

THE ASYMPTOTIC SIZE OF THE LARGEST COMPONENT IN RANDOM GEOMETRIC GRAPHS WITH SOME APPLICATIONS

GE CHEN,*

CHANGLONG YAO** AND

TIANDE GUO,*** Chinese Academy of Sciences

Abstract

In this paper we estimate the expectation of the size of the largest component in a supercritical random geometric graph; the expectation tends to a polynomial on a rate of exponential decay. We sharpen the expectation's asymptotic result using the central limit theorem. Similar results can be obtained for the size of the biggest open cluster, and for the number of open clusters of percolation on a box, and so on.

Keywords: Random geometric graph; percolation; largest component; Poisson–Boolean model; number of open clusters

2010 Mathematics Subject Classification: Primary 60K35

Secondary 60D05; 82B43

1. Introduction

The size of the largest component is a basic property for random geometric graphs (RGGs) and has attracted much interest during the past years, including both theoretical studies (see [7]–[10]) and various applications (see [1], [3], [11], and [12]). In this paper we investigate the asymptotic size of the largest component of a RGG in the supercritical case.

Given a set $\mathcal{X} \subset \mathbb{R}^d$, let $G(\mathcal{X}; r)$ denote the undirected graph with vertex set \mathcal{X} and with undirected edges which connect all those pairs $\{X, Y\}$ with $\|Y - X\| \leq r$, where $\|\cdot\|$ denotes the Euclidean norm (l_2 -norm). The basic model of RGGs can be formulated as $G(\mathcal{X}_n; r_n)$, where \mathcal{X}_n denotes n points which are independently and uniformly distributed in a d -dimensional unit cube. To overcome the lack of spatial independence for the binomial point process \mathcal{X}_n , the model of continuum percolation must be introduced. Following [9, Section 1.7], let \mathcal{H}_λ be a homogeneous Poisson process of intensity λ on \mathbb{R}^d . For $s > 0$, define $B(s) := [0, s]^d$ and $\mathcal{H}_{\lambda,s} := \mathcal{H}_\lambda \cap B(s)$. Following [9], we write the Poisson–Boolean model as $G(\mathcal{H}_{\lambda,s}; 1)$.

We now introduce some notation related to percolation. Following [9, Section 9.6], let $\mathcal{H}_{\lambda,0}$ denote the point process $\mathcal{H}_\lambda \cup \{\mathbf{0}\}$, where $\mathbf{0}$ is the origin in \mathbb{R}^d , and, for $k \in \mathbb{N}$, let $p_k(\lambda)$ denote the probability that the order of the component in $G(\mathcal{H}_{\lambda,0}; 1)$ containing the origin is equal to k . The *percolation probability* $p_\infty(\lambda)$ is defined to be the probability that $\mathbf{0}$ lies in an infinite

Received 3 November 2009; revision received 7 July 2013.

* Postal address: National Center for Mathematics and Interdisciplinary Sciences and Key Laboratory of Systems and Control, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, P. R. China. Email address: chenge@amss.ac.cn

** Postal address: Institute of Applied Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, 100190, P. R. China. Email address: deducemath@126.com

*** Postal address: School of Mathematical Science, University of Chinese Academy of Sciences, Beijing, 100049, P. R. China. Email address: tdguo@ucas.ac.cn

component of the graph $G(\mathcal{H}_{\lambda,0}; 1)$. Therefore, we have $p_\infty(\lambda) = 1 - \sum_{k=1}^\infty p_k(\lambda)$. Let

$$\lambda_c = \inf\{\lambda > 0: p_\infty(\lambda) > 0\}$$

denote the critical intensity of continuum percolation. It is well known that $0 < \lambda_c < \infty$ for $d \geq 2$ (see [2], [5], and [6]).

Following [9, Section 9.6], let $L_j(G)$ denote the order of the j th-largest component of any graph G . Then $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ denotes the order of the largest component of $G(\mathcal{H}_{\lambda,s}; 1)$. The asymptotic properties of $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ have been studied by Penrose [9]. The basic asymptotic result about $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ was provided by Penrose [9, Theorem 10.9], i.e. if $\lambda \neq \lambda_c$ then

$$s^{-d}L_1(G(\mathcal{H}_{\lambda,s}; 1)) \xrightarrow{\mathbb{P}} \lambda p_\infty(\lambda) \quad \text{as } s \rightarrow \infty. \tag{1.1}$$

Also, Penrose [9, Theorem 10.22] gave a central limit theorem for $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ in the supercritical case $\lambda > \lambda_c$, i.e

$$s^{-d/2}(L_1(G(\mathcal{H}_{\lambda,s}; 1)) - \mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))]) \xrightarrow{\mathbb{D}} \mathcal{N}(0, \sigma^2),$$

where $\mathcal{N}(u, \sigma^2)$ denotes the normal distribution with mean u and variance σ^2 . However, the question of how large $\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))]$ should be still remains unanswered. By (1.1) it can be deduced that $\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] = \lambda p_\infty(\lambda)s^d + o(s^d)$, where $f(s) = o(g(s))$ indicates that $\lim_{s \rightarrow \infty} (f(s)/g(s)) = 0$. This result is not precise enough for some theoretic analysis and practical applications.

The corresponding asymptotic results and central limit theorem for $G(\mathcal{X}_n; r_n)$ have also been established by Penrose [9, Theorems 11.9 and 11.16], but we may ask similar questions. In this paper we will study the problem and give a more precise description for the asymptotic sizes of $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ and $L_1(G(\mathcal{X}_n; r_n))$. Our method can be adapted to study some other models and problems.

2. Main results

Our main results can be formulated as the following two theorems.

Theorem 2.1. *Suppose that $d \geq 2$ and $\lambda > \lambda_c$. Then there exist constants $c = c(d, \lambda) > 0$ and $\tau_i = \tau_i(d, \lambda)$, $1 \leq i \leq d$, with $\tau_1 > 0$, such that, for all large enough s ,*

$$\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] = \lambda p_\infty(\lambda)s^d - \sum_{i=1}^d \tau_i s^{d-i} + o(e^{-cs}). \tag{2.1}$$

Also, there exists a constant $\sigma = \sigma(d, \lambda) > 0$ such that

$$L_1(G(\mathcal{H}_{\lambda,s}; 1))s^{-d/2} - \lambda p_\infty(\lambda)s^{d/2} + \sum_{i=1}^{\lfloor d/2 \rfloor} \tau_i s^{d/2-i} \xrightarrow{\mathbb{D}} \mathcal{N}(0, \sigma^2) \tag{2.2}$$

as $s \rightarrow \infty$. Here $\lfloor \cdot \rfloor$ denotes the integer-part function.

Theorem 2.2. *Suppose that $d \geq 2$ and $\lambda > \lambda_c$. Let σ and τ_i be the same constants as defined in Theorem 2.1. There exists a constant $\delta = \delta(d, \lambda)$, with $0 < \delta \leq \sigma$, such that*

$$L_1\left(G\left(\mathcal{X}_n; \left(\frac{n}{\lambda}\right)^{-1/d}\right)\right)\left(\frac{n}{\lambda}\right)^{-1/2} - p_\infty(\lambda)(\lambda n)^{1/2} + \sum_{i=1}^{\lfloor d/2 \rfloor} \tau_i \left(\frac{n}{\lambda}\right)^{1/2-i/d} \xrightarrow{\mathbb{D}} \mathcal{N}(0, \delta^2)$$

as $n \rightarrow \infty$.

To prove these two theorems, we estimate the value of $\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))]$, and then, using the central limit theorem for $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ and $L_1(G(\mathcal{X}_n; (n/\lambda)^{-1/d}))$, we can prove (2.2) and Theorem 2.2.

We need some more notation before we prove our results. For any $x \in \mathbb{R}^d$, we write its l_∞ -norm with $\|x\|_\infty$ given by the maximum absolute value of its coordinates. For any finite set $A \subset \mathbb{R}^d$, the diameter of A is $\text{diam}(A) = \sup_{x,y \in A} \|x - y\|_\infty$. Also, let $|A|$ denote the cardinality of A . Let \oplus denote the Minkowski addition of sets. Let $\text{Leb}(\cdot)$ denote the Lebesgue measure. To simplify the expression, we will omit the dependence of all constants on d and λ ; for example, the constant c stands for $c(d, \lambda)$.

