can be replaced by a $2 + \varepsilon$ moments condition in Theorems 2.2 and 2.3. If this can be done, the total length of the on-line NNG on $\mathcal{U}_{n,1}$ will satisfy a central limit theorem for $d = 4$ or $d = 3$. We suspect that a central limit theorem also holds for $d = 2$, but do not have a proof. We believe that the limiting distribution of is non-normal for $d = 1$, and intend to address this in future work.

Theorem 3.4 also carries through to the on-line k-nearest neighbour graph, although we give a proof only in the case $k = 1$.

The proof of Theorem 3.4 uses the following two lemmas. The first of these is geometric in nature. Given distinct points $x, y \in \mathbb{R}^d$, let $C_{x,y}$ denote the cone with its point at x and with angular radius $\pi/12$, centred on the half-line from x passing through y.

Lemma 3.2 *Suppose* B_0 *is convex. Then*

$$\inf_{x,y \in B_0 : x \neq y} \frac{|C_{x,y} \cap B_{|y-x|}(x) \cap B_0|}{|y-x|^d} > 0.$$

Proof. Take $x, y \in B_0$ with $|y-x| = r > 0$. Let $z = (x+y)/2$. By convexity $z \in B_0$, and geometrical considerations show that

$$B_{(r/2)\sin(\pi/12)}(z) \subseteq C_{x,y} \cap B_r(x),$$

so the result follows from Lemma 3.1. ∎

Recall from Section 2.3 that if a graph $G(\mathcal{X})$ is defined for locally finite point sets $\mathcal{X} \subset \mathbb{R}^d$, then for $x \in \mathcal{X}$, $\mathcal{E}^+(x;\mathcal{X})$ denotes the set of edges of $G(\mathcal{X})$ which are not edges of $G(\mathcal{X} \setminus \{x\})$, and $\mathcal{E}^-(x;\mathcal{X})$ denotes the set of edges of $G(\mathcal{X} \setminus \{x\})$ which are not edges of $G(\mathcal{X})$. Also, $|e|$ denotes the Euclidean length of edge e.

Lemma 3.3 *Let $G(\mathcal{X})$ denote the on-line nearest neighbour graph on the marked point set $\mathcal{X}$, and suppose B_0 is convex. Let $\lambda > 0$ and let $\mu_t := \lambda t^d |B_0|$. Let $0 \leq \alpha < d/4$. Then*

$$\sup_{t \geq 1, x \in tB_0} \sup_{m \in [\mu_t/2, 3\mu_t/2] \cap \mathbb{Z}} \mathbb{E}\left[\left(\sum_{e \in \mathcal{E}^+(x;\mathcal{U}_{m,t} \cup \{x\})} (1+|e|)^\alpha\right)^4\right] < \infty \tag{3.6}$$

and

$$\sup_{t \geq 1, x \in tB_0} \sup_{m \in [\mu_t/2, 3\mu_t/2] \cap \mathbb{Z}} \mathbb{E}\left[\left(\sum_{e \in \mathcal{E}^-(x;\mathcal{U}_{m,t} \cup \{x\})} (1+|e|)^\alpha\right)^4\right] < \infty. \tag{3.7}$$

Proof. Fix t, m and x with $t \geq 1$, $m/\mu_t \in [1/2, 3/2]$, and $x \in tB_0$. Take cones $C_1, C_2, \ldots, C_K$, each with angular radius $\pi/12$ with point at x, and with union $\mathbb{R}^d$, where K is a constant depending only on d. For $1 \leq i \leq K$, let C_i^+ be the cone of angular radius $\pi/6$ with point at x, concentric to C_i. Let $1 \leq i \leq K$, and let R_i be the distance from x to the nearest point Y of $\mathcal{U}_{m,t}$ lying in the cone C_i^+, and carrying a mark $T_Y < T_x$. If no such point of $\mathcal{U}_{m,t}$ exists, set R_i to be $\sup_{y \in tB_0 \cap C_i} |y-x|$, the furthest distance from x to any element of $tB_0 \cap C_i$. Note that all incoming edges to x from points of $\mathcal{U}_{m,t}$ in C_i must be from points at a distance at most R_i from x.

Let $\mathcal{Y}_i$ be the set of points of $\mathcal{U}_{m,t}$ lying in the cone C_i at a distance less than R_i from x. By definition of R_i, necessarily all points of $\mathcal{Y}_i$ carry a mark greater than T_x. Listing the points of $\mathcal{Y}_i$ as $Y_{i,1}, \ldots, Y_{i,\nu(i)}$, in order of increasing mark, let M_i be the number of points $Y_{i,j}$ of $\mathcal{Y}_i$ such that $Y_{i,j}$ lies closer to x than any of the points $Y_{i,1}, \ldots, Y_{i,j-1}$ (we include $Y_{i,1}$ in this set of points). Then each incoming edge at x with an endpoint in the cone C_i is from such a point, so that

$$\begin{aligned}\left(\sum_{e \in \mathcal{E}^+(x;\mathcal{U}_{m,t} \cup \{x\})} (1+|e|)^\alpha\right)^4 &\leq \left(\sum_{i=1}^K (1+M_i)(1+R_i)^\alpha\right)^4 \\ &\leq K^4 \sum_{i=1}^K (1+M_i)^4 (1+R_i)^{4\alpha}. \end{aligned} \tag{3.8}$$

Also, when x is inserted into $\mathcal{U}_{m,t}$, all removed edges are outgoing from points that are connected to x after it is inserted. Hence, all removed edges outgoing from points in C_i, except possibly edge from the point with lowest mark, are of length at most R_i, while the removed edge from the point with lowest mark is of length at most $2\max_{1\le i\le K} R_i$. Hence

$$\left(\sum_{e\in\mathcal{E}^-(x;\mathcal{U}_{m,t}\cup\{x\})}(1+|e|)^\alpha\right)^4 \le \left(K(1+2\max_{i\le K}R_i)^\alpha+\sum_{i=1}^K M_i(1+R_i)^\alpha\right)^4$$
$$\le \left(\sum_{i=1}^K(1+R_i)^\alpha(2^\alpha K+M_i)\right)^4$$
$$\le K^4\sum_{i=1}^K(2^\alpha K+M_i)^4(1+R_i)^{4\alpha}. \quad (3.9)$$

Conditional on the set of points $\mathcal{Y}_i$ (but not their marks), any of the $\nu(i)!$ possible orderings of the marks of points of $\mathcal{Y}$ is equally likely. Note that $\binom{M_i}{4}$ is the number of collections of four distinct points $Y_{i,j_1},\dots,Y_{i,j_4}$ such that each of Y_{i,j_k}, $1\le k\le 4$, lies closer to x than any point of $Y_{i,1},\dots,Y_{i,j_k-1}$.

Given $\nu(i)=\ell$, for any $j_1<j_2<j_3<j_4\le\ell$, the probability that each of Y_{i,j_k}, $1\le k\le 4$ lies closer to x than any point of $Y_{i,1},\dots,Y_{i,j_k-1}$ is equal to $j_1^{-1}j_2^{-1}j_3^{-1}j_4^{-1}$. Hence,

$$\mathbb{E}\left[\binom{M_i}{4}\middle|\nu(i)=\ell\right]=\sum_{1\le i_1<i_2<i_3<i_4\le\ell}\frac{1}{i_1i_2i_3i_4}\le(1+\log\ell)^4/4!,\qquad \ell>0,$$

and since trivially

$$M_i^4\le 256\binom{M_i}{4}\mathbf{1}_{\{M_i\ge 4\}}+81\mathbf{1}_{\{M_i<4\}}\le 256\binom{M_i}{4}+81,$$

we obtain

$$\mathbb{E}[M_i^4|\nu(i)=\ell]\le 81+11(1+\log\ell),\qquad \ell>0.$$

Choose $\varepsilon>0$ with $4\alpha+d\varepsilon<d$. Conditional on $\nu(i)$, the distribution of M_i does not depend on the value of R_i, so that for some constant $c_1>0$,

$$\mathbb{E}[M_i^4|\nu(i),R_i]\le 81+11(1+\log\nu(i))\mathbf{1}_{\{\nu(i)>0\}}\le c_1(1+\nu(i)^\varepsilon),$$

The conditional distribution of $\nu(i)$, given $R_i=r$, is stochastically dominated by the Binomial $\mathrm{Bi}(m,\pi_d r^d/(t^d|B_0|))$, so that, since we assume $m\le 2\lambda t^d|B_0|$,

$$\mathbb{E}[M_i^4|R_i]\le c_1\mathbb{E}[1+(\mathrm{Bi}(m,\pi_d R_i^d/(t^d|B_0|)))^\varepsilon]$$
$$\le c_1(1+(2\lambda\pi_d R_i^d)^\varepsilon)$$

by Jensen's inequality. Hence, for suitable c_2,

$$\mathbb{E}[(1+M_i)^4(1+R_i)^{4\alpha}] \leq c_2\mathbb{E}[\max(1, R_i^{4\alpha+d\varepsilon})]. \tag{3.10}$$

For any $r > 0$, we have $P[R_i > r] = 0$ unless there exists $y \in tB_0 \cap C_i$ with $|y-x| = r$, in which case by Lemma 3.2,

$$|C_i^+ \cap B_r(x) \cap tB_0| \geq |C_{x,y} \cap B_r(x) \cap tB_0| \geq c_3 r^d,$$

for some constant $c_3 > 0$ depending only on B_0. Hence, by conditioning on the value of T_x and using (3.1), we have for some constant c_3 and large enough s that

$$\begin{aligned} P[R_i^{4\alpha+d\varepsilon} > s] &\leq \int_0^1 \left(1 - \frac{c_3 u s^{d/(4\alpha+d\varepsilon)}}{t^d|B_0|}\right)^m du \\ &\leq \int_0^1 \exp(-mc_3 u s^{d/(4\alpha+d\varepsilon)}/(t^d|B_0|))du \end{aligned}$$

and since we assume $m \geq \lambda t^d|B_0|/2$, this is bounded by a constant times $s^{-d/(4\alpha+d\varepsilon)}$. Hence, there is a constant c_4 such that

$$\mathbb{E}[\max(1, R_i^{4\alpha+d\varepsilon})] = \int_1^\infty P[R_i^{4\alpha+d\varepsilon} > s]ds < c_4.$$

Hence, using (3.8) and (3.10), we obtain (3.6). The proof of (3.7) using (3.9) is similar. ■

Proof of Theorem 3.4. We assert that there exists an almost surely finite radius of stabilization R satisfying (2.26) and (2.27). To see this, take a finite collection of cones C_i with point at $\mathbf{0}$ and angular radius $\pi/6$; let R_i^* be the distance from $\mathbf{0}$ to the nearest point in C_i to $\mathbf{0}$ with a lower mark than $T_{\mathbf{0}}$. It is not hard to see that R_i^* is almost surely finite. Then no Poisson point lying in C_i at a distance greater than R_i^* from $\mathbf{0}$ will be connected to $\mathbf{0}$ in the on-line NNG, nor will any edge involving such a point be removed when $\mathbf{0}$ is inserted. The existence of a radius of stabilization follows. Thus we can apply Lemma 2.1 to get the conditions (2.12) and (2.13) for either $h(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ or $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$.

