

ON THE PROBABILITY OF THE EXISTENCE OF FIXED-SIZE COMPONENTS IN RANDOM GEOMETRIC GRAPHS

J. DÍAZ,* ** Universitat Politècnica de Catalunya

D. MITSCHKE,*** ETH Zürich

X. PÉREZ-GIMÉNEZ,* **** Universitat Politècnica de Catalunya

Abstract

In this work we give precise asymptotic expressions for the probability of the existence of fixed-size components at the threshold of connectivity for random geometric graphs.

Keywords: Random geometric graph; connectivity threshold; component size

2000 Mathematics Subject Classification: Primary 68R10

Secondary 60C05

1. Introduction and basic results on random geometric graphs

Recently, quite a bit of work has been done on *random geometric graphs*, owing to the importance of these graphs as theoretical models for ad hoc networks (for applications, we refer the reader to [4, Chapter 3]). Most of the theoretical results on random geometric graphs can be found in the book by Penrose [6]. In this section we succinctly recall the results needed to motivate and prove our main theorem.

Given a set of n vertices and a nonnegative real $r = r(n)$, each vertex is placed at some random position in the unit torus $[0, 1)^2$ selected independently and uniformly at random. We denote by $X_i = (x_i, y_i)$ the random position of vertex i for $i \in \{1, \dots, n\}$, and let $\mathcal{X} = \mathcal{X}(n) = \{X_1, \dots, X_n\}$. Note that, with probability 1, no two vertices choose the same position and, thus, we restrict our attention to the case in which $|\mathcal{X}| = n$. We define $G(\mathcal{X}; r)$ as the random graph having \mathcal{X} as the vertex set, and with an edge connecting each pair of vertices X_i and X_j in \mathcal{X} at distance $d(X_i, X_j) \leq r$, where $d(\cdot, \cdot)$ denotes the Euclidean distance in the torus.

Unless otherwise specified, all our stated results are asymptotic as $n \rightarrow \infty$. We use the following standard notation for the asymptotic behaviour of sequences of nonnegative numbers a_n and b_n : $a = O(b)$ if there exist constants C and n_0 such that $a_n \leq Cb_n$ for $n \geq n_0$. Furthermore, $a = \Omega(b)$ if $b = O(a)$, $a = \Theta(b)$ if $a = O(b)$ and $a = \Omega(b)$, and, finally, $a = o(b)$ if $a_n/b_n \rightarrow 0$ as $n \rightarrow \infty$. As usual, the abbreviation a.a.s. stands for *asymptotically almost surely*, i.e. with probability $1 - o(1)$. All logarithms in this paper are natural logarithms.

Let K_1 be the random variable counting the number of isolated vertices in $G(\mathcal{X}; r)$. By multiplying the probability that one vertex is isolated by the number of vertices we obtain

$$E(K_1) = n(1 - \pi r^2)^{n-1} = n \exp\{-\pi r^2 n - O(r^4 n)\}. \quad (1)$$

Received 8 May 2007; revision received 9 February 2009.

* Postal address: Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, UPC, 08034 Barcelona, Spain.

** Email address: diaz@lsi.upc.edu

*** Postal address: Institut für Theoretische Informatik, ETH Zürich, 8092 Zürich, Switzerland.

Email address: dmitsche@inf.ethz.ch

**** Email address: xperez@lsi.upc.edu

Define $\mu := n \exp\{-\pi r^2 n\}$. Observe from the previous expression that μ is closely related to $E(K_1)$. In fact, $\mu = o(1)$ if and only if $E(K_1) = o(1)$, and if $\mu = \Omega(1)$ then $E(K_1) \sim \mu$. Moreover, the asymptotic behaviour of μ characterises the connectivity of $G(\mathcal{X}; r)$. The following proposition is well known: a result similar to item 1 can be found in Corollary 3.1 of [3] and it can also be found in Section 1.4 of [6], item 2 is Theorem 13.11 of [6], and item 3 can also be found in Section 1.4 of [6]. For the sake of completeness, we give a simple proof of Proposition 1 in Section 4.