Given $\lambda > \lambda_c$, by the uniqueness of the infinite component in continuum percolation (see [9, Theorem 9.19]), the infinite graph $G(\mathcal{H}_\lambda; 1)$ has precisely one infinite component \mathcal{C}_∞ with probability 1. Let C_1, C_2, \dots, C_M denote the components of $G(\mathcal{C}_\infty \cap B(s); 1)$, taken in a decreasing order. We give a result on the rate of subexponential decay of the difference between $\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))]$ and $\mathbb{E}[|C_1|]$.

Lemma 2.1. *Suppose that $d \geq 2$ and $\lambda > \lambda_c$. There exists a constant $c > 0$ such that, for large enough s ,*

$$0 \leq \mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] - \mathbb{E}[|C_1|] \leq e^{-cs}. \tag{2.3}$$

Proof. By the definition of $L_1(G(\mathcal{H}_{\lambda,s}; 1))$ and C_1 , clearly $\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] \geq \mathbb{E}[|C_1|]$. Thus, it just remains to prove the second inequality of (2.3).

Given any $x \in \mathbb{R}^d$, let $C_\infty(x)$ denote the infinite connected component of $G(\mathcal{H}_\lambda \cup \{x\}; 1)$. By the Palm theorem for Poisson processes (see [9, Theorem 1.6]), we have

$$\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] = \lambda \int_{B(s)} \mathbb{P}[x \in V_1(x)] dx,$$

where $V_1(x)$ denotes the largest component of $G(\mathcal{H}_{\lambda,s} \cup \{x\}; 1)$, and

$$\mathbb{E}[|C_1|] = \lambda \int_{B(s)} \mathbb{P}[x \in C_1(x)] dx,$$

where $C_1(x)$ denotes the largest component of $C_\infty(x) \cap B(s)$. Therefore,

$$\begin{aligned} \mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] - \mathbb{E}[|C_1|] &= \lambda \int_{B(s)} (\mathbb{P}[x \in V_1(x)] - \mathbb{P}[x \in C_1(x)]) dx \\ &\leq \lambda \int_{B(s)} \mathbb{P}[\{x \in V_1(x)\} \cap \{x \notin C_1(x)\}] dx \\ &= \lambda \int_{B(s)} \mathbb{P}[\{x \in V_1(x)\} \cap \{x \notin C_\infty(x)\}] dx. \end{aligned} \tag{2.4}$$

Suppose that $0 < \varepsilon < \frac{1}{2}$. By [9, Theorem 10.19], there exist constants $c_1 > 0$ and $s_1 > 0$, such that if $s > s_1$ then

$$\begin{aligned} \mathbb{P}[|V_1(x)| < (1 - \varepsilon)\lambda s^d p_\infty(\lambda)] &\leq \mathbb{P}[L_1(G(\mathcal{H}_{\lambda,s}; 1)) < (1 - \varepsilon)\lambda s^d p_\infty(\lambda)] \\ &\leq \exp(-c_1 s^{d-1}). \end{aligned} \tag{2.5}$$

Also, by [9, Theorem 10.15], there exists a constant $c_2 > 0$ such that, for large enough s ,

$$\sum_{k \geq \lceil (1-\varepsilon)\lambda s^d p_\infty(\lambda) \rceil} p_k(\lambda) < \exp(-c_2 [(1 - \varepsilon)\lambda s^d p_\infty(\lambda)]^{(d-1)/d}), \tag{2.6}$$

where $\lceil x \rceil$ denotes the smallest integer not less than x . Therefore, from (2.5) and (2.6) we obtain

$$\begin{aligned} & \mathbb{P}[\{x \in V_1(x)\} \cap \{x \notin C_\infty(x)\}] \\ & \leq \mathbb{P}[|V_1(x)| < (1 - \varepsilon)\lambda s^d p_\infty(\lambda)] \\ & \quad + \mathbb{P}[\{x \in V_1(x)\} \cap \{x \notin C_\infty(x)\} \cap \{|V_1(x)| \geq (1 - \varepsilon)\lambda s^d p_\infty(\lambda)\}] \\ & \leq \exp(-c_1 s^{d-1}) + \sum_{k \geq \lceil (1-\varepsilon)\lambda s^d p_\infty(\lambda) \rceil} p_k(\lambda) \\ & < \exp(-c_1 s^{d-1}) + \exp(-c_2 [(1 - \varepsilon)\lambda p_\infty(\lambda)]^{(d-1)/d} s^{d-1}) \quad \text{as } s \rightarrow \infty. \end{aligned}$$

Combined with (2.4) this yields our result.

To estimate the value of $\mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))]$, by Lemma 2.1 we just need to get the value of $\mathbb{E}[|C_1|]$ instead. Actually, by the Palm theory for infinite Poisson processes (see [9, Theorem 9.22]), we have

$$\mathbb{E}\left[\sum_{i=1}^M |C_i|\right] = \mathbb{E}[|\mathcal{C}_\infty \cap B(s)|] = \lambda p_\infty(\lambda) s^d, \tag{2.7}$$

so we just need to estimate the value of $\mathbb{E}[\sum_{i=2}^M |C_i|]$. Let $L(s) := B(s) \setminus [1, s - 1]^d$. For any $2 \leq i \leq M$, since $C_i \subset \mathcal{C}_\infty$, there exists at least one point in $L(s) \cap C_i$ which connects to $\mathcal{C}_\infty \setminus B(s)$ directly; we choose the nearest one to the boundary of $B(s)$ as the *out-connect point*. We can see that each component of C_2, \dots, C_M contains exactly one out-connect point.

For any region $R \subseteq B(s)$ and $2 \leq i \leq M$, define

$$\chi_i(R) := \begin{cases} 1 & \text{if the out-connect point of } C_i \text{ is contained by } R, \\ 0 & \text{otherwise,} \end{cases}$$

and define

$$\xi(R) = \xi(R, s) := \sum_{i=2}^M \chi_i(R) |C_i|.$$

By the definition of $\xi(\cdot)$, it is easy to see that, for any $R, \tilde{R} \subset B(s)$, if $\text{Leb}(R \cap \tilde{R}) = 0$ then $\mathbb{E}[\xi(R \cap \tilde{R})] = 0$ and $\mathbb{E}[\xi(R \cup \tilde{R})] = \mathbb{E}[\xi(R)] + \mathbb{E}[\xi(\tilde{R})]$.

For $0 \leq i \leq d - 1$, define

$$R_i = R_i(s) := [0, 1] \times \underbrace{\left[0, \frac{s}{2}\right] \times \cdots \times \left[0, \frac{s}{2}\right]}_{d-1-i} \times \underbrace{\left[1, \frac{s}{2}\right] \times \cdots \times \left[1, \frac{s}{2}\right]}_i.$$

Note that $[1, s/2]^d \cap L(s) = \emptyset$. Then, by symmetry,

$$\begin{aligned} \mathbb{E}\left[\sum_{i=2}^M |C_i|\right] &= \mathbb{E}[\xi(B(s))] \\ &= 2^d \mathbb{E}\left[\xi\left(\left[0, \frac{s}{2}\right]^d\right)\right] \end{aligned}$$

$$\begin{aligned}
 &= 2^d \left\{ \mathbb{E}[\xi(R_0)] + \mathbb{E} \left[\xi \left(\left[1, \frac{s}{2} \right] \times \left[0, \frac{s}{2} \right]^{d-1} \right) \right] \right\} \\
 &= 2^d \sum_{i=0}^{d-1} \mathbb{E}[\xi(R_i)]. \tag{2.8}
 \end{aligned}$$

Thus, we just need to estimate the value of $\mathbb{E}[\xi(R_i)]$. Lemmas 2.2–2.5, below, are now stated to get the desired estimation.