The case $\alpha = 0$ of Lemma 3.3 gives us the condition (2.5) for the functional $h(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$ for any bounded function ψ of rooted connected graphs of radius at most 1. Also, the uniform bound (2.16) is obvious for any such h, and by scale invariance of the on-line NNG, the homogeneity condition (2.23) with $\gamma = 0$ also holds. Thus Theorems 2.2, 2.3 and 2.4 are all applicable in this case.

Turning to the case where $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$, with ϕ satisfying the growth bound (3.5), once again Lemma 3.3 gives us the condition (2.5). Also, (2.16) is again obvious in this case, so that Theorems 2.2 and 2.3 are applicable in this case. By scale invariance the homogeneity condition (2.23) holds (with $\gamma = \alpha$) for the case $\phi(r) = r^\alpha$, so that Theorem 2.4 is also applicable in this case. ■

4 Proof of the general CLT for lattice systems

Assume throughout this section that $X = (X_z, z \in \mathbb{Z}^d)$ is as described in Section 2.2. Assume also that for $i = 1, 2, \ldots, k$, $(H_t^i(A), t \geq 1, A \in \mathcal{R}(B_0))$ is a random set function as described in Section 2.2, satisfying the stabilization conditions (2.3) and (2.4) along with the moments condition (2.5) for some $\gamma > 2$. Assume also that $(t_n)_{n\geq 1}$ is an arbitrary $[1,\infty)$-valued sequence which tends to infinity as $n \to \infty$.

For $y \in \mathbb{Z}^d$, and $i \in \{1, 2, \ldots, k\}$, since the shifted family of i.i.d. variables $\tau_y X$ has the same joint distribution as X, by (2.3) there exists a random variable $\Delta^i_{\infty,y}$ such that for $A \in \mathcal{R}(B_0)$ and $z_n \in \mathbb{Z}^d, n \geq 1$ with $\liminf_{n\to\infty}(\tau_{z_n}(t_n A)) = \mathbb{R}^d$, we have

$$H^i_{t_n,z_n}(\tau_y X, A) - H^i_{t,z_n}((\tau_y X)^{\mathbf{0}}, A) \xrightarrow{P} \Delta^i_{\infty,y}. \tag{4.11}$$

In other words, $\Delta^i_{\infty,y}$ is defined in just the same manner as $\Delta^{H^i}_\infty$ at (2.3) but using the shifted family of i.i.d. variables $\tau_y X$.

For $y \in \mathbb{Z}^d$, let

$$F^i_y := \mathbb{E}[\Delta^i_{\infty,y} | \mathcal{F}_y]. \tag{4.12}$$

By the conditional Jensen inequality, Fatou's lemma, and the moments condition (2.5),

$$\mathbb{E}[(F^i_y)^2] = \mathbb{E}[(\mathbb{E}[\Delta^i_{\infty,y}|\mathcal{F}_y])^2] \leq \mathbb{E}[(\Delta^i_{\infty,y})^2] = \mathbb{E}[(\Delta^{H^i}_\infty)^2] < \infty. \tag{4.13}$$

Lemma 4.4 *Let $A \in \mathcal{R}(B_0)$, with $|A| > 0$, and let $i, j \in \{1, 2, \ldots, k\}$. Then*

$$(t_n^d|A|)^{-1} \sum_{y \in (t_n A)\cap \mathbb{Z}^d} F^i_y F^j_y \xrightarrow{L^1} \mathbb{E}[F^i_{\mathbf{0}} F^j_{\mathbf{0}}] = \mathbb{E}[\mathbb{E}(\Delta^i_{\mathbf{0}}(\infty)|\mathcal{F}_{\mathbf{0}})\mathbb{E}(\Delta^j_{\mathbf{0}}(\infty)|\mathcal{F}_{\mathbf{0}})]. \tag{4.14}$$

Proof. Since $\mathcal{F}_y$ is the σ-field generated by $(\tau_y X_z)_{z \preccurlyeq \mathbf{0}}$, the definition of $F^i_y F^j_y$ in terms of $\tau_y X$ is the same as that of $F^i_{\mathbf{0}} F^j_{\mathbf{0}}$ in terms of X. Hence the random field $(F^i_y F^j_y, y \in \mathbb{Z}^d)$ is a stationary family of random variables. Also, each variable F^i_y has has finite second moment by (4.13), and likewise for F^j_y, so that $F^i_y F^j_y$ has finite first moment by the Cauchy-Schwarz inequality.

Also, the σ-field of translation-invariant $\sigma(X)$-measurable events is trivial (see Durrett [4], chapter 6, lemma 4.3).

The result follows from the classical Ergodic Theorem ([4], chapter 6, section 2). For details, see the proof of eqn (2.8) of [11]. In the terminology of [11], the sequence of sets $(t_n A)_{n\geq 1}$ has vanishing relative boundary because of the assumption that A is Riemann measurable. This assumption also implies that $\mathrm{card}(t_n A \cap \mathbb{Z}^d) \sim t_n^d |A|$ as $n \to \infty$. ■

For $y \in \mathbb{Z}^d$, let X^y be the random field X with the value X_y at site y replaced by the independent copy X_* (i.e., $X^y = (X^y_z, z \in \mathbb{Z}^d)$ with $X^y_y = X_*$ and $X^y_z = X_z$ for $z \neq y$). For $t > 0$, and $1 \leq i \leq k$, set

$$\tilde{\Delta}^i_{t,y}(A) := H^i_t(X, A) - H^i_t(X^y, A). \tag{4.15}$$

Observe that $X^y = \tau_{-y}((\tau_y X)^{\mathbf{0}})$, so that

$$\tilde{\Delta}^i_{t,y}(A) \;=\; H^i_{t,-y}(\tau_y X, A) - H^i_{t,-y}((\tau_y X)^{\mathbf{0}}, A). \tag{4.16}$$

Therefore by the definition (2.2), since the translated random field $\tau_y X$ has the same distribution as X,

$$\tilde{\Delta}^i_{t,y}(A) \;\stackrel{\mathcal{D}}{=}\; \Delta^{H^i}_{t,-y}(A). \tag{4.17}$$

Lemma 4.5 *Let $A \in \mathcal{R}(B_0)$, and $i \in \{1, 2, \ldots, k\}$. Suppose $(y_n)_{n\geq 1}$ is a $\mathbb{Z}^d$-valued sequence. Then*

$$\lim_{n\to\infty} E[(\tilde{\Delta}^i_{t_n,y_n}(A) - \Delta^i_{\infty,y_n})^2] = 0 \quad \text{ if } \quad \liminf(\tau_{-y_n}(t_n A)) = \mathbb{R}^d \tag{4.18}$$

and

$$\lim_{n\to\infty} E[(\tilde{\Delta}^i_{t_n,y_n}(A))^2] = 0 \quad \text{ if } \quad \liminf(\tau_{-y_n}(t_n(B_0 \setminus A))) = \mathbb{R}^d. \tag{4.19}$$

Proof. The second limiting expression (4.19) follows from the distributional identity (4.17) along with the second stabilization condition (2.4) and the moments condition (2.5) (see [19] A 13.2(f)).

To prove (4.18), observe that since $\Delta^i_{\infty,y}$ is defined in terms of $\tau_y X$ in the same manner as $\Delta^{H^i}_\infty$ is defined in terms of X, we have by (4.16) that

$$\begin{aligned} & E[(\tilde{\Delta}^i_{t_n,y_n}(A) - \Delta^i_{\infty,y_n})^2] \\ &= E[(H^i_{t_n,-y_n}(\tau_{y_n} X, A) - H^i_{t_n,-y_n}((\tau_{y_n} X)^{\mathbf{0}}, A) - \Delta^i_{\infty,y_n})^2] \\ &= E[(H^i_{t_n,-y_n}(X, A) - H^i_{t_n,-y_n}(X^{\mathbf{0}}, A) - \Delta^{H^i}_\infty)^2]. \end{aligned}$$

If $\liminf(\tau_{-y_n}(t_n A)) = \mathbb{R}^d$, then this tends to zero as $n \to \infty$ by the stabilization and moments conditions (2.3) and (2.5) (again see [19] A 13.2(f)). ■

Recalling the definition of $\widetilde{A}$ at (2.1), define the sequence of sets $(B_n)_{n\geq 1}$ in $\mathbb{Z}^d$ by

$$B_n := \widetilde{t_n B_0}. \tag{4.20}$$

For $i \in \{1, 2, \ldots, k\}$, $y \in \mathbb{Z}^d$, $t \geq 1$, and $A \in \mathcal{R}(B_0)$, let

$$F^i_{t,y}(A) := \mathbb{E}[\tilde{\Delta}^i_{t,y}(A) | \mathcal{F}_y]. \tag{4.21}$$

For $\gamma > 1$, the conditional Jensen inequality implies that

$$E[|F^i_{t,y}(A)|^\gamma] = E[|E[\tilde{\Delta}^i_{t,y}(A)|\mathcal{F}_y]|^\gamma] \leq E[|\tilde{\Delta}^i_{t,y}(A)|^\gamma] \tag{4.22}$$

and therefore the distributional identity (4.17) together with the moments condition (2.5) imply that for some $\gamma > 2$,

$$\sup\{E[|F^i_{t,y}(A)|^\gamma] : A \in \mathcal{R}(B_0), t \geq 1, y \in \widetilde{tB_0}\} < \infty. \tag{4.23}$$

Lemma 4.6 *For any $A \in \mathcal{R}(B_0)$, $A' \in \mathcal{R}(B_0)$, and any $i, j \in \{1, \ldots, k\}$, as $n \to \infty$ we have*

$$t_n^{-d} \sum_{x \in (t_n(A \cap A') \cap \mathbb{Z}^d)} (F^i_{t_n,x}(A) F^j_{t_n,x}(A') - F^i_x F^j_x) \xrightarrow{L^1} 0 \tag{4.24}$$

and

$$t_n^{-d} \sum_{x \in B_n \setminus t_n(A \cap A')} F^i_{t_n,x}(A) F^j_{t_n,x}(A') \xrightarrow{L^1} 0. \tag{4.25}$$

Proof. By the triangle and Cauchy-Schwarz inequalities,

$$\begin{aligned} \mathbb{E}[|F^i_{t,x}(A) F^j_{t,x}(A') - F^i_x F^j_x|] &\leq (E[F^i_{t,x}(A)^2])^{1/2} (E[(F^j_{t,x}(A') - F^j_x)^2])^{1/2} \\ &\quad + (E[(F^i_{t,x}(A) - F^i_x)^2])^{1/2} (E[(F^j_x)^2])^{1/2}. \end{aligned} \tag{4.26}$$