Proposition 1. *In terms of μ , the connectivity can be characterised as follows.*

1. *If $\mu \rightarrow 0$ then, a.a.s., $G(\mathcal{X}; r)$ is connected.*
2. *If $\mu = \Theta(1)$ then, a.a.s., $G(\mathcal{X}; r)$ consists of one giant component of size greater than $n/2$ and a Poisson number (with parameter μ) of isolated vertices.*
3. *If $\mu \rightarrow \infty$ then, a.a.s., $G(\mathcal{X}; r)$ is disconnected.*

From the definition of μ we have

$$\mu = \Theta(1) \quad \text{if and only if} \quad r = \sqrt{\frac{\log n \pm O(1)}{\pi n}}.$$

Therefore, we conclude that the property of connectivity of $G(\mathcal{X}; r)$ exhibits a sharp threshold at $r = \sqrt{\log n / \pi n}$. Note that the previous classification of the connectivity of $G(\mathcal{X}; r)$ indicates that, if $\mu = \Theta(1)$, the components of size 1 are predominant, and these components provide the main contribution to the connectivity of $G(\mathcal{X}; r)$. In fact, if $\mu = \Theta(1)$, the probability that $G(\mathcal{X}; r)$ has some component of size greater than 1 other than the giant component is $o(1)$.

On the other hand, Penrose [6] studied the number of components in $G(\mathcal{X}; r)$ that are isomorphic to a given fixed graph; equivalently, he studied the probability of finding *components* of a given size in $G(\mathcal{X}; r)$. However, the range of radii r covered by Penrose does not exceed the *thermodynamical threshold* $\Theta(\sqrt{1/n})$ where a giant component appears at $G(\mathcal{X}; r)$, which is below the connectivity threshold treated in the present paper. In fact, a percolation argument in [6] only shows that, with probability $1 - o(1)$, no components other than isolated vertices and the giant component exist at the connectivity threshold, without giving accurate bounds on this probability (see Section 1.4, Proposition 13.12, and Proposition 13.13 of [6]).

Throughout the paper, we will consider $G(\mathcal{X}; r)$ with $r = \sqrt{(\log n \pm O(1)) / \pi n}$. We prove that, for such a choice of r , given a fixed $\ell > 1$, the probability of having components of size exactly ℓ is $\Theta(1 / \log^{\ell-1} n)$. Moreover, in the process of the proof we characterise the different types of components that could exist for such a value of r .

2. Basic definitions and statements of results

Given a component Γ of $G(\mathcal{X}; r)$, Γ is *embeddable* if it can be mapped into the square $[r, 1 - r]^2$ by a translation in the torus. Embeddable components do not wrap around the torus.

Components which are not embeddable must have a large size (at least $\Omega(1/r)$). Sometimes several nonembeddable components can coexist together (see Figure 1).

However, there are some nonembeddable components which are so spread around the torus that they do not allow any room for other nonembeddable components. Call these components *solitary*. Clearly, we can have at most one solitary component. We cannot disprove the existence of a solitary component, since, with probability $1 - o(1)$, there exists a giant component of this nature; see Corollary 2.1 of [3] (implicitly it is also in Theorem 13.11 of [6]). For components

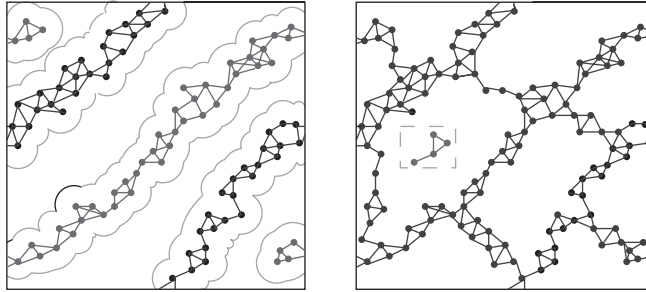


FIGURE 1: Nonembeddable components on the unit torus. *Left*: two nonembeddable and nonsolitary components. *Right*: a solitary nonembeddable component and an embeddable component.