Lemma 2.2. *Suppose that $d \geq 2$ and $\lambda > \lambda_c$. Let $V_x = V_x(s)$ denote the connected component containing x of $G(\mathcal{H}_{\lambda,s} \cup \{x\}; 1)$. There exist constants $c > 0$ and $n_0 > 0$ such that, if $n > n_0$ and $s > 2n$ then, for any point $x \in B(s)$,*

$$\mathbb{P} \left[n \leq \text{diam}(V_x) \leq \frac{s}{2} \right] < e^{-cn} \tag{2.9}$$

and

$$\mathbb{P} \left[\{|V_x| \geq n\} \cap \left\{ \text{diam}(V_x) \leq \frac{s}{2} \right\} \right] < \exp(-cn^{(d-1)/d}). \tag{2.10}$$

Proof. The proof uses ideas from the latter part of the proof of Theorem 10.18 in [9]. Given $x \in \mathbb{R}^d$, let \tilde{z} denote the point in $B'_{\tilde{z}}(n(s))$ satisfying $x \in B_{\tilde{z}}$, where the definitions of $B'_{\tilde{z}}(n(s))$ and $B_{\tilde{z}}$ are given in [9, p. 216] and [9, p. 217], respectively. Also, $C_x, D_{\text{ext}}C_x, M_0, n(s)$, and $M(s)$ are defined as in [9, pp. 218–219]. Penrose [9, p. 219] proved that $D_{\text{ext}}C_x$ is $*$ -connected and if $|C_x| < n(s)^d/2$ then

$$|D_{\text{ext}}C_x| \geq (2d)^{-1} \left(1 - \left(\frac{2}{3} \right)^{1/d} \right) |C_x|^{(d-1)/d}. \tag{2.11}$$

Let $\mathcal{A}_{m,s}$ denote the collection of $*$ -connected subsets of cardinality m which disconnects the point \tilde{z} from the giant component of $B'_{\tilde{z}}(n(s))$. Then $\mathcal{A}_{m,s}$ is restricted by the box of $B'_{\tilde{z}}(n(s)) \cap ([-m, m]^d \oplus \tilde{z})$ and $D_{\text{ext}}C_x \in \mathcal{A}_{|D_{\text{ext}}C_x|,s}$. By a Peierls argument (see [9, Corollary 9.4]), the cardinality $|\mathcal{A}_{m,s}|$ is bounded by $(2m+1)^d \gamma^m$, with $\gamma := 2^{3^d}$. Therefore, there exists a constant k_0 such that, for any integer $k > k_0$,

$$\begin{aligned}
 \mathbb{P}[|D_{\text{ext}}C_x| \geq k] &\leq \mathbb{P} \left[\bigcup_{m \geq k} \bigcup_{\sigma \in \mathcal{A}_{m,s}} \{X_z = 0 \text{ for all } z \in \sigma\} \right] \\
 &\leq \sum_{m \geq k} (2m+1)^d \gamma^m (1-p_1)^m \\
 &< \left(\frac{2}{3} \right)^k. \tag{2.12}
 \end{aligned}$$

By the definition of C_x and $D_{\text{ext}}C_x$, if $n \leq \text{diam}(V_x) \leq s/2$ then

$$\frac{n}{M(s)} - 1 \leq \text{diam}(C_x) \leq \frac{n(s)}{2} + 2;$$

therefore, we obtain $|C_x| < n(s)^d/2$ and $|D_{\text{ext}}C_x| \geq n/M(s) - 1$ for large s . Therefore, by (2.12), there exists a constant $n_0 > 0$ such that if $n > n_0$ then

$$\mathbb{P} \left[n \leq \text{diam}(V_x) \leq \frac{s}{2} \right] \leq \mathbb{P} \left[|D_{\text{ext}}C_x| \geq \frac{n}{M(s)} - 1 \right] < \left(\frac{2}{3} \right)^{n/2M_0-1}.$$

This yields (2.9).

It remains to consider the case of $|V_x| > n$. Since C_x is a $*$ -connected component containing \tilde{z} in $B'_Z(n(s))$, by a Peierls argument (see [9, Lemma 9.3]), for all k , the number of $*$ -connected subsets of $B'_Z(n(s))$ of cardinality k containing \tilde{z} is at most γ^k . Let $c_2 \geq e^2(2M_0)^d\lambda$. If $|C_x| < k$ and $|V_x| \geq c_2k + 1$, then for at least one of these subsets of $B'_Z(n(s))$ the union of the associated boxes B_z contains at least c_2k points of \mathcal{H}_λ . Therefore, by [9, Lemma 1.2], we have

$$\begin{aligned} \mathbb{P}[\{|C_x| < k\} \cap \{|V_x| \geq c_2k + 1\}] &< \gamma^k \mathbb{P}[\text{Po}(k(2M_0)^d\lambda) \geq c_2k] \\ &\leq \gamma^k \exp\left\{-\left(\frac{c_2k}{2}\right) \log\left(\frac{c_2}{(2M_0)^d\lambda}\right)\right\}. \end{aligned} \tag{2.13}$$

So if c_2 is chosen to be large enough, this probability decays exponentially in k .

Set $\beta := (2d)^{-1}(1 - (\frac{2}{3})^{1/d})$. By (2.11) and (2.12), we have

$$\mathbb{P}\left[\left\{\text{diam}(V_x) \leq \frac{s}{2}\right\} \cap \{|C_x| \geq k\}\right] \leq \mathbb{P}[|D_{\text{ext}}C_x| \geq \beta k^{(d-1)/d}] < \left(\frac{2}{3}\right)^{\beta k^{(d-1)/d}}.$$

Combined with (2.13), this gives (2.10).

For $x \in B(s)$ and $0 < a \leq 1$, define the box

$$B_i(x, a) := x \oplus \underbrace{([0, 1] \times \cdots \times [0, 1])}_i \times \underbrace{[0, a] \times \cdots \times [0, a])}_{d-i}.$$

Also, for any region $R \subseteq B(s)$, define

$$D(R) = D(R, s) := \max_{2 \leq j \leq M, \chi_j(R)=1} \text{diam}(C_j).$$

Lemma 2.3. *Suppose that $d \geq 2$ and $\lambda > \lambda_c$. There exist constants $c > 0$ and $n_0 > 0$, such that if $x \in B(s)$, $a \in (0, 1]$, and $n > n_0$, then*

$$\mathbb{P}[D(B_i(x, a)) \geq n] < e^{-cn} \tag{2.14}$$

and

$$\mathbb{P}[\xi(B_i(x, a)) \geq n] < \exp(-cn^{(d-1)/d}) + e^{-cs}. \tag{2.15}$$

Proof. Let W_1 denote the number of the connected components which intersect with $B_i(x, a)$ and have metric diameter not greater than $s/2$ but not smaller than n . By Markov’s inequality, we obtain

$$\mathbb{P}\left[\{D(B_i(x, a)) \geq n\} \cap \left\{D(B_i(x, a)) \leq \frac{s}{2}\right\}\right] \leq \mathbb{P}[W_1 > 0] \leq \mathbb{E}[W_1]. \tag{2.16}$$

By the Palm theory for Poisson processes and Lemma 2.2, if $n > n_0$ then

$$\begin{aligned} \mathbb{E}[W_1] &= \lambda \int_{B_i(x,a)} \mathbb{P}\left[\{\text{diam}(V_x(s)) \geq n\} \cap \left\{\text{diam}(V_x(s)) \leq \frac{s}{2}\right\}\right] dx \\ &< \lambda a^{d-i} e^{-cn}. \end{aligned} \tag{2.17}$$

Also, since C_i ($2 \leq i \leq M$) is not the largest component of $G(\mathcal{H}_{\lambda,s}; 1)$, by [9, Proposition 10.13], there exist constants $c_1 > 0$ and $s_1 > 0$ such that if $s > s_1$ then

$$\mathbb{P}\left[D(B_i(x, a)) > \frac{s}{2}\right] < e^{-c_1s}. \tag{2.18}$$

Together with (2.16), (2.17), and (2.18), we obtain

$$\mathbb{P}[D(B_i(x, a)) \geq n] < e^{-cn} + e^{-c_1s}.$$

Since $\mathbb{P}[D(B_i(x, a)) > s] = 0$, (2.14) follows.

Note that $B_i(x, a)$ contains at most 2^d connected components. Thus, if $\xi(B_i(x, a)) \geq n$, by the definition of $\xi(\cdot)$ there exists at least one component intersecting with $B_i(x, a)$ such that it contains no less than $2^{-d}n$ points. Let W_2 be the number of the connected components which intersect with $B_i(x, a)$ and have more than $2^{-d}n$ elements and not larger than $s/2$ metric diameter. With a similar argument as used to obtain (2.16) and (2.17), it follows that if $n > n_0$ then

$$\begin{aligned} &\mathbb{P}\left[\{\xi(B_i(x, a)) \geq n\} \cap \left\{D(B_i(x, a)) \leq \frac{s}{2}\right\}\right] \\ &\leq \mathbb{E}[W_2] \\ &= \lambda \int_{B_i(x, a)} \mathbb{P}\left[\{|V_x(s)| \geq 2^{-d}n\} \cap \left\{\text{diam}(V_x(s)) \leq \frac{s}{2}\right\}\right] dx \\ &< \lambda a^{d-i} \exp(-c2^{-d}n); \end{aligned}$$

together with (2.18), this gives (2.15).