By (4.23) and (4.13), $\mathbb{E}[(F^i_{t,x}(A))^2]$ and $\mathbb{E}[(F^j_x)^2]$ are uniformly bounded. Moreover, by definitions (4.12), (4.21) and by the conditional Jensen inequality,

$$\begin{aligned} \mathbb{E}[(F^i_{t,x}(A) - F^i_x)^2] &= \mathbb{E}[(\mathbb{E}[\tilde{\Delta}^i_{t,x}(A) - \Delta^i_{\infty,x}|\mathcal{F}_x])^2] \\ &\leq \mathbb{E}[\mathbb{E}[(\tilde{\Delta}^i_{t,x}(A) - \Delta^i_{\infty,x})^2|\mathcal{F}_x]] = \mathbb{E}[(\tilde{\Delta}^i_{t,x}(A) - \Delta^i_{\infty,x})^2] \end{aligned} \tag{4.27}$$

and similarly,

$$\mathbb{E}[(F^j_{t,x}(A') - F^j_x)^2] \leq \mathbb{E}[(\tilde{\Delta}^j_{t,x}(A') - \Delta^j_{\infty,x})^2]. \tag{4.28}$$

For $A \in \mathcal{R}(B_0)$, define 'interior' and 'exterior' lattice sets for the set $t_n A$ by

$$\mathrm{int}_n(A) = \{z \in \mathbb{Z}^d : B_{t_n^{1/2}}(z) \subseteq t_n A\}; \quad \mathrm{ext}_n(A) = \mathrm{int}_n(B_0 \setminus A),$$

and the 'boundary' lattice set

$$\partial_n(A) = \{z \in \mathbb{Z}^d : B_{t_n^{1/2}}(z) \cap t_n A \neq \emptyset \text{ and } B_{t_n^{1/2}}(z) \cap t_n(B_0 \setminus A) \neq \emptyset\}.$$

We assert that

$$\lim_{n \to \infty} \sup_{x \in \mathrm{int}_n(A \cap A')} \mathbb{E}[|F^i_{t_n,x}(A) F^j_{t_n,x}(A') - F^i_x F^j_x|] = 0. \tag{4.29}$$

Indeed, if this were untrue we could take a sequence $(x_n)_{n\geq 1}$ with $x_n \in \mathrm{int}_n(A\cap A')$ and

$$\limsup \mathbb{E}[|F^i_{t_n,x_n}(A)F^j_{t_n,x_n}(A') - F^i_{x_n}F^j_{x_n}|] > 0.$$

This would imply by (4.26), (4.27) and (4.28) that

$$\limsup_{n\to\infty} \max(\mathbb{E}[(\tilde{\Delta}^i_{t_n,x_n}(A) - \Delta^i_{\infty,x_n})^2], \mathbb{E}[(\tilde{\Delta}^j_{t_n,x_n}(A') - \Delta^j_{\infty,x_n})^2]) > 0,$$

which contradicts eqn (4.18) from Lemma 4.5.

By a similar argument to the proof of (4.29), this time using (4.19) and (4.22), we obtain

$$\lim_{n\to\infty} \sup_{x\in \mathrm{ext}_n(A)} \mathbb{E}[F^i_{t_n,x}(A)^2] = 0, \qquad \lim_{n\to\infty} \sup_{x\in \mathrm{ext}_n(A')} \mathbb{E}[F^j_{t_n,x}(A')^2] = 0. \tag{4.30}$$

By (4.30) and the Cauchy-Schwarz inequality,

$$\lim_{n\to\infty} \sup_{x\in \mathrm{ext}_n(A)\cup \mathrm{ext}_n(A')} \mathbb{E}[|F^i_{t_n,x}(A)F^j_{t_n,x}(A')|] = 0. \tag{4.31}$$

Using the uniform boundedness of both $\mathbb{E}[|F^i_{t_n,x}(A)F^j_{t_n,x}(A') - F^i_xF^j_x|]$ and $\mathbb{E}[|F^i_{t_n,x}(A)F^j_{t_n,x}(A')|]$ (see (4.23) and (4.13)) we may deduce (4.24) from (4.29), and (4.25) from (4.31). We here elaborate only on the argument for (4.25). The absolute value of the sum in the left hand side of (4.25) is bounded by four terms, namely a sum over $x \in \mathrm{ext}_n(A)$, a sum over $x \in \mathrm{ext}_n(A')$, a sum over x in a subset of $\partial_n(A)$, and a sum over x in a subset of $\partial_n(A')$. The first two of these terms tend to zero by (4.31), while the other terms tends to zero by the uniform boundedness of the terms in the sum and the fact that the number of sites in $\partial_n(A)$ is small relative to t_n^d (by Riemann measurability of A), and likewise for A'. ■

Proof of Theorem 2.1. We consider linear combinations. Recalling that $(t_n)_{n\geq 1}$ is an arbitrary sequence tending to infinity, let $b_1, \ldots, b_k$ be arbitrary constants. By the Cramér-Wold device (see, e.g., [4]) it suffices to prove that with

$$\sigma^*_{j\ell} := \mathbb{E}(\mathbb{E}[\Delta^{H^j}_\infty|\mathcal{F}_\mathbf{0}]\mathbb{E}[\Delta^{H^\ell}_\infty|\mathcal{F}_\mathbf{0}]), \tag{4.32}$$

we have

$$t_n^{-d/2}\sum_{j=1}^k b_j(H^j_{t_n}(A_j) - \mathbb{E}H^j_{t_n}(A_j)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \sum_{j=1}^k\sum_{\ell=1}^k b_jb_\ell|A_j\cap A_\ell|\sigma^*_{j\ell}\right), \tag{4.33}$$

and that the variance of the left hand side of (4.33) converges to that of the right hand side. We shall represent the left hand side of (4.33) as a sum of martingale differences.

For $j, \ell \in \{1, 2, \ldots, k\}$, let $A_{n,j,\ell} := t_n(A_j \cap A_\ell) \cap \mathbb{Z}^d$. Recall that $B_n := \widetilde{t_n B_0}$ (see (4.20)). Let $\nu_n = \mathrm{card}(B_n)$ and $\nu_{n,j,\ell} := \mathrm{card}(A_{n,j,\ell})$. Since $B_0, A_1, A_2, \ldots, A_k$ and are all Riemann measurable we have (for each j, ℓ)

$$\lim_{n\to\infty} (\nu_n / t_n^d) = |B_0|; \qquad \lim_{n\to\infty} (\nu_{n,j,\ell}/t_n^d) = |A_j \cap A_\ell|. \tag{4.34}$$

Define the filtration $(\mathcal{G}_0, \mathcal{G}_1, \ldots, \mathcal{G}_{\nu_n})$ as follows: let $\mathcal{G}_0$ be the trivial σ-field, label the elements of B_n in lexicographic order as $x_1, \ldots, x_{\nu_n}$, and let $\mathcal{G}_i = \mathcal{F}_{x_i}$ for $1 \le i \le \nu_n$. Then

$$\sum_{j=1}^{k} b_j (H_{t_n}^j(A_j) - \mathbb{E}\, H_{t_n}^j(A_j)) = \sum_{i=1}^{\nu_n} D_i$$

where we set $D_i := \sum_{j=1}^k b_j D_{i,j}$ with

$$D_{i,j} := \mathbb{E}[H_{t_n}^j(A_j)|\mathcal{G}_i] - \mathbb{E}[H_{t_n}^j(A_j)|\mathcal{G}_{i-1}].$$

By orthogonality of martingale differences,

$$\mathrm{Var}\left[\sum_{j=1}^{k} b_j H_{t_n}^j(A_j)\right] = \mathbb{E} \sum_{i=1}^{\nu_n} D_i^2.$$

By this representation of the variance, along with the central limit theorem for martingale difference arrays (Theorem (2.3) of McLeish [10], or Theorem 2.10 of Penrose [12]) it suffices to prove the conditions

$$\sup_{n\ge 1} \mathbb{E}\left[\max_{1\le i\le \nu_n} \left(t_n^{-d/2}|D_i|\right)^2\right] < \infty, \tag{4.35}$$

$$t_n^{-d/2} \max_{1\le i\le \nu_n} |D_i| \xrightarrow{P} 0, \tag{4.36}$$

and

$$t_n^{-d} \sum_{i=1}^{\nu_n} D_i^2 \xrightarrow{L^1} \sum_{j=1}^{k}\sum_{\ell=1}^{k} b_j b_\ell |A_j \cap A_\ell| \sigma_{j\ell}^*. \tag{4.37}$$

With $\tilde{\Delta}_{t,y}^j(A)$ defined at (4.15), and $F_{t,y}^j(A)$ defined at (4.21), we have

$$D_{i,j} = E[\tilde{\Delta}_{t_n,x_i}^j(A_j)|\mathcal{F}_{x_i}] = F_{t_n,x_i}^j(A_j). \tag{4.38}$$

First we check (4.35). By (4.38), we have

$$t_n^{-d}\mathbb{E}[\max_{i\le\nu_n} D_i^2] \le t_n^{-d} \sum_{i=1}^{\nu_n} \mathbb{E}[D_i^2] = t_n^{-d} \sum_{i=1}^{\nu_n} \mathbb{E}\left[\left(\sum_{j=1}^{k} b_j F_{t_n,x_i}^j(A_j)\right)^2\right]$$

which is uniformly bounded by (4.34) and (4.23).

For the second condition (4.36), let $\varepsilon > 0$ and use Boole's and Markov's inequalities to obtain

$$P\left[\max_{1\le i\le \nu_n} |D_i| \ge t_n^{d/2}\varepsilon\right] \le \sum_{i=1}^{\nu_n} \frac{\mathbb{E}[|D_i|^\gamma]}{t_n^{\gamma d/2}\varepsilon^\gamma},$$

which tends to zero, by (4.34) and the fact that for some $\gamma > 2$, $E[|D_i|^\gamma]$ is uniformly bounded by (4.38) and (4.23).