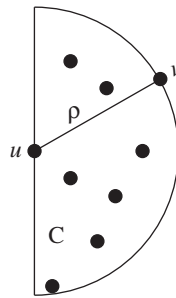


FIGURE 2: A component Γ belonging to $K'_{\varepsilon,9}$.

which are not solitary, we give asymptotic bounds on the probability of their existence according to their size.

Given a fixed integer $\ell \geq 1$, let K_ℓ be the number of components in $G(\mathcal{X}; r)$ of size exactly ℓ . For large enough n , we can assume these to be embeddable, since $r = o(1)$. Moreover, for any fixed $\varepsilon > 0$, let $K'_{\varepsilon,\ell}$ be the number of components of size exactly ℓ , which have all their vertices a distance at most εr from their leftmost vertex. Finally, \tilde{K}_ℓ denotes the number of components of size at least ℓ and which are not solitary. In Figure 2 an example of a component Γ of size exactly $\ell = 9$ is given, which has all its vertices a distance at most εr from the leftmost vertex u .

Note that $K'_{\varepsilon,\ell} \leq K_\ell \leq \tilde{K}_\ell$. However, in the following we show that, asymptotically, all the weight in the probability that $\tilde{K}_\ell > 0$ comes from components which also contribute to $K'_{\varepsilon,\ell}$ for arbitrarily small ε . This means that the more common components of size at least ℓ are cliques of size exactly ℓ with all their vertices close together.

We now have all the necessary definitions to state our main theorem, which is proved in Section 3.

Theorem 1. *Let $\ell \geq 2$ be a fixed integer. Let $0 < \varepsilon < \frac{1}{2}$ be fixed. Assume that $\mu = \Theta(1)$. Then*

$$\Pr[\tilde{K}_\ell > 0] \sim \Pr[K_\ell > 0] \sim \Pr[K'_{\varepsilon,\ell} > 0] = \Theta\left(\frac{1}{\log^{\ell-1} n}\right).$$

Given a random set \mathcal{X} of n points in $[0, 1)^2$, let $(G(\mathcal{X}; r))_{r \in \mathbb{R}^+}$ be the continuous random graph process describing the evolution of $G(\mathcal{X}; r)$ for r between 0 and $+\infty$ (\mathcal{X} remains fixed

for the whole process). Observe that the graph process starts at $r = 0$ with all n vertices being isolated, then edges are progressively added, and, finally, at $r \geq \sqrt{2}/2$ we have the complete graph on n vertices. In this context, consider the random variables $r_c = r_c(n) = \inf\{r \in \mathbb{R}^+ : G(\mathcal{X}; r) \text{ is connected}\}$ and $r_i = r_i(n) = \inf\{r \in \mathbb{R}^+ : G(\mathcal{X}; r) \text{ has no isolated vertex}\}$.

As a corollary of Theorem 1, we obtain an alternative proof of the following well-known result (see Theorem 1 of [5]): intuitively speaking, we show that, a.a.s., $(G(\mathcal{X}; r))_{r \in \mathbb{R}^+}$ becomes connected exactly at the same moment when the last isolated vertex disappears. Note that this is stronger than the results stated in the introduction, which just say that the properties of connectivity and having no isolated vertex have a sharp threshold with the same asymptotic characterization (see Proposition 1).

Corollary 1. *With probability $1 - o(1)$, we have $r_c = r_i$.*

The proof of Corollary 1 is given in Section 4.

3. Proof of Theorem 1

We state and prove three lemmas from which Theorem 1 will follow easily.