Let real numbers $s_1 > 2$ and $s_2 > 2$ be given. Let points $x = (x_1, x_2, \dots, x_d) \in [0, s_1/2]^d$ and $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_d) \in [0, s_2/2]^d$ be given. For all $1 \leq j \leq d$, define

$$N_{x, \tilde{x}}^j(s_1, s_2) := \begin{cases} \min(s_1, s_2) - x_j - 1 & \text{if } x_j = \tilde{x}_j, \\ \min(x_j, \tilde{x}_j, s_1 - x_j - 1, s_2 - \tilde{x}_j - 1) & \text{otherwise,} \end{cases}$$

and let

$$N_{x, \tilde{x}}(s_1, s_2) := \min_{1 \leq j \leq d} [N_{x, \tilde{x}}^j(s_1, s_2)]. \tag{2.19}$$

Lemma 2.4. *Let us assume that $d \geq 2$, $\lambda > \lambda_c$, $1 \leq i \leq d$, and $0 < a \leq 1$. There exist constants $c > 0$ and $n_0 > 0$, such that if $x \in [0, s_1/2]^d$, $\tilde{x} \in [0, s_2/2]^d$, and $N_{x, \tilde{x}}(s_1, s_2) > n_0$ then*

$$|\mathbb{E}[\xi(B_i(x, a), s_1)] - \mathbb{E}[\xi(B_i(\tilde{x}, a), s_2)]| < \exp(-cN_{x, \tilde{x}}(s_1, s_2)).$$

Proof. Let $B'(s_2) := B(s_2) \oplus \{x - \tilde{x}\}$, and let $\tilde{C}_1, \tilde{C}_2, \dots, \tilde{C}_{\tilde{M}}$ denote the components of $G(\mathcal{C}_\infty \cap B'(s_2); 1)$, taken in decreasing order. For any region $R \subseteq B'(s_2)$ and $2 \leq i \leq \tilde{M}$, define

$$\tilde{\chi}_i(R) := \begin{cases} 1 & \text{if the out-connect point of } \tilde{C}_i \text{ is contained by } R, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\tilde{\xi}(R, s_2) := \sum_{i=2}^{\tilde{M}} \tilde{\chi}_i(R) |\tilde{C}_i|$, and define

$$\tilde{D}(R, s_2) := \max_{2 \leq j \leq \tilde{M}, \tilde{\chi}_j(R)=1} \text{diam}(\tilde{C}_j).$$

According to the ergodicity of Poisson point processes, we obtain

$$\mathbb{P}[\tilde{\xi}(B_i(x, a), s_2) = k] = \mathbb{P}[\xi(B_i(\tilde{x}, a), s_2) = k] \quad \text{for all } k \geq 1. \tag{2.20}$$

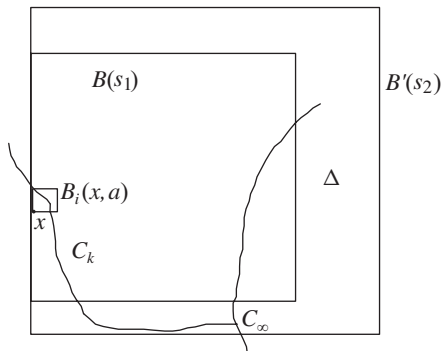


FIGURE 1: If C_k connects with $\mathcal{H}_\lambda \cap \Delta$, the event $\xi(B_i(x, a), s_1) \neq \tilde{\xi}(B_i(x, a), s_2)$ may happen.

Let $\Delta := B(s_1) \cup B'(s_2) - B(s_1) \cap B'(s_2)$. If $\xi(B_i(x, a), s_1) \neq \tilde{\xi}(B_i(x, a), s_2)$, then there exists at least one component among $C_2, \dots, C_M, \tilde{C}_2, \dots, \tilde{C}_M$ which connects directly with $\mathcal{H}_\lambda \cap \Delta$; see Figure 1. For simplicity of exposition, we take $N = N_{x, \tilde{x}}(s_1, s_2)$, $\xi_1 = \xi(B_i(x, a), s_1)$, and $\xi_2 = \tilde{\xi}(B_i(x, a), s_2)$. Therefore, by (2.14), if $N > n_0 + 1$ then

$$\begin{aligned} \mathbb{P}[\xi_1 \neq \xi_2] &\leq \mathbb{P}[\{D(B_i(x, a), s_1) \geq N - 1\} \cup \{\tilde{D}(B_i(x, a), s_2) \geq N - 1\}] \\ &< 2e^{-c(N-1)}. \end{aligned} \tag{2.21}$$

Also,

$$\begin{aligned} &\mathbb{P}[\{\xi_1 = k\} \cap \{\xi_2 \neq k\}] + \mathbb{P}[\{\xi_1 \neq k\} \cap \{\xi_2 = k\}] \\ &= \mathbb{P}[\{\xi_1 = k\}] + \mathbb{P}[\{\xi_2 = k\}] - 2\mathbb{P}[\{\xi_1 = k\} \cap \{\xi_2 = k\}] \\ &\geq |\mathbb{P}[\{\xi_1 = k\}] - \mathbb{P}[\{\xi_2 = k\}]|, \end{aligned} \tag{2.22}$$

so by (2.21) and (2.22) we have

$$\begin{aligned} \sum_{k=1}^{\infty} |\mathbb{P}[\{\xi_1 = k\}] - \mathbb{P}[\{\xi_2 = k\}]| &\leq \sum_{k=1}^{\infty} (\mathbb{P}[\{\xi_1 = k\} \cap \{\xi_1 \neq \xi_2\}] + \mathbb{P}[\{\xi_2 = k\} \cap \{\xi_1 \neq \xi_2\}]) \\ &= \mathbb{P}[\{\xi_1 \geq 1\} \cap \{\xi_1 \neq \xi_2\}] + \mathbb{P}[\{\xi_2 \geq 1\} \cap \{\xi_1 \neq \xi_2\}] \\ &< 4e^{-c(N-1)}. \end{aligned} \tag{2.23}$$

Thus, by (2.20) and (2.23) we obtain

$$\begin{aligned} |\mathbb{E}[\xi_1] - \mathbb{E}[\xi(B_i(\tilde{x}, a), s_2)]| &= \left| \sum_{n=1}^{\infty} \sum_{k=n}^{\infty} (\mathbb{P}[\xi_1 = k] - \mathbb{P}[\xi_2 = k]) \right| \\ &< 4N^{d/(d-1)} e^{-c(N-1)} + \sum_{n=N^{d/(d-1)}}^{\infty} (\mathbb{P}[\xi_1 \geq n] + \mathbb{P}[\xi_2 \geq n]). \end{aligned} \tag{2.24}$$

In the following we estimate the upper bound of $\sum_{n=N^{d/(d-1)}}^{\infty} \mathbb{P}[\xi_1 \geq n]$. Firstly, by (2.15), for large enough N , we obtain

$$\sum_{n=N^{d/(d-1)}}^{e^2 \lambda s_1^d} \mathbb{P}[\xi_1 \geq n] < \sum_{n=N^{d/(d-1)}}^{e^2 \lambda s_1^d} \exp(-cn^{(d-1)/d}) + e^2 \lambda s_1^d e^{-cs_1}. \tag{2.25}$$

Set $\alpha := \exp(-cN)$. Then

$$\begin{aligned} \sum_{n=N^{d/(d-1)}}^{e^2\lambda s_1^d} \exp(-cn^{(d-1)/d}) &= \sum_{n=N^{d/(d-1)}}^{e^2\lambda s_1^d} \alpha^{(nN^{-d/(d-1)})^{(d-1)/d}} \\ &< N^{d/(d-1)} \sum_{k=1}^{\infty} \alpha^{k^{(d-1)/d}} \\ &= N^{d/(d-1)} \alpha \sum_{k=1}^{\infty} \alpha^{k^{(d-1)/d-1}} \\ &< MN^{d/(d-1)}\alpha, \end{aligned} \tag{2.26}$$

where $M = \sum_{k=1}^{\infty} \exp(-c(k^{(d-1)/d} - 1)) < \infty$ is a constant.