It remains to prove (4.37). It suffices to prove that for each $j, \ell \in \{1, 2, \dots, k\}$ we have

$$t_n^{-d} \sum_{i=1}^{\nu_n} D_{i,j} D_{i,\ell} \xrightarrow{L^1} |A_j \cap A_\ell| \sigma^*_{j\ell}. \tag{4.39}$$

Using (4.14), (4.24), (4.25), and (4.34), we obtain

$$t_n^{-d} \sum_{x\in B_n} F^j_{t_n,x}(A_j) F^\ell_{t_n,x}(A_\ell) \xrightarrow{L^1} |A_j \cap A_\ell| \mathbb{E}[F^j_{\mathbf{0}} F^\ell_{\mathbf{0}}]. \tag{4.40}$$

By the definitions (4.12) and (4.32), the right-hand side of (4.40) equals $|A_j \cap A_\ell|\sigma^*_{j\ell}$. By (4.38), eqn (4.40) gives us (4.39). The proof is complete. ■

5 Proof of general continuum results

Before proving the results stated in Section 2.3, we give some consequences of the stabilization and moments conditions given in that section. Observe first that any point process set function h satisfying the strong stabilization conditions (2.12) and (2.13) with stabilizing limit $\delta(\infty)$ must satisfy the following continuum version of the conditions (2.3), (2.4) whereby for any $A \in \mathcal{R}(B_0)$ and any $([1,\infty) \times \mathbb{R}^d)$-valued sequence $(t_n, x_n)_{n\ge1}$ we have

$$\delta(\tau_{x_n}(t_n A), \mathcal{P} \cap (\tau_{x_n}(B_0))) \xrightarrow{\text{a.s.}} \delta(\infty) \quad \text{if} \quad \liminf_{n\to\infty}(\tau_{x_n}(t_n A)) = \mathbb{R}^d \tag{5.1}$$

and

$$\delta(\tau_{x_n}(t_n A), \mathcal{P} \cap (\tau_{x_n}(B_0))) \xrightarrow{\text{a.s.}} 0 \quad \text{if} \quad \liminf_{n\to\infty}(\tau_{x_n}(t_n (B_0 \setminus A))) = \mathbb{R}^d. \tag{5.2}$$

Also, conditions (2.15) and (2.16) imply the following Poisson point process version of the condition (2.5), namely

$$\sup_{A\in\mathcal{R}(B_0), t\ge1, x\in\mathbb{R}^d : \mathbf{0}\in\tau_x(tB_0)} \{\mathbb{E}[\delta(\tau_x(tA), \mathcal{P} \cap \tau_x(tB_0))^4]\} < \infty. \tag{5.3}$$

This can be deduced from (2.15) and (2.16), by a similar argument to the proof of Lemma 4.1 of [14].

Suppose we are given a point process set function h that satisfies (2.12). For any locally finite set $\mathcal{X} \subset \mathbb{R}^d$ and any $x \in \mathbb{R}^d$, define

$$\delta_\infty(x, \mathcal{X}) := \limsup_{r \to \infty} (h((\mathcal{X} \cap B_r(\mathbf{0})) \cup \{x\}, \mathbb{R}^d) - h(\mathcal{X} \cap B_r(\mathbf{0}), \mathbb{R}^d)). \tag{5.4}$$

Observe that $\delta_\infty(\mathbf{0}, \mathcal{P})$ equals the stabilizing limit $\delta(\infty)$ of h.

Lemma 5.1 *Given $\varepsilon \in (0, 1]$, let the random vector $\xi(\varepsilon)$ be uniformly distributed over the cube $Q_{\mathbf{0}}^{\varepsilon}$, independent of $\mathcal{P}$. Then*

$$\delta_\infty(\xi(\varepsilon), \mathcal{P}) \xrightarrow{P} \delta(\infty) \quad \text{as} \quad \varepsilon \downarrow 0. \tag{5.5}$$

Proof. As $K \to \infty$, we have by (2.12) that

$$h((\mathcal{P} \cap B_K(\mathbf{0})) \cup \{\mathbf{0}\}, \mathbb{R}^d) - h(\mathcal{P} \cap B_K(\mathbf{0}), \mathbb{R}^d) \xrightarrow{P} \delta(\infty)$$

(in fact with almost sure convergence), and by translation-invariance,

$$h((\mathcal{P} \cap B_K(\mathbf{0})) \cup \{\xi(\varepsilon)\}, \mathbb{R}^d) - h(\mathcal{P} \cap B_K(\mathbf{0}), \mathbb{R}^d) \xrightarrow{P} \delta_\infty(\xi(\varepsilon), \mathcal{P}).$$

Hence, it suffices to prove that for any integer $j \geq 1$, and any $K > 0$, if $X_1, \ldots X_j$ are uniformly distributed over $B_K(\mathbf{0})$, independent of each other and of $\xi(\varepsilon)$, then

$$\begin{aligned} &h(\{\xi(\varepsilon), X_1, \ldots, X_j\}, \mathbb{R}^d) - h(\{X_1, \ldots, X_j\}, \mathbb{R}^d) \\ &\xrightarrow{P} h(\{\mathbf{0}, X_1, \ldots, X_j\}, \mathbb{R}^d) - h(\{X_1, \ldots, X_j\}, \mathbb{R}^d) \quad \text{as} \quad \varepsilon \downarrow 0. \end{aligned} \tag{5.6}$$

By (2.16), the above random variables are uniformly bounded by a constant (dependent on j and K). Define $\tilde{h} : (\mathbb{R}^d)^j \to \mathbb{R}$ by

$$\tilde{h}(y_1, \ldots, y_j) = h(\{\mathbf{0}, y_1, y_2 + y_1, y_3 + y_1, \ldots, y_j + y_1\}, \mathbb{R}^d).$$

If x_1 lies at a Lebesgue point (see e.g. [18]) of $\tilde{h}(\cdot, x_2 - x_1, x_3 - x_1, \ldots, x_j - x_1)$ then

$$\begin{aligned} &\varepsilon^{-d} \int_{Q_{\mathbf{0}}^{\varepsilon}} |h(\{x, x_1, x_2, \ldots, x_j\}, \mathbb{R}^d) - h(\{\mathbf{0}, x_1, x_2, \ldots, x_j\}, \mathbb{R}^d)| dx \\ &= \varepsilon^{-d} \int_{Q_{\mathbf{0}}^{\varepsilon}} |h(\{\mathbf{0}, x_1 - x, x_2 - x, \ldots, x_j - x\}, \mathbb{R}^d) - h(\{\mathbf{0}, x_1, x_2, \ldots, x_j\}, \mathbb{R}^d)| dx \\ &= \varepsilon^{-d} \int_{Q_{\mathbf{0}}^{\varepsilon}} |\tilde{h}(x_1 - x, x_2 - x_1, \ldots, x_j - x_1) - \tilde{h}(x_1, x_2 - x_1, \ldots, x_j - x_1)| dx \\ &\to 0 \quad \text{as} \quad \varepsilon \downarrow 0, \end{aligned} \tag{5.7}$$

where the last line comes from the definition of a Lebesgue point.

Since we assume $h(\{x_1, \ldots, x_j\})$ is a Borel-measurable function of $(x_1, \ldots, x_j)$, it follows that for all $(y_2, \ldots, y_j) \in (\mathbb{R}^d)^{j-1}$ the function $\tilde{h}(\{\cdot, y_2, \ldots, y_j\})$ is Borel-measurable, and hence, by the Lebesgue Density Theorem (see [17] or [12]), that almost every $x \in \mathbb{R}^d$ is a Lebesgue point of $\tilde{h}(\cdot, y_2, y_3, \ldots, y_j)$.

Suppose $X_1,\dots,X_j$ are independent and uniformly distributed over $B_K(\mathbf{0})$. Then for almost every possible collection of values for $(X_2-X_1,\dots,X_j-X_1)$ the conditional distribution of X_1 conditional on these values of $(X_2-X_1,\dots,X_j-X_1)$ is absolutely continuous with respect to Lebesgue measure on $\mathbb{R}^d$ (in fact, uniform over a certain region). Hence, given the values of $(X_2-X_1,\dots,X_j-X_1)$, the conditional probability that X_1 lies at a Lebesgue point of $\tilde{h}(\cdot,X_2-X_1,X_3-X_1,\dots,X_j-X_1)$ is 1. Thus, with probability 1, X_1 lies at a Lebesgue point of $\tilde{h}(\cdot,X_2-X_1,X_3-X_2,\dots,X_j-X_1)$. Hence by (5.7), and the Dominated convergence theorem,

$$\begin{aligned}&\mathbb{E}[|h(\{\xi(\varepsilon),X_1,\dots,X_j\},\mathbb{R}^d)-h(\{\mathbf{0},X_1,\dots,X_j\},\mathbb{R}^d)|]\\ =\;&\mathbb{E}\left[\varepsilon^{-d}\int_{Q_{\mathbf{0}}^{\varepsilon}}|h(\{x,X_1,\dots,X_j\},\mathbb{R}^d)-h(\{\mathbf{0},X_1,\dots,X_j\},\mathbb{R}^d)|dx\right]\\ &\qquad\to 0 \quad\text{as}\quad \varepsilon\downarrow 0.\end{aligned}$$

Since convergence in L^1 implies convergence in probability, (5.6) then follows. ■

Next we use discretization and application of Theorem 2.1 to prove a weaker statement of Theorem 2.2, which does not include the expression (2.17) for σ_{ij}. In the proof we introduce a parameter ε which we shall later on make tend to zero to establish (2.17).

Proposition 5.1 *Suppose that $h^1,\dots,h^k$ are point process set functions which satisfy the stabilization conditions* (2.12), (2.13), *the moments condition* (2.15), *and the polynomial bound* (2.16). *Then there exists a $k\times k$ matrix $\Sigma=(\sigma_{ij})_{i,j=1}^k$ such that if $A_1,\dots,A_k$ are sets in $\mathcal{R}(B_0)$, then as $t\to\infty$,*

$$t^{-d}\mathrm{Cov}(h^i(\mathcal{P}_t,tA_i),h^j(\mathcal{P}_t,tA_j))\to\lambda\sigma_{ij}|A_i\cap A_j| \tag{5.8}$$

and

$$t^{-d/2}(h^i(\mathcal{P}_t,A_i)-\mathbb{E}h^i(\mathcal{P}_t,A_i))_{i=1}^k\xrightarrow{\mathcal{D}}\mathcal{N}(\mathbf{0},(\lambda\sigma_{ij}|A_i\cap A_j|)_{i,j=1}^k). \tag{5.9}$$

Proof. Fix $\varepsilon\in(0,1]$. To apply Theorem 2.1, for $z\in\mathbb{Z}^d$ define X_z to be the point process $\tau_{-\varepsilon z}(\mathcal{P}\cap Q_z^{\varepsilon})$. Then X_z $(z\in\mathbb{Z}^d)$ are independent and identically distributed. Also, define the random set function

$$H_t^i(A):=h^i(\mathcal{P}_t,t\varepsilon A),\qquad t\geq 1, A\in\mathcal{R}(\varepsilon^{-1}B_0), \tag{5.10}$$

which is a function of $\{X_z,z\in\widetilde{(t/\varepsilon)B_0}\}$. Set $H_{t,y}^i:=H_t^i(\tau_yX,A)$, as at (2.1). By the translation invariance property (2.10) of h^i we have

$$H_{t,y}^i(A)=h^i(\mathcal{P}\cap\tau_{\varepsilon y}(tB_0),\tau_{\varepsilon y}(t\varepsilon A)). \tag{5.11}$$

We need to check conditions (2.3), (2.4), and (2.5) in this context. These refer to the increment

$$\Delta_{t,y}^{H^i}(A)=H_{t,y}^i(X,A)-H_{t,y}^i(X^{\mathbf{0}},A)$$

which, by (5.11), is (minus) the increment in $h^i(\mathcal{P} \cap \tau_{\varepsilon y}(tB_0), \tau_{\varepsilon y}(t\varepsilon A))$ when we resample the Poisson process $\mathcal{P}$ in the cube $Q_{\mathbf{0}}^{\varepsilon}$. The stabilization condition (5.1) refers instead to the insertion of a single point at the origin; however, the required stabilization (2.3) (in the present Poissonian context) can be deduced from (5.1) by the argument used to prove (3.2) of [14]. Moreover, a virtually identical argument can be used to deduce (2.4) from (5.2).