Lemma 1. *Let $\ell \geq 2$ be a fixed integer, and let $0 < \varepsilon < \frac{1}{2}$ also be fixed. Assume that $\mu = \Theta(1)$. Then,*

$$E(K'_{\varepsilon, \ell}) = \Theta\left(\frac{1}{\log^{\ell-1} n}\right).$$

Proof. First observe that, with probability 1, for each component Γ which contributes to $K'_{\varepsilon, \ell}$, Γ has a unique leftmost vertex X_i and the vertex X_j in Γ at greatest distance from X_i is also unique. Hence, we can restrict our attention to this case.

Fix an arbitrary set of indices $J \subset \{1, \dots, n\}$ of size $|J| = \ell$, with two distinguished elements i and j . Denote by $\mathcal{Y} = \bigcup_{k \in J} X_k$ the set of random points in \mathcal{X} with indices in J . Let \mathcal{E} be the following event. All vertices in \mathcal{Y} are a distance at most εr from X_i and to the right of X_i ; vertex X_j is the one in \mathcal{Y} with greatest distance from X_i ; and the vertices of \mathcal{Y} form a component Γ of $G(\mathcal{X}; r)$. If $\Pr[\mathcal{E}]$ is multiplied by the number of possible choices of i, j , and the remaining $\ell - 2$ elements of J , we obtain

$$E(K'_{\varepsilon, \ell}) = n(n - 1) \binom{n - 2}{\ell - 2} \Pr[\mathcal{E}]. \tag{2}$$

In order to bound the probability of \mathcal{E} , we need some definitions. Let $\rho = d(X_i, X_j)$, and let \mathcal{B} be the set of all points in the torus $[0, 1)^2$ which are a distance at most r from some vertex in \mathcal{Y} (see Figure 3). Note that ρ and \mathcal{B} depend on the set of random points \mathcal{Y} .

We first need bounds of $\text{Area}(\mathcal{B})$ in terms of ρ . Observe that \mathcal{B} is contained in the circle of radius $r + \rho$ and centre X_i , and, thus,

$$\text{Area}(\mathcal{B}) \leq \pi(r + \rho)^2. \tag{3}$$

Let $i_L = i, i_R, i_T$, and i_B be respectively the indices of the leftmost, rightmost, topmost, and bottommost vertices in \mathcal{Y} (some of these indices possibly equal). Assume without loss of generality that the vertical length of \mathcal{Y} (i.e. the vertical distance between X_{i_T} and X_{i_B}) is at least $\rho/\sqrt{2}$. Otherwise, the horizontal length of \mathcal{Y} has this property and we can rotate the descriptions in the argument. The upper halfcircle with centre X_{i_T} and the lower halfcircle with centre X_{i_B} are disjoint, and are contained in \mathcal{B} . If X_{i_R} is at greater vertical distance from X_{i_T}

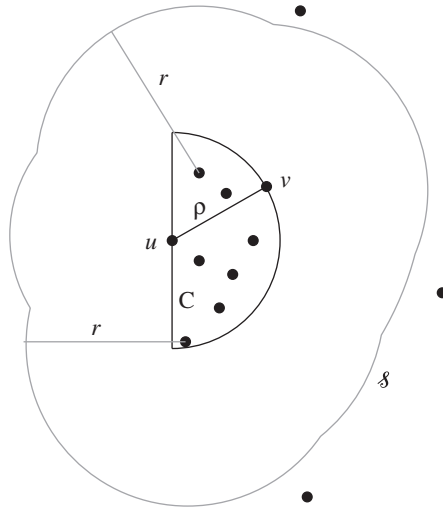


FIGURE 3: The set δ for the component Γ of Figure 2.

than from X_{i_B} , then consider the rectangle of height $\rho/(2\sqrt{2})$ and width $r - \rho/(2\sqrt{2})$ with one corner on X_{i_R} , and above and to the right of X_{i_R} . Otherwise, consider the same rectangle below and to the right of X_{i_R} . This rectangle is also contained in δ and its interior does not intersect the previously described halfcircles. Analogously, we can find another rectangle of height $\rho/(2\sqrt{2})$ and width $r - \rho/(2\sqrt{2})$ to the left of X_{i_L} and either above or below X_{i_L} with the same properties. Hence,

$$\text{Area}(\delta) \geq \pi r^2 + 2\left(\frac{\rho}{2\sqrt{2}}\right)\left(r - \frac{\rho}{2\sqrt{2}}\right). \tag{4}$$