Secondly, by [9, Lemma 1.2] we have

$$\begin{aligned} \sum_{n=e^2\lambda s_1^d+1}^{\infty} \mathbb{P}[\xi_1 \geq n] &< \sum_{n=e^2\lambda s_1^d+1}^{\infty} \mathbb{P}[\text{Po}(\lambda s_1^d) \geq n] \\ &\leq \sum_{n=e^2\lambda s_1^d+1}^{\infty} \exp\left(-\left(\frac{n}{2}\right) \log\left(\frac{n}{\lambda s_1^d}\right)\right) \\ &< \frac{e^{-(e^2\lambda s_1^d+1)}}{1 - e^{-1}}. \end{aligned} \tag{2.27}$$

Thus, by (2.25), (2.26), and (2.27), there exists a constant $c_1 > 0$ such that, for large N ,

$$\sum_{n=N^{d/(d-1)}}^{\infty} \mathbb{P}[\xi_1 \geq n] < e^{-c_1N}. \tag{2.28}$$

Using the ergodicity of Poisson point processes, in a similar way we obtain

$$\sum_{n=N^{d/(d-1)}}^{\infty} \mathbb{P}[\xi_2 \geq n] < e^{-c_1N}. \tag{2.29}$$

Combining (2.24), (2.28), and (2.29) yields the result.

Lemma 2.5. *Suppose that $d \geq 2$ and $\lambda > \lambda_c$. Let $i \in [1, d]$ be an integer and $a \in (0, 1]$ and $x_j \in [0, \infty)$, $1 \leq j \leq i$, be constants. Define the point*

$$\tilde{x}_{s,a} = \tilde{x}_{s,a}(x_1, \dots, x_i) := \left(x_1, \dots, x_i, \frac{s}{2} - a, \dots, \frac{s}{2} - a\right) \in \mathbb{R}^d.$$

Then the limit of $\mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))]$ exists and

$$\lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))] = a^{d-i} \lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_i(\tilde{x}_{s,1}, 1))]. \tag{2.30}$$

Also, if $\min_{1 \leq j \leq i} \{x_j\} = 0$ then $\lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))] > 0$.

Proof. For large enough s_1 and s_2 , suppose that $s_2 > s_1$. By (2.19), it is easy to obtain $N_{\tilde{x}_{s_1,a}, \tilde{x}_{s_2,a}}(s_1, s_2) > s_1/2 - 2$. Therefore, by Lemma 2.4 and Cauchy’s criterion for convergence, the limit of $\mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))]$ exists as $s \rightarrow \infty$.

For any constant $b \in [0, 1]$, let

$$y_{s,b} = y_{s,b}(x_1, \dots, x_i) := \left(x_1, \dots, x_i, \frac{s}{2} - 1, \dots, \frac{s}{2} - 1, \frac{s}{2} - b\right) \in \mathbb{R}^d.$$

Similarly, by Lemma 2.4 and Cauchy’s criterion, the limit of $\mathbb{E}[\xi(B_{d-1}(y_{s,b}, b))]$ exists. Define

$$f_{x_1, \dots, x_i}(b) := \lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_{d-1}(y_{s,b}, b))].$$

Since $\text{Leb}(B_{d-1}(y_{s,b}, b) \cap B_{d-1}(y_{s,1}, 1 - b)) = 0$, by the definition of ξ we have

$$\mathbb{E}[\xi(B_{d-1}(y_{s,1}, 1))] = \mathbb{E}[\xi(B_{d-1}(y_{s,1}, 1 - b))] + \mathbb{E}[\xi(B_{d-1}(y_{s,b}, b))]. \tag{2.31}$$

By (2.19), $N_{y_{s,1}, y_{s,1-b}}(s, s) > s/2 - 2$. Using Lemma 2.4 and Cauchy’s criterion, we have

$$\lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_{d-1}(y_{s,1}, 1 - b))] = \lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_{d-1}(y_{s,1-b}, 1 - b))] = f_{x_1, \dots, x_i}(1 - b).$$

Therefore, taking the limit of both sides of (2.31), we obtain

$$f_{x_1, \dots, x_i}(1) = f_{x_1, \dots, x_i}(1 - b) + f_{x_1, \dots, x_i}(b),$$

which indicates that $f_{x_1, \dots, x_i}(b) = b f_{x_1, \dots, x_i}(1)$. With a similar method, we obtain

$$\lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))] = a^{d-i} f_{x_1, \dots, x_i}(1),$$

which gives (2.30).

It remains to prove that $\lim_{s \rightarrow \infty} \mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))] > 0$ if $\min_{1 \leq j \leq i} \{x_j\} = 0$. For simplicity of exposition, we restrict ourselves to the case of $d = 2$, and the proof of this result has no essential difficulty when $d \geq 3$.

Let $\partial B(s)$ denote the boundary of $B(s)$. If $\min_{1 \leq j \leq i} \{x_j\} = 0$, then $\tilde{x}_{s,a} \in \partial B(s)$. For $x \in B_i(\tilde{x}_{s,a}, a)$, let d_x be the Euclidean distance from x to $\partial B(s)$, then $0 \leq d_x \leq 1$. Let V_x denote the connected component containing x of $G(\mathcal{H}_{\lambda,s} \cup \{x\}; 1)$. Firstly, we will show that there exists a constant $c > 0$ such that

$$\mathbb{P}[\{|V_x| = 1\} \cap \{x \in \mathcal{C}_\infty\}] \geq c \left[1 - \exp\left(\lambda \left(d_x \sqrt{1 - d_x^2} - \arccos d_x\right)\right)\right] p_\infty(\lambda). \tag{2.32}$$

Define B_x^- to be the rectangle of $(1 + d_x) \times 2$ centred at x and B_x^+ to be the rectangle of $(\frac{7}{3} + d_x) \times \frac{10}{3}$ centred at x . Divide the region of $B_x^+ \setminus B_x^-$ into 64 small rectangles with two different sizes: one size (denoted by R_1) is $\frac{1}{3} \times \frac{1}{3}$, and the other size (denoted by R_2) is $(1 + d_x)/6 \times \frac{1}{3}$; see Figure 2. The number of small rectangles with size R_1 is 40, and the number of small rectangles with size R_2 is 24. Let A_1 be the event that each of these 64 small rectangles includes at least one point of \mathcal{H}_λ . By the properties of Poisson point processes, we have

$$\begin{aligned} \mathbb{P}(A_1) &= (1 - e^{-\lambda/9})^{40} (1 - e^{-\lambda(1+d_x)/18})^{24} \\ &\geq (1 - e^{-\lambda/9})^{40} (1 - e^{-\lambda/18})^{24}. \end{aligned} \tag{2.33}$$

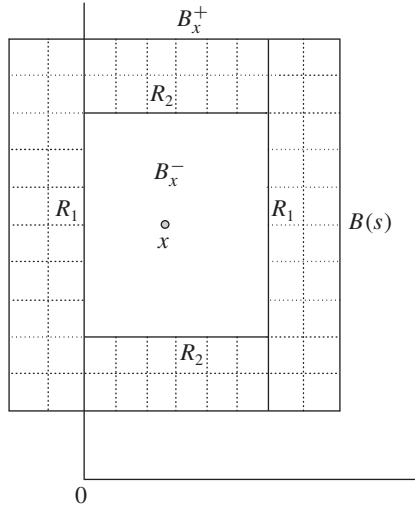


FIGURE 2: The placements of B_x^- , B_x^+ , R_1 , and R_2 .

If A_1 happens then there exists a connected component in $B_x^+ \setminus B_x^-$ which contains all the points in these small rectangles. Also, for any point in $\mathbb{R}^d \setminus B_x^-$ which can connect directly with a point in B_x^- , it must connect directly with this connected component. Let A_2 denote the event that there exists at least one point in $B_x^+ \setminus B_x^-$ contained by \mathcal{C}_∞ . So, according to the above discussion, the event $A_1 \cap A_2$ is independent with the distribution of the points of \mathcal{H}_λ in B_x^- . Therefore,

$$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1)\mathbb{P}(A_2 \mid A_1) \geq \mathbb{P}(A_1)p_\infty(\lambda). \tag{2.34}$$

Let A_3 be the event that there exists at least one point of \mathcal{H}_λ in $B(x; 1) \cap B(s)^c$, where $B(x; 1)$ denotes the d -dimensional unit ball centred at point x . By the properties of Poisson point processes we can compute

$$\mathbb{P}(A_3) = 1 - \exp\left(\lambda\left(d_x\sqrt{1 - d_x^2} - \arccos d_x\right)\right). \tag{2.35}$$

Because A_3 and $A_1 \cap A_2$ are both increasing events in $G(\mathcal{H}_\lambda; 1)$, by the Fortuin—Kasteleyn—Ginibre inequality (see [5, Theorem 2.2]) we have

$$\mathbb{P}(A_3 \cap A_1 \cap A_2) \geq \mathbb{P}(A_3)\mathbb{P}(A_1 \cap A_2). \tag{2.36}$$

If the event $A_3 \cap A_1 \cap A_2$ happens, it must be true that $x \in \mathcal{C}_\infty$. Also, the event A_3 is independent with the distribution of the points of \mathcal{H}_λ in B_x^- , so we have

$$\begin{aligned} \mathbb{P}(\{|V_x| = 1\} \cap \{x \in \mathcal{C}_\infty\}) &\geq \mathbb{P}[A_3 \cap A_1 \cap A_2 \cap \{\mathcal{H}_\lambda \cap B_x^- = \emptyset\}] \\ &= e^{-2(1+x)\lambda}\mathbb{P}(A_3 \cap A_1 \cap A_2) \\ &\geq e^{-4\lambda}\mathbb{P}(A_3 \cap A_1 \cap A_2). \end{aligned} \tag{2.37}$$

Set $c := e^{-4\lambda}(1 - e^{-\lambda/9})^{40}(1 - e^{-\lambda/18})^{24}$. Then (2.33), (2.34), (2.35), (2.36), and (2.37) yield (2.32).