The proof of (2.5), in this context, from the assumed condition (5.3), proceeds essentially by the argument given to prove (3.3) of [14]; because of this proximity we do not give further details. Having established conditions (2.3), (2.4) and (2.5) we may apply Theorem 2.1 to deduce the results (5.8) and (5.9) (see (5.13) and (5.14) below). ■

The proof of proposition 5.1 just given actually provides us with some information about the limiting variance matrix $(\sigma_{ij})_{i,j=1}^k$. In the context of this proof, the σ-field $\mathcal{F}_{\mathbf{0}}$ appearing in Theorem 2.1 is, in effect, the σ-field generated by the restriction of the Poisson configuration $\mathcal{P}$ to $\cup_{z \in \mathbb{Z}^d, z \preccurlyeq \mathbf{0}} Q_z^{\varepsilon}$, i.e., to cubes in the division of $\mathbb{R}^d$ into cubes Q_z^{ε} of side ε, up to and including $Q_{\mathbf{0}}^{\varepsilon}$ in the lexicographic ordering. To emphasize its dependence on ε, we denote this σ-field by $\mathcal{F}_{\mathbf{0}}^{\varepsilon}$. With the random set function H_t^i defined by (5.10), set

$$\sigma_{ij}^*(\varepsilon) := \mathbb{E}\,[\mathbb{E}\,(\Delta_\infty^{H^i} | \mathcal{F}_{\mathbf{0}}^{\varepsilon}) \mathbb{E}\,(\Delta_\infty^{H^j} | \mathcal{F}_{\mathbf{0}}^{\varepsilon})]. \tag{5.12}$$

Then the application of Theorem 2.1 in the preceding proof gives us

$$\begin{aligned} \lim_{t \to \infty} t^{-d} \mathrm{Cov}(h^i(\mathcal{P}_t, tA_i), h^j(\mathcal{P}_t, tA_j)) &= \lim_{t \to \infty} t^{-d} \mathrm{Cov}(H_t^i(\varepsilon^{-1} A_i), H_t^j(\varepsilon^{-1} A_j))) \\ &= \varepsilon^{-d} |A_i \cap A_j| \sigma_{ij}^*(\varepsilon) \end{aligned} \tag{5.13}$$

and as $t \to \infty$,

$$(t^{-d/2}(h_t^i(A_i) - \mathbb{E}\, h_t^i(A_i)))_{i=1}^k \xrightarrow{\mathcal{D}} \mathcal{N}(\mathbf{0}, (\varepsilon^{-d} \sigma_{ij}^*(\varepsilon) |A_i \cap A_j|)_{i,j=1}^k). \tag{5.14}$$

In other words, the matrix $(\sigma_{i,j})_{i,j=1}^k$ in the statement of Proposition 5.1 is given, for any $\varepsilon \in (0, 1]$, by

$$\sigma_{ij} = \lambda^{-1} \varepsilon^{-d} \sigma_{ij}^*(\varepsilon). \tag{5.15}$$

Proof of Theorem 2.2. In view of Proposition 5.1 and the discussion above, it remains to prove that σ_{ij}, given by (5.15) for any $\varepsilon \in (0, 1]$, is also given by (2.17).

With i and j fixed, define point process set functions $h := h^i + h^j$ and $h' := h^i - h^j$, along with the corresponding random set functions $H := H^i + H^j$ and $H' := H^i - H^j$ (where H^i and H^j are given at (5.10)). The definition of H^i also depends on ε, as does the limiting increment $\Delta_\infty^{H^i}$; from now on we denote the latter quantity by $\Delta_\infty^{H^i, \varepsilon}$, and define $\Delta_\infty^{H, \varepsilon}$ and $\Delta_\infty^{H', \varepsilon}$ analogously. By linearity, for all $\varepsilon > 0$ we have

$$\sigma_{ij}^*(\varepsilon) = (1/4) \mathbb{E} \left[(\mathbb{E}\,[\Delta_\infty^{H, \varepsilon} | \mathcal{F}_{\mathbf{0}}^{\varepsilon}])^2 - (\mathbb{E}\,[\Delta_\infty^{H', \varepsilon} | \mathcal{F}_{\mathbf{0}}^{\varepsilon}])^2 \right].$$

To prove (2.17), we use the fact that the value of $\varepsilon^{-d}\sigma^*_{ij}(\varepsilon)$ does not depend on the choice of ε, since the left hand side of (5.13) does not depend on ε and therefore neither does the right hand side. The aim is to show, by taking $\varepsilon \downarrow 0$ in (5.15) that σ_{ij} equals the expression

$$\mathbb{E}\,[\mathbb{E}\,(\delta^i(\infty)|\mathcal{F})\mathbb{E}\,(\delta^j(\infty)|\mathcal{F})] = \frac{1}{4}\,\mathbb{E}\,\left[(\mathbb{E}\,[\delta(\infty)|\mathcal{F}])^2 - (\mathbb{E}\,[\delta'(\infty)|\mathcal{F}])^2\right],$$

where $\delta(\infty)$ (respectively $\delta'(\infty)$) is the stabilizing limit of the point process set function $h^i + h^j$ (respectively $h^i - h^j$). In other words, it remains to prove that

$$\lim_{\varepsilon\downarrow 0}\varepsilon^{-d}\mathbb{E}\,[(\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}|\mathcal{F}_{\mathbf{0}}^\varepsilon])^2] = \lambda\mathbb{E}\,[(\mathbb{E}\,[\delta(\infty)|\mathcal{F}])^2], \tag{5.16}$$

and also a similar limit for H', for which the proof will be identical.

By following the proof of ([14], Lemma 3.1) and observing that $c(\mu)$ in that proof tends to zero as $\mu \downarrow 0$, we see that

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\,[(\Delta_\infty^{H,\varepsilon})^4] = 0. \tag{5.17}$$

Let N_ε (respectively N'_ε) be the number of points of $\mathcal{P}$ in the cube $Q_{\mathbf{0}}^\varepsilon$ (respectively the number of resampled Poisson points in $Q_{\mathbf{0}}^\varepsilon$). If $N_\varepsilon = N'_\varepsilon = 0$ then $\Delta_\infty^{H,\varepsilon} = 0$. Also, N_ε is $\mathcal{F}_{\mathbf{0}}^\varepsilon$-measurable. Hence,

$$\begin{aligned}\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}\mathbf{1}_{\{N'_\varepsilon=0\}}|\mathcal{F}_{\mathbf{0}}^\varepsilon]\mathbf{1}_{\{N_\varepsilon=0\}} &= \mathbb{E}\,[\Delta_\infty^{H,\varepsilon}\mathbf{1}_{\{N'_\varepsilon=0\}}\mathbf{1}_{\{N_\varepsilon=0\}}|\mathcal{F}_{\mathbf{0}}^\varepsilon]\\ &= E[0|\mathcal{F}_{\mathbf{0}}^\varepsilon] = 0, \quad \text{a.s.}\end{aligned}$$

Hence

$$\begin{aligned}(\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}|\mathcal{F}_{\mathbf{0}}^\varepsilon]\mathbf{1}_{\{N_\varepsilon=0\}})^2 &= (\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}(\mathbf{1}_{\{N'_\varepsilon=0\}} + \mathbf{1}_{\{N'_\varepsilon>0\}})|\mathcal{F}_{\mathbf{0}}^\varepsilon]\mathbf{1}_{\{N_\varepsilon=0\}})^2\\ &= (\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}\mathbf{1}_{\{N'_\varepsilon>0\}}|\mathcal{F}_{\mathbf{0}}^\varepsilon]\mathbf{1}_{\{N_\varepsilon=0\}})^2 \le (\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}\mathbf{1}_{\{N'_\varepsilon>0\}}|\mathcal{F}_{\mathbf{0}}^\varepsilon])^2, \quad \text{a.s..}\end{aligned}$$

Hence, by the conditional Cauchy-Schwarz inequality (see e.g. [4]), and the independence of N'_ε and $\mathcal{F}_{\mathbf{0}}^\varepsilon$,

$$(\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}|\mathcal{F}_{\mathbf{0}}^\varepsilon]\mathbf{1}_{\{N_\varepsilon=0\}})^2 \le \mathbb{E}\,[(\Delta_\infty^{H,\varepsilon})^2|\mathcal{F}_{\mathbf{0}}^\varepsilon]P[N'_\varepsilon > 0], \quad \text{a.s.}$$

Taking expectations, then using Jensen's inequality and (5.17), we obtain

$$\begin{aligned}\varepsilon^{-d}\mathbb{E}\,[(\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}|\mathcal{F}_{\mathbf{0}}^\varepsilon])^2\mathbf{1}_{\{N_\varepsilon=0\}}] &\le \varepsilon^{-d}P[N'_\varepsilon > 0]\mathbb{E}\,[(\Delta_\infty^{H,\varepsilon})^2]\\ &\le \lambda\mathbb{E}\,[(\Delta_\infty^{H,\varepsilon})^4]^{1/2}\\ &\to 0 \quad \text{as } \varepsilon \downarrow 0.\end{aligned} \tag{5.18}$$

Let $Y_\varepsilon = (\mathbb{E}\,(\Delta_\infty^{H,\varepsilon}|\mathcal{F}_{\mathbf{0}}^\varepsilon))^2$. By the Cauchy-Schwarz and Jensen inequalities, and (5.17),

$$\begin{aligned}\varepsilon^{-d}\mathbb{E}\,[Y_\varepsilon\mathbf{1}_{\{N_\varepsilon\ge 2\}}] &\le \varepsilon^{-d}(P[N_\varepsilon \ge 2])^{1/2}(\mathbb{E}\,[Y_\varepsilon^2])^{1/2}\\ &\le \text{const.} \times \mathbb{E}\,[(\Delta_\infty^{H,\varepsilon})^4]^{1/2} \to 0 \quad \text{as } \ \varepsilon \downarrow 0.\end{aligned} \tag{5.19}$$

Similarly,

$$\varepsilon^{-d}\mathbb{E}\,[Y_\varepsilon \mathbf{1}_{\{N_\varepsilon=1,N'_\varepsilon\ge 1\}}] \to 0 \quad \text{as} \quad \varepsilon \downarrow 0.$$

It remains to consider $\mathbb{E}\,[Y_\varepsilon \mathbf{1}_{\{N_\varepsilon=1,N'_\varepsilon=0\}}]$. Since $P[N_\varepsilon = 1, N'_\varepsilon = 0] \sim \lambda\varepsilon^d$ as $\varepsilon \downarrow 0$, to establish (5.16) we must show that

$$\lim_{\varepsilon\downarrow 0}\mathbb{E}\,[(\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}|\mathcal{F}_0^\varepsilon])^2|N_\varepsilon = 1, N'_\varepsilon = 0] = \mathbb{E}\,[(\mathbb{E}\,[\delta(\infty)|\mathcal{F}])^2]. \tag{5.20}$$

Given ε, let $\mathbf{Y}_\varepsilon$ (respectively $\mathbf{Y}'_\varepsilon$) be the restriction of the Poisson process $\mathcal{P}$ to the union of cubes Q_z^ε with $z \in \mathbb{Z}^d$ and $z \prec \mathbf{0}$ (respectively, with $\mathbf{0} \prec z$).