From (3), (4), and the fact that $\rho < r/2$, we can write

$$\pi r^2\left(1 + \frac{1}{6} \frac{\rho}{r}\right) < \text{Area}(\delta) < \pi r^2\left(1 + \frac{5}{2} \frac{\rho}{r}\right) < \frac{9\pi}{4} r^2. \tag{5}$$

Now consider the probability P that the $n - \ell$ vertices not in \mathcal{Y} lie outside δ . Clearly, $P = (1 - \text{Area}(\delta))^{n-\ell}$. Moreover, by (5) and using the fact that $\exp\{-x - x^2\} \leq 1 - x \leq e^{-x}$ for all $x \in [0, \frac{1}{2}]$, we obtain

$$\exp\left\{-\left(1 + \frac{5\rho}{2r}\right)\pi r^2 n - \left(\frac{9\pi r^2}{4}\right)^2 n\right\} < P < \frac{\exp\{-(1 + \rho/(6r))\pi r^2 n\}}{(1 - 9\pi r^2/4)^\ell},$$

and after plugging in the definition of μ (recall that $\mu = n \exp\{-r^2\pi n\}$) we have

$$\left(\frac{\mu}{n}\right)^{1+5\rho/(2r)} \exp\left\{-\left(\frac{9\pi r^2}{4}\right)^2 n\right\} < P < \left(\frac{\mu}{n}\right)^{1+\rho/(6r)} \frac{1}{(1 - 9\pi r^2/4)^\ell}. \tag{6}$$

Event \mathcal{E} can also be described as follows. There is some nonnegative real $\rho \leq \varepsilon r$ such that X_j is placed at distance ρ from X_i and to the right of X_i ; all the remaining vertices in \mathcal{Y} are inside the halfcircle of centre X_i and radius ρ ; and the $n - \ell$ vertices not in \mathcal{Y} lie outside δ . Hence, $\text{Pr}[\mathcal{E}]$ can be bounded from above or below by integrating with respect to ρ the probability

density function of $d(X_i, X_j)$ times the probability that the remaining $\ell - 2$ selected vertices lie inside the right halfcircle of centre X_i and radius ρ times the upper or, respectively, lower bound on P we obtained in (6):

$$\Theta(1)I\left(\frac{5}{2}\right) \leq \Pr(\mathcal{E}) \leq \Theta(1)I\left(\frac{1}{6}\right), \tag{7}$$

where

$$I(\beta) = \int_0^{\varepsilon r} \pi \rho \left(\frac{\pi}{2} \rho^2\right)^{\ell-2} \frac{1}{n^{1+\beta\rho/r}} d\rho = \frac{2}{n} \left(\frac{\pi}{2} r^2\right)^{\ell-1} \int_0^\varepsilon x^{2\ell-3} n^{-\beta x} dx. \tag{8}$$

Since ℓ is fixed, for $\beta = \frac{5}{2}$ or $\beta = \frac{1}{6}$,

$$\begin{aligned} I(\beta) &= \Theta\left(\frac{\log^{\ell-1} n}{n^\ell}\right) \int_0^\varepsilon x^{2\ell-3} n^{-\beta x} dx \\ &= \Theta\left(\frac{\log^{\ell-1} n}{n^\ell}\right) \frac{(2\ell - 3)!}{(\beta \log n)^{2\ell-2}} \\ &= \Theta\left(\frac{1}{n^\ell \log^{\ell-1} n}\right). \end{aligned} \tag{9}$$

The statement follows from (2), (7), and (9).