Let W denote the number of the points of $\mathcal{H}_\lambda \cap B_i(\tilde{x}_{s,a}, a)$ which belong to \mathcal{C}_∞ but are isolated in $B(s)$. By the definition of $\xi(B_i(\tilde{x}_{s,a}, a))$ and Palm theory for Poisson processes, we have

$$\mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))] \geq \mathbb{E}[W] = \lambda \int_{B_i(\tilde{x}_{s,a}, a)} \mathbb{P}[\{|V_x| = 1\} \cap \{x \in \mathcal{C}_\infty\}] dx.$$

Combining this with (2.32), we obtain $\mathbb{E}[\xi(B_i(\tilde{x}_{s,a}, a))] > \frac{1}{2}c(1 - e^{(1-\pi)\lambda/4})\lambda p_\infty(\lambda)$. Our result follows.

Proof of Theorem 2.1. For simplicity of exposition, we shall prove (2.1) only in the case of $d = 3$; this proof has no essential difficulty in the $d = 2$ or $d \geq 4$ case.

Let $\eta_{ij}(s) := \mathbb{E}[\xi([0, 1] \times [i, i + 1] \times [j, j + 1], s)]$ and take $n = \lfloor s/2 \rfloor$. By symmetry, we have $\eta_{ij}(s) = \eta_{ji}(s)$; therefore,

$$\begin{aligned} \mathbb{E}[\xi([0, 1] \times [0, n]^2)] &= \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \eta_{ij}(s) \\ &= \eta_{00}(s) + \sum_{k=1}^{n-1} \left(2 \sum_{i=0}^{k-1} \eta_{ik}(s) + \eta_{kk}(s) \right). \end{aligned} \tag{2.38}$$

Set

$$a_1(s) := \eta_{00}(s) + \sum_{k=1}^{n-1} \left(2 \sum_{i=0}^{k-1} (\eta_{ik}(s) - \eta_{i,n-1}(s)) + \eta_{kk}(s) - \eta_{k,n-1}(s) \right).$$

Then, for large s and s_2 satisfying $s_2 > s$, by Lemma 2.4 we have

$$\begin{aligned} &|a_1(s) - a_1(s_2)| \\ &< 2n^2 e^{-cs/2} + \sum_{k=n}^{n_2-1} \left(2 \sum_{i=0}^{k-1} |\eta_{ik}(s_2) - \eta_{i,n_2-1}(s_2)| + |\eta_{kk}(s_2) - \eta_{k,n_2-1}(s_2)| \right) \\ &< 2n^2 e^{-cs/2} + \sum_{k=n}^{n_2-1} \left(2 \sum_{i=0}^{k-1} e^{-ck} + e^{-ck} \right) \\ &= o(e^{-cs/3}), \end{aligned} \tag{2.39}$$

where $n_2 = \lfloor s_2/2 \rfloor$ and c is the same constant appearing in Lemma 2.4. Then, by Cauchy’s criterion, the limit of $a_1(s)$ exists.

Define the point $y_i = (0, i, n) \in \mathbb{R}^3$. For any $i \in [0, n - 1]$ and large s , using Lemmas 2.4 and 2.5, we obtain

$$\begin{aligned} &\left| \mathbb{E} \left[\xi \left(B_2 \left(y_i, \frac{s}{2} - n \right) \right) \right] - \left(\frac{s}{2} - n \right) \eta_{i,n-1}(s) \right| \\ &\leq \left| \mathbb{E} \left[\xi \left(B_2 \left(y_i, \frac{s}{2} - n \right) \right) \right] - \left(\frac{s}{2} - n \right) \mathbb{E} \left[\xi \left([0, 1] \times [i, i + 1] \times \left[\frac{s}{2} - 1, \frac{s}{2} \right] \right) \right] \right| \\ &\quad + \left(\frac{s}{2} - n \right) \left| \mathbb{E} \left[\xi \left([0, 1] \times [i, i + 1] \times \left[\frac{s}{2} - 1, \frac{s}{2} \right] \right) \right] - \eta_{i,n-1}(s) \right| \\ &= o(e^{-cs/3}). \end{aligned} \tag{2.40}$$

Similarly, we obtain

$$\mathbb{E}\left[\xi\left([0, 1] \times \left[n, \frac{s}{2}\right]^2\right)\right] = \left(\frac{s}{2} - n\right)^2 \eta_{n-1,n-1}(s) + o(e^{-cs/3}). \tag{2.41}$$

Recall that $R_0 = [0, 1] \times [0, s/2]^2$. Then (2.38), (2.39), (2.40), and (2.41) yield

$$\begin{aligned} \mathbb{E}[\xi(R_0)] &= \mathbb{E}[\xi([0, 1] \times [0, n]^2)] + 2 \sum_{i=0}^{n-1} \mathbb{E}\left[\xi\left(B_2\left(y_i, \frac{s}{2} - n\right)\right)\right] \\ &\quad + \mathbb{E}\left[\xi\left([0, 1] \times \left[n, \frac{s}{2}\right]^2\right)\right] \\ &= \sum_{k=1}^{n-1} \left(2 \sum_{i=0}^{k-1} \eta_{i,n-1}(s) + \eta_{k,n-1}(s)\right) + (s - 2n) \sum_{i=0}^{n-1} \eta_{i,n-1}(s) \\ &\quad + \left(\frac{s}{2} - n\right)^2 \eta_{n-1,n-1}(s) + a_1 + o(e^{-cs/3}), \end{aligned} \tag{2.42}$$

where $a_1 := \lim_{s \rightarrow \infty} a_1(s)$. Let $b_i(s) := \eta_{i,n-1}(s) - \eta_{n-1,n-1}(s)$. Then by (2.42) we have

$$\begin{aligned} \mathbb{E}[\xi(R_0)] &= \left(\frac{s^2}{4} - 1\right) \eta_{n-1,n-1}(s) + \sum_{k=1}^{n-1} \left(2 \sum_{i=0}^{k-1} b_i(s) + b_k(s)\right) \\ &\quad + (s - 2n) \sum_{i=0}^{n-1} b_i(s) + a_1 + o(e^{-cs/3}) \\ &= \left(\frac{s^2}{4} - 1\right) \eta_{n-1,n-1}(s) + s \sum_{i=0}^{n-2} b_i(s) - 2b_0(s) - \sum_{i=1}^{n-2} (2i + 1)b_i(s) \\ &\quad + a_1 + o(e^{-cs/3}). \end{aligned} \tag{2.43}$$

Set

$$a_2(s) := \sum_{i=0}^{n-2} b_i(s) \quad \text{and} \quad a_3(s) := 2b_0(s) + \sum_{i=1}^{n-2} (2i + 1)b_i(s).$$

With a similar argument as used to obtain (2.40), it follows that there exist constants a_2 and a_3 such that

$$|a_2(s) - a_2| < 3ne^{-cs/2} \quad \text{and} \quad |a_3(s) - a_3| = o(e^{-cs/3}).$$

Also, by Lemma 2.4 and Cauchy’s criterion, there exists a constant $a_0 > 0$ such that

$$|\eta_{n-1,n-1}(s) - a_0| < e^{-c(n-1)}.$$

Substituting $a_0, a_2,$ and a_3 into (2.43), we have

$$\mathbb{E}[\xi(R_0)] = \left(\frac{s^2}{4} - 1\right) a_0 + sa_2 - a_3 + a_1 + o(e^{-cs/3}). \tag{2.44}$$

Now, using a similar argument as above, there exist constants $a_4, a_5, a_6,$ and a_7 such that

$$\mathbb{E}[\xi(R_1)] = \frac{s^2}{4} a_0 + sa_4 + a_5 + o(e^{-cs/3}) \tag{2.45}$$

and

$$\mathbb{E}[\xi(R_2)] = \frac{s^2}{4}a_0 + sa_6 + a_7 + o(e^{-cs/3}). \tag{2.46}$$

Combining (2.44), (2.45), and (2.46) with (2.7), (2.8), and Lemma 2.1, we obtain (2.1), where $\tau_1 = 6a_0 > 0$.