Given that $N_\varepsilon = 1$ and $N'_\varepsilon = 0$, the restriction of $\mathcal{P}$ to $Q_{\mathbf{0}}^\varepsilon$ consists of a single point uniformly distributed over $Q_{\mathbf{0}}^\varepsilon$; we denote this random point by ξ. Then, given that $N_\varepsilon = 1$ and $N'_\varepsilon = 0$, almost surely $\Delta_\infty^{H,\varepsilon}$ equals the increment $\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon)$ (using notation defined at (5.4)). Thus,

$$\mathbb{E}\,[(\mathbb{E}\,[\Delta_\infty^{H,\varepsilon}|\mathcal{F}_0^\varepsilon])^2|N_\varepsilon = 1, N'_\varepsilon = 0] = \mathbb{E}\,[(\mathbb{E}\,[\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon)|\xi, \mathbf{Y}_\varepsilon])^2] \tag{5.21}$$

where ξ is uniformly distributed over $Q_{\mathbf{0}}^\varepsilon$ and is independent of $\mathcal{P}$.

By the Cauchy-Schwarz and Jensen inequalities,

$$\begin{aligned}
&\mathbb{E}\,\{(\mathbb{E}\,[\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon)|\xi, \mathbf{Y}_\varepsilon])^2 - (\mathbb{E}\,[\delta(\infty)|\xi, \mathbf{Y}_\varepsilon])^2\} \\
&= \mathbb{E}\,\{\mathbb{E}\,[\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) + \delta(\infty)|\xi, \mathbf{Y}_\varepsilon]\mathbb{E}\,[\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) - \delta(\infty)|\xi, \mathbf{Y}_\varepsilon]\} \\
&\le \mathbb{E}\,[(\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) + \delta(\infty))^2]^{1/2}\mathbb{E}\,[(\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) - \delta(\infty))^2]^{1/2}.
\end{aligned} \tag{5.22}$$

By the moments condition (5.3), the stabilization condition (5.1), and Fatou's lemma, $\mathbb{E}\,[\delta(\infty)^4] < \infty$. Also, by definition $\delta(\infty)$ is almost surely the same as $\delta_\infty(\mathbf{0}, \mathcal{P})$ which has the same distribution as $\delta(\xi, \mathcal{P})$ by translation-invariance, so that

$$\begin{aligned}
\mathbb{E}\,[\delta(\infty)^4] &= \mathbb{E}\,[\delta_\infty(\xi, \mathcal{P})^4] \\
\ge P[\mathcal{P} \cap Q_{\mathbf{0}}^\varepsilon = \emptyset]&\mathbb{E}\,[\delta_\infty(\xi, \mathcal{P})^4|\mathcal{P} \cap Q_{\mathbf{0}}^\varepsilon = \emptyset] \\
&= e^{-\lambda\varepsilon^d}\mathbb{E}\,[\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon)^4],
\end{aligned}$$

so that $\mathbb{E}\,[\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon)^4]$ remains uniformly bounded as $\varepsilon \downarrow 0$. Combining all these estimates, we obtain

$$\limsup_{\varepsilon\downarrow 0}\mathbb{E}\,[(\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) + \delta(\infty))^2] < \infty. \tag{5.23}$$

As $\varepsilon \downarrow 0$, it is the case that $P[\mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon \neq \mathcal{P}]$ tends to zero, and hence

$$\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) - \delta_\infty(\xi, \mathcal{P}) \xrightarrow{P} 0.$$

Combined with (5.5) from Lemma 5.1, this implies that $\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) \xrightarrow{P} \delta(\infty)$, and hence, using also the fact that $\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon)$ has uniformly bounded fourth moments, we obtain the limit

$$\mathbb{E}\,[(\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) - \delta(\infty))^2] \to 0 \quad \text{as} \quad \varepsilon \downarrow 0.$$

Combined with (5.22) and (5.23), this shows that

$$\lim_{\varepsilon \downarrow 0} \left(\mathbb{E} \{ (\mathbb{E} [\delta_\infty(\xi, \mathbf{Y}_\varepsilon \cup \mathbf{Y}'_\varepsilon) | \xi, \mathbf{Y}_\varepsilon])^2 - (\mathbb{E}[\delta(\infty)|\xi, \mathbf{Y}_\varepsilon])^2 \} \right) = 0 \tag{5.24}$$

Also, by the independence of ξ from $\delta(\infty)$ and $\mathbf{Y}_\varepsilon$, along with the martingale convergence theorem, as $\varepsilon \downarrow 0$

$$\mathbb{E}[\delta(\infty)|\xi, \mathbf{Y}_\varepsilon] = \mathbb{E}[\delta(\infty)|\mathbf{Y}_\varepsilon] \to \mathbb{E}[\delta(\infty)|\mathcal{F}], \qquad \text{a.s.}$$

Since $\mathbb{E}[\delta(\infty)^4] < \infty$, the variables $(\mathbb{E}[\delta(\infty)|\xi, \mathbf{Y}_\varepsilon])^2$ are uniformly integrable, so that

$$\mathbb{E}[(\mathbb{E}[\delta(\infty)|\xi, \mathbf{Y}_\varepsilon])^2] \to \mathbb{E}[(\mathbb{E}[\delta(\infty)|\mathcal{F}])^2] \text{ as } \varepsilon \downarrow 0.$$

Combining this with (5.24) and (5.21), we obtain (5.20) as required. ■

To de-Poissonize the limits (2.18), (2.19) and obtain (2.21) and (2.22), we use a technique related to that used in [7] and [8]. Given point process set functions h, h', we make the definition

$$R_{m,t}(A) = h(\mathcal{U}_{m+1,t}, tA) - h(\mathcal{U}_{m,t}, tA), \tag{5.25}$$
$$R'_{m,t}(A) = h'(\mathcal{U}_{m+1,t}, tA) - h'(\mathcal{U}_{m,t}, tA). \tag{5.26}$$

Recall from (2.14) the definition

$$\mu_t := \lambda t^d |B_0|.$$

We shall use the following coupling lemma, which resembles Lemma 4.2 of [14].

Lemma 5.2 *Suppose h and h' are strongly stabilizing point process set functions. Suppose $A \in \mathcal{R}(B_0)$. Let the random d-vector Y be uniformly distributed over B_0 and independent of $\mathcal{P}$. Let $\varepsilon > 0$. Then there exists $\eta > 0$ and $t_0 \geq 1$ such that for all $t \geq t_0$ and all integer $m, m' \in [(1-\eta)\mu_t, (1+\eta)\mu_t]$ with $m < m'$, there exists a coupled family of variables D, D', R, R' with following properties:*

- *D and D' each have the same distribution as $\delta(\infty)\mathbf{1}_{\{Y \in A\}}$;*
- *D and D' are independent;*
- *(R, R') have the same joint distribution as $(R_{m,t}(A), R'_{m',t}(A))$;*
- *$P[\{D \neq R\} \cup \{D' \neq R'\}] < \varepsilon$.*

Proof. Suppose we are given t. On a suitable probability space, let $\mathcal{P}$ and $\mathcal{P}'$ be independent homogeneous Poisson point processes in $\mathbb{R}^d$ of intensity λ; let $U, U', V_1, V_2, \ldots$ be independent variables uniformly distributed over tB_0, independent of $\mathcal{P}$ and $\mathcal{P}'$.

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

Let N denote the number of points of $\mathcal{P}''$ lying in tB_0 (a Poisson variable with mean μ_t). Choose an ordering on the points of $\mathcal{P}''$ lying in tB_0, uniformly at random from all $N!$ possible such orderings. Use this ordering to list the points of $\mathcal{P}''$ in tB_0 as $W_1, W_2, \ldots, W_N$. Also, set $W_{N+1} = V_1, W_{N+2} = V_2, W_{N+3} = V_3$ and so on. Define the point process $\mathcal{W}_n := \{W_1, \ldots, W_n\}$ (for each $n \geq 1$), and the increments

$$R := h(\mathcal{W}_m \cup \{U\}, tA) - h(\mathcal{W}_m, tA);$$

$$R' = h'(\mathcal{W}_{m'-1} \cup \{U, U'\}, tA) - h'(\mathcal{W}_{m'-1} \cup \{U\}, tA).$$

The variables $U, U', W_1, W_2, W_3, \ldots,$ are independent uniformly distributed variables on tB_0, and therefore the pairs (R, R') and $(R_{m,t}(A), R'_{m',t}(A))$ have the same joint distribution as claimed.

Let $\tilde{\mathcal{P}}$ be the translated point process $\tau_{-U}(\mathcal{P})$. Similarly, let $\tilde{\mathcal{P}}' := \tau_{-U'}(\mathcal{P}')$. Then $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$ are independent homogeneous Poisson processes of intensity λ on $\mathbb{R}^d$. Moreover, U and U' are independent of $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$. Let S, S' be radii of stabilization for $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$, respectively, and define

$$D = \delta(\mathbb{R}^d, \tilde{\mathcal{P}} \cap B_S(\mathbf{0}))\mathbf{1}_{\{U \in tA\}}; \qquad D' = \delta(\mathbb{R}^d, \tilde{\mathcal{P}}' \cap B_{S'}(\mathbf{0}))\mathbf{1}_{\{U' \in tA\}}.$$

Then D and D' are independent, and each have the same distribution as $\delta(\infty)\mathbf{1}_{\{Y \in A\}}$.