Lemma 2. *Let $\ell \geq 2$ be a fixed integer. Let $\varepsilon > 0$ also be fixed. Assume that $\mu = \Theta(1)$. Then*

$$\Pr[\tilde{K}_\ell - K'_{\varepsilon,\ell} > 0] = O\left(\frac{1}{\log^\ell n}\right).$$

Proof. We assume throughout this proof that $\varepsilon \leq 10^{-18}$, and we prove the claim for this case. The case in which $\varepsilon > 10^{-18}$ follows from the fact that $(\tilde{K}_\ell - K'_{\varepsilon,\ell}) \leq (\tilde{K}_\ell - K'_{10^{-18},\ell})$.

Consider all the possible components in $G(\mathcal{X}; r)$ which are not solitary. Remove from these components the ones of size at most ℓ and diameter at most εr , and denote by M the number of remaining components. By construction, $\tilde{K}_\ell - K'_{\varepsilon,\ell} \leq M$, and, therefore, it is sufficient to prove that $\Pr[M > 0] = O(1/\log^\ell n)$. The components counted by M are classified into several types according to their size and diameter. We deal with each type separately.

Part 1. Consider all the possible components in $G(\mathcal{X}; r)$ which have diameter at most εr and size between $\ell + 1$ and $\log n/37$. Call them components of type 1, and let M_1 denote their number.

For each k , $\ell + 1 \leq k \leq \log n/37$, let E_k be the expected number of components of type 1 and size k . We observe that these components have all of their vertices a distance at most εr from the leftmost vertex. Therefore, we can apply the same argument we used for bounding $E(K'_{\varepsilon,\ell})$ in the proof of Lemma 1. Note that (2), (7), and (8) are also valid for sizes not fixed but depending on n . Thus, we obtain

$$E_k \leq O(1)n(n - 1) \binom{n - 2}{k - 2} I\left(\frac{1}{6}\right),$$

where $I(\frac{1}{6})$ is defined in (8). We use the fact that

$$\binom{n - 2}{k - 2} \leq \left(\frac{ne}{k - 2}\right)^{k-2}$$

and obtain

$$E_k = O(1) \log n \left(\frac{e \log n}{2k-2} \right)^{k-2} \int_0^\varepsilon x^{2k-3} n^{-x/6} dx. \tag{10}$$

The expression $x^{2k-3} n^{-x/6}$ can be maximized for $x \in \mathbb{R}^+$ by elementary techniques, and we deduce that

$$x^{2k-3} n^{-x/6} \leq \left(\frac{2k-3}{(e/6) \log n} \right)^{2k-3}.$$

We can bound the integral in (10) and obtain

$$\begin{aligned} E_k &= O(1) \log n \left(\frac{e \log n}{2k-2} \right)^{k-2} \varepsilon \left(\frac{2k-3}{(e/6) \log n} \right)^{2k-3} \\ &= O(1) \left(\frac{36}{2e} \frac{(2k-3)^2}{(k-2) \log n} \right)^{k-2} k. \end{aligned}$$

Note that, for $k \leq \log n/37$, the expression

$$k \left(\frac{36}{2e} \frac{(2k-3)^2}{(k-2) \log n} \right)^{k-2}$$

is decreasing with k . Hence, we can write

$$E_k = O\left(\frac{1}{\log^{\ell+1} n} \right) \text{ for all } k: \ell + 3 \leq k \leq \frac{1}{37} \log n.$$

Moreover, the bounds $E_{\ell+1} = O(1/\log^\ell n)$ and $E_{\ell+2} = O(1/\log^{\ell+1} n)$ are obtained from Lemma 1, and, hence,

$$E(M_1) = \sum_{k=\ell+1}^{\log n/37} E_k = O\left(\frac{1}{\log^\ell n} \right) + O\left(\frac{1}{\log^{\ell+1} n} \right) + \frac{\log n}{37} O\left(\frac{1}{\log^{\ell+1} n} \right) = O\left(\frac{1}{\log^\ell n} \right),$$

and then $\Pr[M_1 > 0] \leq E(M_1) = O(1/\log^\ell n)$.