Using [9, Theorems 10.22 and 11.16] (which shows that $\delta > 0$), (2.1) yields (2.2).

Proof of Theorem 2.2. Given the discussion in the proof of Theorem 11.16 of [9], Equation (2.45) of [9] yields

$$\left(\frac{n}{\lambda}\right)^{-1/2} \left(L_1 \left(G \left(\mathcal{X}_n; \left(\frac{n}{\lambda}\right)^{-1/d} \right) \right) - \mathbb{E}[L_1(G(\mathcal{H}_{\lambda,s}; 1))] \right) \xrightarrow{D} \mathcal{N}(0, \delta^2),$$

where $s = (n/\lambda)^{1/d}$. Combining this and (2.1) yields the result.

3. Some applications

Our method used in the proof of Theorem 2.1 can be applied to estimate the expectation of many other random variables restricted to a box B as B becomes large; for example, the size of the biggest *open cluster* for percolation, the coverage area of the largest component for the Poisson–Boolean model, the number of open clusters or connected components for percolation and the Poisson–Boolean model, the number of open clusters or connected components with order k for percolation and the Poisson–Boolean model, the final size of a spatial epidemic mentioned in [9], and so on. We will give similar results to Theorem 2.1 for the size of the biggest open cluster and the number of open clusters for site percolation but the method can be adapted to bond percolation.

Following [2, Chapter 1], let $\mathbb{L}^d = (\mathbb{Z}^d, \mathbb{E}^d)$ denote the integer lattice with vertex set \mathbb{Z}^d and edges \mathbb{E}^d between all vertex pairs at an l_1 -distance of 1. For $d \geq 2$ we take $X = (X_x, x \in \mathbb{Z}^d)$ to be a family of independent and identically distributed Bernoulli random variables with parameter $p \in (0, 1)$. Sites $x \in \mathbb{Z}^d$ with $X_x = 1$ are called *open*; sites $x \in \mathbb{Z}^d$ with $X_x = 0$ are called *closed*. The corresponding probability measure on $\{0, 1\}^{\mathbb{Z}^d}$ is denoted by \mathbb{P}_p . The open clusters are denoted by the connected components of the subgraph of \mathbb{L}^d induced by the set of open vertices. Let C_0 denote the open cluster containing the origin. The percolation probability is $\theta(p) = \mathbb{P}_p(|C_0| = \infty)$ and the critical probability is $p_c = p_c(d) := \sup\{p : \theta(p) = 0\}$. It is well known [2] that $p_c \in (0, 1)$. If $p > p_c$, by [2, Theorem 8.1], with probability 1, there exists exactly one infinite open cluster C_∞ .

Given an integer $n > 0$, we denote by open clusters in $B(n)$ the connected components of the subgraph of the integer lattice \mathbb{L}^d induced by the set of open vertices lying in $B(n)$. Similar results to Theorem 2.1 concerned with the order of the biggest open cluster in $B(n)$ can be given as follows.

Theorem 3.1. *Suppose that $d \geq 2$ and $p \in (p_c, 1)$. Let $H(X; B(n))$ be the order of the biggest open cluster in $B(n)$. Then there exist constants $c = c(d, p) > 0$ and $\tau_i = \tau_i(d, p)$, $1 \leq i \leq d$, with $\tau_1 > 0$, such that, for all large enough n ,*

$$\mathbb{E}_p[H(X; B(n - 1))] = \theta(p)n^d - \sum_{i=1}^d \tau_i n^{d-i} + o(e^{-cn}). \tag{3.1}$$

Also, there exists a constant $\sigma = \sigma(d, p) > 0$ such that

$$H(X; B(n - 1))n^{-d/2} - \theta(p)n^{d/2} + \sum_{i=1}^{\lfloor d/2 \rfloor} \tau_i n^{d/2-i} \xrightarrow{D} \mathcal{N}(0, \sigma^2) \tag{3.2}$$

as $n \rightarrow \infty$.

Proof. Using similar methods as above, we have $\mathbb{E}_p[|\mathcal{C}_\infty \cap B(n - 1)|] = \theta(p)n^d$. Let C_1, C_2, \dots, C_M denote the components of $\mathcal{C}_\infty \cap B(n - 1)$, taken in decreasing order. Let $L(n - 1) = B(n - 1) \setminus [1, n - 2]^d$. For any $2 \leq i \leq M$, since $C_i \subset \mathcal{C}_\infty$, there exists at least one point in $L(n - 1) \cap C_i$ which connects to \mathcal{C}_∞ directly; we choose the smallest one according to the lexicographic ordering on \mathbb{Z}^d as the out-connect point. For any $x \in \mathbb{Z}^d \cap L(n - 1)$, define

$$\xi(x) := \begin{cases} |C_i| & \text{if there exists } i \in [2, M] \text{ such that } x \text{ is the out-connect point of } C_i, \\ 0 & \text{otherwise.} \end{cases}$$

Also, for an integer $j \in [0, d - 1]$, let

$$R_j := ([0, 1] \times [0, n - 1]^{d-1-j} \times [1, n - 2]^j) \cap \mathbb{Z}^d.$$

Then

$$\mathbb{E} \left[\sum_{i=2}^M |C_i| \right] = \sum_{x \in \mathbb{Z}^d \cap L(n-1)} \mathbb{E}[\xi(x)] = 2 \sum_{j=0}^{d-1} \sum_{x \in R_j} \mathbb{E}[\xi(x)].$$

With a similar argument as used in the proof of Theorem 2.1, we obtain (3.1), where

$$\tau_1 = 2d \lim_{n \rightarrow \infty} \mathbb{E} \left[\xi \left(\left(0, \left\lfloor \frac{n}{2} \right\rfloor, \dots, \left\lfloor \frac{n}{2} \right\rfloor \right) \right) \right] > 0.$$

Using [8, Theorem 3.2], (3.2) follows.

Following [2, Chapter 1.5], we define the *number of open clusters per vertex* by

$$\kappa(p) = \mathbb{E}_p(|C_0|^{-1}) = \sum_{n=1}^{\infty} \frac{1}{n} \mathbb{P}_p(|C_0| = n),$$

with the convention that $1/\infty = 0$. Similar results to Theorem 2.1 concerned with the number of open clusters in $B(n)$ can also be given as follows.

Theorem 3.2. *Suppose that $d \geq 2$ and $p \in (0, p_c) \cup (p_c, 1)$. Let $H(X; B(n))$ be the number of open clusters in $B(n)$. Then there exist constants $c = c(d, p) > 0$ and $\tau_i = \tau_i(d, p) > 0$, $1 \leq i \leq d$, with $\tau_1 > 0$, such that, for all large enough n ,*

$$\mathbb{E}_p[H(X; B(n - 1))] = \kappa(p)n^d + \sum_{i=1}^d \tau_i n^{d-i} + o(e^{-cn}). \tag{3.3}$$

Also, there exists a constant $\sigma = \sigma(d, p) > 0$ such that

$$H(X; B(n - 1))n^{-d/2} - \kappa(p)n^{d/2} - \sum_{i=1}^{\lfloor d/2 \rfloor} \tau_i n^{d/2-i} \xrightarrow{D} \mathcal{N}(0, \sigma^2) \text{ as } n \rightarrow \infty. \tag{3.4}$$

Moreover, for any constant $\varepsilon \in (0, d/2)$,

$$\begin{aligned} \mathbb{P}_p \left(\frac{H(X; B(n-1)) - \kappa(p)n^d - \sum_{i=1}^d \tau_i n^{d-i}}{\text{var}(H(X; B(n-1)))} \leq x \right) \\ = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy + o(n^{-d/2+\varepsilon}), \end{aligned} \tag{3.5}$$

where $\text{var}(\cdot)$ denotes the variance.