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 assumption that B_0 and A are Riemann measurable, take t 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 Euclidean distance at least $2K$ from $\partial(tB_0)$, from $\partial(tA)$, and from each other. Set $\eta = \varepsilon(2K)^{-d}/(18\lambda)$. We assume $|m - \mu_t| \leq \eta\mu_t$ and $|m' - \mu_t| \leq \eta\mu_t$. Define events E_1, E, E' by

$$E_1 := \{|N - m| > 2\eta\mu_t\} \cup \{|N - m'| > 2\eta\mu_t\},$$

$$E := \{\mathcal{W}_m \cap B_K(U) \neq \mathcal{P}'' \cap tB_0 \cap B_K(U)\},$$
$$E' := \{\mathcal{W}_{m'} \cap B_K(U') \neq \mathcal{P}'' \cap tB_0 \cap B_K(U')\}.$$

Event E occurs either if one or more of the $(N - m)^+$ "discarded" points of $\mathcal{P}''$ lies in $B_K(U)$, or if one or more of the $(m - N)^+$ "added" points of $\{V_1, V_2, \ldots\}$ lies in $B_K(U)$, and similarly for E'. Hence,

$$P[E|E_1^c] \leq (2\eta\mu_t)(2K)^d(\lambda/\mu_t) < \varepsilon/9; \quad P[E'|E_1^c] < \varepsilon/9.$$

Using the defining properties (2.12) and (2.13) of the radii of (strong) stabilization S, S' for $\tilde{\mathcal{P}}$ and $\tilde{\mathcal{P}}'$, and using Boole's inequality, we obtain for large enough t that

$$P[(D, D') \neq (R, R')] \leq P[E_0] + P[E_1] + P[S > K] + P[S' > K] \\ + P[E \setminus E_1] + P[E' \setminus E_1] < \varepsilon. \qquad \blacksquare$$

Lemma 5.3 *Suppose h is strongly stabilizing and satisfies the moments condition (2.15). Suppose $A \in \mathcal{R}(B_0)$. Let $g : [1,\infty) \to (0,\infty)$ be a function with $g(t)/t^d \to 0$ as $t \to \infty$. Then*

$$\lim_{t\to\infty} \sup_{\mu_t - g(t) \le m \le \mu_t + g(t)} \left| \mathbb{E}\, R_{m,t}(A) - \left(\frac{|A|}{|B_0|} \right) \mathbb{E}\, \delta(\infty) \right| = 0. \tag{5.27}$$

Also,

$$\lim_{n\to\infty} \sup_{\mu_t - g(t) \le m < m' \le \mu_t + g(t)} \left| \mathbb{E}\, R_{m,t}(A) R_{m',t}(A) - \left(\frac{|A|}{|B_0|} \mathbb{E}\, \delta(\infty) \right)^2 \right| = 0, \tag{5.28}$$

and

$$\lim_{n\to\infty} \sup_{\mu_t - g(t) \le m \le \mu_t + g(t)} |\mathbb{E}\, R_{m,t}(A)^2| < \infty. \tag{5.29}$$

Proof. Let m be an arbitrary integer satisfying $\mu_t - g(t) \le m \le \mu_t + g(t)$. Let $\varepsilon > 0$. Provided t is large enough, by Lemma 5.2 we can find coupled variables D and R, with D having the same distribution as $\delta(\infty)\mathbf{1}_{\{Y \in A\}}$, with R having the same distribution as $R_{m,t}(A)$, and with $P[D \ne R] < \varepsilon$. Then

$$\mathbb{E}\, R_{m,t}(A) = \mathbb{E}\, R = \mathbb{E}[D] - \mathbb{E}[D\mathbf{1}_{\{D \ne R\}}] + \mathbb{E}[R\mathbf{1}_{\{D \ne R\}}].$$

By the Cauchy-Schwarz inequality and the moments conditions (5.3) and (2.15), there is a constant c, independent of ε, such that $|\mathbb{E}[D\mathbf{1}_{\{D\ne R\}}]| \le c\varepsilon^{1/2}$ and $|\mathbb{E}[R\mathbf{1}_{\{D \ne R\}}]| \le c\varepsilon^{1/2}$. Since ε is arbitrarily small, (5.27) follows. Moreover, (5.29) follows from (2.15).

Next we consider m, m' with $\mu_t - g(t) \le m < m' \le \mu_t + g(t)$. By Lemma 5.2 we can find coupled variables D, D', R, R' such that D and D' are independent and each have the same distribution as $\delta(\infty)\mathbf{1}_{\{Y \in A\}}$, (R, R') have the same joint distribution as $(R_{m,t}(A), R'_{m',t}(A))$, and $P[(D, D') \ne (R, R')] < \varepsilon$. Then

$$\mathbb{E}[RR'] - \mathbb{E}[DD'] = \mathbb{E}[RR'\mathbf{1}\{(D,D') \ne (R,R')\}] - \mathbb{E}[DD'\mathbf{1}\{(D,D') \ne (R,R')\}],$$

and by the Cauchy-Schwarz inequality and the moments conditions (5.3) and (2.15), the right hand side has modulus bounded by a constant multiple of $\varepsilon^{1/2}$. Since ε is arbitrarily small, (5.28) follows. ■

Proof of Theorem 2.3. Assume $(t_n)_{n\ge1}$ is a $(1,\infty)$-valued sequence satisfying $n/(t_n^d |B_0|) \to \lambda$ as $n \to \infty$. For $1 \le j \le k$, let

$$\zeta_n^j := h^j(\mathcal{U}_{n,t_n}, t_n A_j); \qquad \tilde{\zeta}_n^j := h^j(\mathcal{P}_{t_n}, t_n A_j).$$

For each n, assume $\mathcal{P}_{t_n}$ is coupled to $\mathcal{U}_{n,t_n}$ by setting $\mathcal{P}_{t_n} = \{U_{1,t_n}, U_{2,t_n}, \dots, U_{N_n,t_n}\}$ with N_n an independent Poisson variable with mean $\mu_n = \lambda t_n^d |B_0|$. Define the k-vector

$$\alpha := (\alpha_j)_{j=1}^k, \quad \text{with } \alpha_j := \left(\frac{|A_j|}{|B_0|} \right) \mathbb{E}[\delta^j(\infty)].$$

The first step is to prove that as $n \to \infty$,

$$\mathbb{E}\left[(n^{-1/2}(\tilde{\zeta}_n^j - \zeta_n^j - (N_n - n)\alpha_j))^2\right] \to 0. \tag{5.30}$$

To prove this, (writing $t(n)$ for t_n when typographically convenient), note that the expectation in the left hand side is equal to

$$\sum_{m:|m-\mu_{t(n)}|\leq n^{3/4}} \mathbb{E}\left[n^{-1}\left(h^j(\mathcal{U}_{m,t(n)}, t_n A_j) - h^j(\mathcal{U}_{n,t(n)}, t_n A_j) - (m-n)\alpha_j\right)^2\right] P[N_n = m]$$
$$+ n^{-1}\mathbb{E}\left[\left(\tilde{\zeta}_n^j - \zeta_n^j - (N_n - n)\alpha_j\right)^2 \mathbf{1}\{|N_n - \mu_{t(n)}| > n^{3/4}\}\right]. \tag{5.31}$$

Let $\varepsilon > 0$. By (5.25) and Lemma 5.3, there exists $c > 0$ such that for large enough n and all m with $n \leq m \leq \mu_{t(n)} + n^{3/4}$,

$$\mathbb{E}[(h^j(\mathcal{U}_{m,t(n)}, t_n A_j) - h^j(\mathcal{U}_{n,t(n)}, t_n A_j) - (m-n)\alpha_j)^2] = \mathbb{E}\left[\left(\sum_{\ell=n}^{m-1}(R^j_{\ell,t(n)}(A_j) - \alpha_j)\right)^2\right]$$
$$\leq \varepsilon(m-n)^2 + c(m-n),$$

where the bound comes from expanding out the double sum arising from the expectation of the squared sum. A similar argument applies when $\mu_{t(n)} - n^{3/4} \leq m \leq n$, and hence the first term in (5.31) 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 uniform bound (2.16) and the Cauchy-Schwarz inequality, there is a constant β_3 such that the second term in (5.31) is bounded by $\beta_3 n^{\beta_3}(P[|N_n - \mu_{t(n)}| > n^{3/4}])^{1/2}$, which tends to zero, e.g. by Lemma 1.4 of [12]. This completes the proof of (5.30).

Let $b_1, \ldots, b_k$ be arbitrary real constants. Define the column vector $\mathbf{b} := (b_1, \ldots, b_k)'$. Let

$$\zeta_n := \sum_{j=1}^k b_j \zeta_n^j; \qquad \zeta_n' = \sum_{j=1}^k b_j \tilde{\zeta}_n^j.$$

We prove convergence of $n^{-1}\mathrm{Var}(\zeta_n)$, using the identity

$$n^{-1/2}\zeta_n' = n^{-1/2}\zeta_n + n^{-1/2}(N_n - n)\alpha'\mathbf{b} + n^{-1/2}(\zeta_n' - \zeta_n - (N_n - n)\alpha'\mathbf{b}).$$

In the right hand side, the third term has variance tending to zero by (5.30), while the second term has variance $(\alpha'\mathbf{b})^2$ and is independent of the first term. It follows that with the matrix $\Sigma = (\sigma_{ij})_{i,j=1}^k$ given by Theorem 2.2, and the matrix $\Sigma^A = (\sigma_{ij}^A)_{i,j=1}^k$ given by

$$\sigma_{ij}^A := \frac{\sigma_{ij}|A_i \cap A_j|}{|B_0|},$$

we have from Theorem 2.2 that

$$\mathbf{b}'\Sigma^A\mathbf{b} = \lim_{n\to\infty} n^{-1}\mathrm{Var}(\zeta_n') = \lim_{n\to\infty}(n^{-1}\mathrm{Var}(\zeta_n)) + (\alpha'\mathbf{b})^2,$$

so that $\Sigma^A - \alpha\alpha'$ is nonnegative definite and $n^{-1}\mathrm{Var}(\zeta_n) \to \mathbf{b}'(\Sigma^A - \alpha\alpha')\mathbf{b}$. This gives us (2.21).

The proof of theorem 2.2 (since it is derived by taking linear combinations) tells us that $n^{-1/2}(\zeta_n' - \mathbb{E}\,\zeta_n') \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{b}'\Sigma^A\mathbf{b})$. Combined with (5.30) this gives us

$$n^{-1/2}(\zeta_n - \mathbb{E}\,\zeta_n' + (N_n - n)\alpha'\mathbf{b}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{b}'\Sigma^A\mathbf{b}). \tag{5.32}$$

Recall that $\mu_n := \lambda t_n^d|B_0| = \mathbb{E}\,N_n$. Since $n^{-1/2}(N_n - \mu_n)\alpha'\mathbf{b}$ is independent of ζ_n and is asymptotically normal with mean zero and variance $(\alpha'\mathbf{b})^2 = \mathbf{b}'\alpha\alpha'\mathbf{b}$, it follows from (5.32), by considering characteristic functions, that

$$n^{-1/2}(\zeta_n - \mathbb{E}\,\zeta_n' + (\mu_n - n)\alpha'\mathbf{b}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{b}'(\Sigma^A - \alpha\alpha')\mathbf{b}). \tag{5.33}$$

By (5.30), the expectation of $n^{-1/2}(\zeta_n' - \zeta_n - (N_n - n)\alpha'\mathbf{b})$ tends to zero, so in (5.33) we can replace $-\mathbb{E}\,\zeta_n' + (\mu_n - n)\alpha'\mathbf{b}$ by $-\mathbb{E}\,\zeta_n$, which gives us

$$n^{-1/2}(\zeta_n - \mathbb{E}\,\zeta_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathbf{b}'(\Sigma^A - \alpha\alpha')\mathbf{b}).$$

Then (2.22) follows by the Cramér-Wold device. ■

Proof of Lemma 2.1. First set $h(\mathcal{X}, A) = L_\phi^G(\mathcal{X}, A)$, as defined at (2.28). Let the random variable R satisfy (2.26) and (2.27). Let $A \in \mathcal{R}(\mathbb{R}^d)$, and let $\mathcal{A} \subset \mathbb{R}^d \setminus B_R(\mathbf{0})$ be finite. Then if $B_R(\mathbf{0}) \subseteq A$, the increment $\delta(A, (\mathcal{P} \cap B_R(\mathbf{0})) \cup \mathcal{A})$ is equal to

$$\left(\sum_{e\in\mathcal{E}^+(\mathbf{0};\mathcal{P}^0)} \phi(|e|)\right) - \sum_{e\in\mathcal{E}^-(\mathbf{0};\mathcal{P}^0)} \phi(|e|), \tag{5.34}$$

since all added and removed edges have both endpoints in A. Hence (2.12) holds with $\delta(\infty)$ equal to the expression displayed in (5.34). If instead $A \cap B_R(\mathbf{0}) = \emptyset$, then $\delta(A, (\mathcal{P} \cap B_R(\mathbf{0})) \cup \mathcal{A}) = 0$ since added and removed edges have neither endpoint in A. Hence, (2.13) holds.

Next, suppose we set $h(\mathcal{X}, A) = V_\psi^G(\mathcal{X}, A)$, where ψ is a bounded functional on the set $\mathcal{K}_\kappa$ of rooted graphs of radius at most κ, with $\kappa \in \mathbb{N}$.

We assert that there exists an almost surely finite random variable R such that (2.26) and (2.27) hold, and such that for every vertex X of $G(\mathcal{P}^0 \cap B_R(\mathbf{0}))$ at a graph distance at most κ from $\mathbf{0}$, the set of edges incident to X is unaffected by changes outside $B_R(\mathbf{0})$, and for every vertex Y of $G(\mathcal{P} \cap B_R(\mathbf{0}))$ at a graph distance at most κ from some endpoint of some edge in $\mathcal{E}^-(\mathbf{0}; \mathcal{P}^0 \cap B_R(\mathbf{0}))$ the set of edges incident to Y is unaffected by changes outside $B_R(\mathbf{0})$. The existence of such an R follows from the proof of Lemma 3.3 of [15].

Then for $A \in \mathcal{R}(\mathbb{R}^d)$ with $B_R(\mathbf{0}) \subseteq A$, only for vertices X of $G(\mathcal{P}^0 \cap B_R(\mathbf{0}))$ at a graph distance at most κ from the origin, and vertices X of of $G(\mathcal{P} \cap B_R(\mathbf{0}))$ at a graph distance at most κ from some endpoint of some edge in $\mathcal{E}^-(\mathbf{0}; \mathcal{P}^0)$, have the possibility for $\psi(G_X(\mathcal{P}^0))$ to be different from $\psi(G_X(\mathcal{P}))$, and moreover for these vertices it is the case that the value of $\psi(G_X(\mathcal{P}^0)) - \psi(G_X(\mathcal{P}))$ is unaffected by changes outside $B_R(\mathbf{0})$. The conditions (2.12) and (2.13) follow for this case.

Suppose now that uniqueness of the infinite component holds, and set $h(\mathcal{X}, A) = K^G(\mathcal{X}, A)$. The stabilization conditions are proved, essentially by a slight modification of the proof of Proposition 6.1 of [14]. For the convenience of the reader, we describe the argument in the present, more general context.

Let R be a radius of stabilization, as given at (2.26) and (2.27). Choose a finite $R' > R$ such that for any two points of $\mathcal{P}$ in $B_R(\mathbf{0})$, either they can be connected by a path in $G(\mathcal{P})$ all of whose nodes lie in $B_{R'}(\mathbf{0})$, or at least one of them lies in a finite component contained in $B_{R'}(\mathbf{0})$, and such that a similar statement holds for $\mathcal{P}^0$. The proof that we can choose such an R' is based on the uniqueness of the infinite component in $G(\mathcal{P})$ and $G(\mathcal{P}^0)$, and is given in more detail in [14].

By Lemma 3.3 of [15], there almost surely exists $R'' > R'$ such that for all $X \in \mathcal{P} \cap B_{R'}(\mathbf{0})$, the set of edges incident to X in $G(\mathcal{P}^0)$ is unaffected by additions or deletions of points outside $B_{R''}$, and moreover, by the proof of Lemma 3.3 of [15], we can choose R'' to be so large that in addition, the set of edges incident to X in $G(\mathcal{P})$ is unaffected by additions or deletions of points outside $B_{R''}$.

Suppose that $A \in \mathcal{R}(\mathbb{R}^d)$ and $B_{R''}(\mathbf{0}) \subseteq A$. Suppose $\mathcal{A}$ is disjoint from $B_{R''}(\mathbf{0})$. When we change from $G(\mathcal{P} \cap B_{R''}(\mathbf{0}) \cup \mathcal{A})$ to $G(\mathcal{P}^0 \cap B_{R''}(\mathbf{0}) \cup \mathcal{A})$, the effect is first to add a vertex at the origin, then to add the edges of $\mathcal{E}^+(\mathbf{0}; \mathcal{P}^0)$, and then to remove the edges of $\mathcal{E}^-(\mathbf{0}; \mathcal{P}^0)$. Consider adding successive edges, in some specified order. Each edge reduces the number of components that intersect A by 1 if it joins two points that were previously not connected by a path, and otherwise does not affect the number of components. The question of whether a particular added edge changes the number of components is determined by the graph structure of the restriction of $G(\mathcal{P})$ to vertices in $B_{R'}(\mathbf{0})$, and therefore does not depend on A or $\mathcal{A}$ (always presuming $A \in \mathcal{R}$ and $B_{R''}(\mathbf{0}) \subseteq A$). A similar argument applies with deleted edges.

It follows from the above that if we set $\delta(\infty) := \delta(B_{R''}(\mathbf{0}), \mathcal{P} \cap B_{R''}(\mathbf{0}))$ and $S = R''$, then (2.12) holds.

Now suppose that $A \cap B_{R''}(\mathbf{0}) = \emptyset$ (and $\mathcal{A}$ is also disjoint from $B_{R''}(\mathbf{0})$ as before). Consider again the process of successive additions and deletions described above. If an added edge connects two previously disconnected components, then at least one of them has a vertex set entirely contained in $B_{R'}(\mathbf{0})$, and therefore does not have any vertices in A, and so this change does not cause any increment in the number of components that have at least one vertex in A. A similar argument applies with removed vertices; hence, if $B_{R''}(\mathbf{0}) \cap A = \emptyset$ we have $\delta(A, (\mathcal{P} \cap B_{R''}(\mathbf{0})) \cup \mathcal{A}) = 0$, so that (2.13) holds. ■

References

[1] D. Aldous and J.M. Steele (1992) Asymptotics for Euclidean minimal spanning trees on random points. *Probab. Theory Related Fields* **92**, 247-258.

[2] Baryshnikov, Y. and Yukich, J. E. (2002) Central limit theorems for random measures in geometric probability. Preprint, electronically available from `http://www.lehigh.edu/ math`

[3] Berger, N., Bollobás, B., Borgs, C., Chayes, J., and Riordan, O. (2003). Degree Distribution of the FKP model. *Automata, Languages and Programming: 30th International Colloquium, ICALP 2003, Lecture Notes in Computer Science 2719*, eds. J. C. M. Baeten, J. K. Lenstra, J. Parrow, and G. J. Woeginger, 725–738. Springer, Heidelberg.

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

[5] Fabrikant, A., Koutsoupias, E. and Papadimitriou, C. M. Heuristically optimized trade-offs: a new paradigm for power laws in the Internet. ICALP 2002, *Lecture Notes in Computer Science* **2380**, Springer, pp. 110–122.

[6] Grimmett, G. (1999) *Percolation (Second Edition).* Springer, Berlin

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

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

[9] Lee, S. (1999) The central limit theorem for Euclidean minimal spanning trees II. *Adv. Appl. Probab.* **31**, 969-984.

[10] McLeish, D. L. (1974) Dependent central limit theorems and invariance principles. *Ann. Probab.* **2**, 620-628.

[11] Penrose, M. D. (2001) A central limit theorem with applications to percolation, epidemics and boolean models. *Ann. Probab.* **29**, 1515-1546.

[12] Penrose, M. (2003) *Random Geometric Graphs.* Oxford University Press, Oxford.

[13] Penrose, M.D. and Pisztora, A. (1996) Large deviations for discrete and continuous percolation. *Adv. Appl. Probab.* **28**, 29-52.

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

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

[16] Penrose, M.D. and Yukich, J. E. (2004) Normal approximation in geometric probability. In preparation.

[17] Rudin, W. (1976) *Principles of Mathematical Analysis (Third Edition).* McGraw-Hill International.

[18] Rudin, W. (1987) *Real and Complex Analysis (Third Edition).* McGraw-Hill, New York.

[19] Williams, D. (1991) *Probability with Martingales.* Cambridge University Press, Cambridge.

[20] Yukich, J. E. (1998) *Probability Theory of Classical Euclidean Optimization Problems.* Lecture Notes in Mathematics **1675**, Springer, Berlin.

[21] Yukich, J.E. (1999) Asymptotics for weighted minimal spanning trees on random points. *Stochastic Process. Appl.* **85** 123-138.

[22] Zhang, Y. (2001) A martingale approach in the study of percolation clusters in the Z^d lattice. *J. Theoret. Probab.* **14**, 165-187.

Department of Mathematical Sciences
University of Bath
Bath BA2 7AY
United Kingdom
M.D.Penrose@bath.ac.uk