Proof. Let $L(n-1) = B(n-1) \setminus [1, n-2]^d$. For any $x \in B(n-1) \cap \mathbb{Z}^d$, let C_x denote the open cluster including x , and let $C_x(B(n-1))$ denote the open cluster including x in $B(n-1)$. Then $C_x(B(n-1)) \subseteq C_x$. For all open clusters C in $B(n-1)$, if $C \cap L(n-1) \neq \emptyset$, according to the lexicographic ordering on \mathbb{Z}^d we choose the smallest element of $C \cap L(n-1)$ as the indicated vertex of C . For any $x \in \mathbb{Z}^d \cap L(n-1)$, define

$$\xi(x, B(n-1)) := \begin{cases} 1 - \frac{|C_x(B(n-1))|}{|C_x|} & \text{if } x \text{ is the indicated vertex of } C_x(B(n-1)), \\ 0 & \text{otherwise.} \end{cases}$$

Note that, for any $y \in \mathbb{Z}^d \cap B(n-1)$,

$$\sum_{x \in C_y(B(n-1))} (|C_y(B(n-1))|^{-1} - |C_y|^{-1}) = 1 - \frac{|C_y(B(n-1))|}{|C_y|}.$$

Then by [2, Equation (4.7)] we have

$$\begin{aligned} H(X; B(n-1)) &= \sum_{x \in \mathbb{Z}^d \cap B(n-1)} |C_x(B(n-1))|^{-1} \\ &= \sum_{x \in \mathbb{Z}^d \cap B(n-1)} |C_x|^{-1} + \sum_{x \in \mathbb{Z}^d \cap L(n-1)} \xi(x, B(n-1)). \end{aligned} \tag{3.6}$$

Therefore, taking the expectation of both sides of (3.6), we obtain

$$\mathbb{E}_p[H(X; B(n-1))] = \kappa(p)n^d + \sum_{x \in \mathbb{Z}^d \cap L(n-1)} \mathbb{E}_p[\xi(x, B(n-1))].$$

Suppose that $1 \leq i \leq d$ and $x_j \in [0, K/2 - 1] \cap \mathbb{Z}$ for $1 \leq j \leq i$. For large integers n_1, n_2 , let

$$x = \left(x_1, \dots, x_i, \left\lfloor \frac{n_1}{2} \right\rfloor, \dots, \left\lfloor \frac{n_1}{2} \right\rfloor \right) \in \mathbb{Z}^d$$

and

$$\tilde{x} = \left(x_1, \dots, x_i, \left\lfloor \frac{n_2}{2} \right\rfloor, \dots, \left\lfloor \frac{n_2}{2} \right\rfloor \right) \in \mathbb{Z}^d.$$

Set $\tilde{B}(n_2) := B(n_2) \oplus \{x - \tilde{x}\}$. Since ξ is stationary under translations of the lattice \mathbb{L}^d , then $\xi(\tilde{x}, B(n_2))$ and $\xi(x, \tilde{B}(n_2))$ have the same distribution function. Now, let

$$n_0 = \min \left\{ \left\lfloor \frac{n_1}{2} \right\rfloor, \left\lfloor \frac{n_2}{2} \right\rfloor \right\};$$

by the definition of ξ we have

$$\begin{aligned} \mathbb{P}_p[\xi(x, B(n_1)) \neq \xi(x, \tilde{B}(n_2))] &= \mathbb{P}_p[\xi(x, B(n_1)) \neq \xi(x, \tilde{B}(n_2)), C_x \neq C_\infty] \\ &\leq \mathbb{P}_p[\text{diam}(C_x) \geq n_0, C_x \neq C_\infty] \\ &< e^{-cn_0}, \end{aligned}$$

where the last inequality follows from [2, Theorem 6.1] for $p < p_c$ and [2, Theorem 8.18] for $p > p_c$, respectively. Thus,

$$\begin{aligned} &|\mathbb{E}_p[\xi(x, B(n_1))] - \mathbb{E}_p[\xi(\tilde{x}, B(n_2))]| \\ &\leq \sum_t t |\mathbb{P}_p[\xi(x, B(n_1)) = t] - \mathbb{P}_p[\xi(x, \tilde{B}(n_2)) = t]| \\ &\leq \sum_t (\mathbb{P}_p[\xi(x, B(n_1)) = t, \xi(x, B(n_1)) \neq \xi(x, \tilde{B}(n_2))] \\ &\quad + \mathbb{P}_p[\xi(x, \tilde{B}(n_2)) = t, \xi(x, B(n_1)) \neq \xi(x, \tilde{B}(n_2))]) \\ &= 2\mathbb{P}_p[\xi(x, B(n_1)) \neq \xi(x, \tilde{B}(n_2))] \\ &< 2e^{-cn_0}. \end{aligned}$$

Therefore, $\lim_{n \rightarrow \infty} \mathbb{E}_p[\xi(x, B(n))]$ exists. In fact, a similar result to Theorem 2.4 can be deduced. Let

$$\tilde{\tau}_i(K) = \binom{d}{i} \sum_{x_j \in [0, K-1] \cup [n-K, n-1], 1 \leq j \leq i} \lim_{n \rightarrow \infty} \mathbb{E}_p \left[\xi \left(\left(x_1, \dots, x_i, \left\lfloor \frac{n}{2} \right\rfloor, \dots, \left\lfloor \frac{n}{2} \right\rfloor \right) \right) \right],$$

and let $\tau_i(K) = \sum_{j=1}^i \tilde{\tau}_j(K) \binom{d-j}{i-j} (-2K)^{i-j}$. In a similar way, we obtain (3.3). Combining (3.3) with [8, Theorem 3.1], (3.4) follows immediately. By [4, Theorem 2.1], [8, Theorem 3.1], and (3.3), (3.5) can be deduced.

It is worth noting that our results do have significance for some practical applications. In fact, the initial motivation of this paper was to provide theoretical foundation and guidance for the design of *wireless multihop networks*. Wireless multihop networks, e.g. vehicular ad hoc networks, mobile ad hoc networks, and wireless sensor networks, typically consist of a group of decentralized and self-organized nodes that communicate with each other in a peer-to-peer manner over wireless channels, and are increasingly being used in military and civilian applications [12]. Large scale wireless multihop networks are usually formulated by random geometric graphs, and the size of the largest component is a fundamental variable for a network, which plays a key role for the topology control in wireless multihop networks. However, this variable cannot be described very precisely by both former theoretic results and even computer simulations as the scale of the network grows to a very large size. Theorem 2.1 and Theorem 2.2 provide a precise estimation for this variable. Using simulations, the approximate values of the parameters $p_\infty(\lambda)$, τ_i , σ , and δ can be obtained; thus, the expression of the asymptotic size of the largest component can be well established, which has guiding significance to the topology control in wireless multihop networks.

Acknowledgements

This research was supported by the National Natural Science Foundation of China under grant numbers 61203141 and 71271204, and the Innovation Program of the Chinese Academy of Sciences under grant number kjcx-yw-s7.

References

- [1] GLAUCHE, I., KRAUSE, W., SOLLACHER, R. AND GREINER, M. (2003). Continuum percolation of wireless ad hoc communication networks. *Physica A* **325**, 577–600.
- [2] GRIMMETT, G. (1999). *Percolation*, 2nd edn. Springer, Berlin.
- [3] HEKMAT, R. AND VAN MIEGHEM, P. (2006). Connectivity in Wireless Ad Hoc Networks with a Log-normal Radio Model. *Mobile Networks Appl.* **11**, 351–360.
- [4] JIANG, J., ZHANG, S. AND GUO, T. (2010). A convergence rate in a martingale CLT for percolation clusters. *J. Graduate School Chinese Acad. Sci.* **27**, 577–583.
- [5] MEESTER, R. AND ROY, R. (1996). *Continuum Percolation*. Cambridge University Press.
- [6] PENROSE, M. D. (1991). On a continuum percolation model. *Adv. Appl. Prob.* **23**, 536–556.
- [7] PENROSE, M. D. (1995). Single linkage clustering and continuum percolation. *J. Multivariate Anal.* **53**, 94–109.
- [8] PENROSE, M. D. (2001). A central limit theorem with applications to percolation, epidemics and Boolean models. *Ann. Prob.* **29**, 1515–1546.
- [9] PENROSE, M. D. (2003). *Random Geometric Graphs*. Oxford University Press.
- [10] PENROSE, M. D. AND PISZTORA, A. (1996). Large deviations for discrete and continuous percolation. *Adv. Appl. Prob.* **28**, 29–52.
- [11] PISHRO-NIK, H., CHAN, K. AND FEKRI, F. (2009). Connectivity properties of large-scale sensor networks. *Wireless Networks* **15**, 945–964.
- [12] TA, X., MAO, G. AND ANDERSON, B. D. O. (2009). On the giant component of wireless multihop networks in the presence of shadowing. *IEEE Trans. Vehicular Tech.* **58**, 5152–5163.