Part 2. Consider all the possible components in $G(\mathcal{X}; r)$ which have diameter at most εr and size greater than $\log n/37$. Call them components of type 2, and let M_2 denote their number.

We tessellate the torus with square cells of side $y = \lfloor (\varepsilon r)^{-1} \rfloor^{-1}$ ($y \geq \varepsilon r$, but also $y \sim \varepsilon r$). We define a box to be a square of side $2y$ consisting of the union of four cells of the tessellation. Consider the set of all possible boxes. Note that any component of type 2 must be fully contained in some box (see Figure 4).

Let us fix a box b . Let W be the number of vertices which are contained inside b . Note that W has a binomial distribution with mean $E(W) = (2y)^2 n \sim (2\varepsilon)^2 \log n/\pi$. By setting $\delta = \log n/37 E(W) - 1$ and applying the Chernoff inequality to W (see, e.g. [2, Theorem 12.7]), we have

$$\begin{aligned} \Pr[W > \frac{1}{37} \log n] &= \Pr[W > (1 + \delta) E(W)] \\ &\leq \left(\frac{e^\delta}{(1 + \delta)^{1+\delta}} \right)^{E(W)} \\ &= n^{-(\log(1+\delta) - \delta/(1+\delta))/37}. \end{aligned}$$

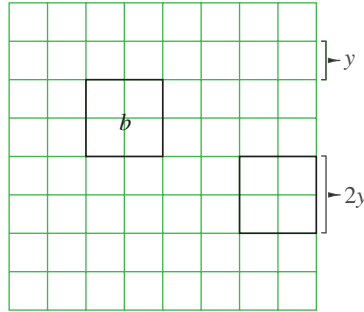


FIGURE 4: The tessellation for counting components of type 2 with two particular boxes marked.

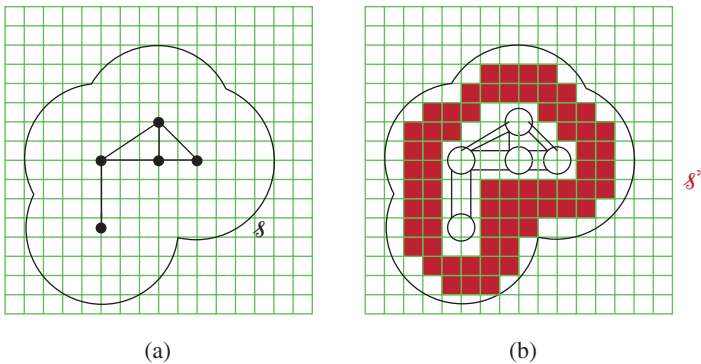


FIGURE 5: The tessellation for counting components of type 3.

Note that $\delta \sim \pi/148\epsilon^2 - 1 > e^{79}$; therefore,

$$\Pr[W > \frac{1}{37} \log n] < n^{-2.1}.$$

Taking a union bound over the set of all $\Theta(r^{-1}) = \Theta(n/\log n)$ boxes, the probability that there is some box with more than $\log n/37$ vertices is $O(1/(n^{1.1} \log n))$. Since each component of type 2 is contained in some box, we have

$$\Pr[M_2 > 0] = O\left(\frac{1}{n^{1.1} \log n}\right).$$

Part 3. Consider all the possible components in $G(\mathcal{X}; r)$ which are embeddable and have diameter at least ϵr . Call them components of type 3, and let M_3 denote their number.

We tessellate the torus into square cells of side αr for some $\alpha = \alpha(\epsilon) > 0$ fixed but sufficiently small. Let Γ be a component of type 3. Let $\mathcal{S} = \mathcal{S}_\Gamma$ be the set of all points in the torus $[0, 1)^2$ which are a distance at most r from some vertex in Γ . Remove from \mathcal{S} the vertices of Γ and the edges (represented by straight line segments), and denote by \mathcal{S}' the outer connected topologic component of the remaining set. By construction, \mathcal{S}' must contain no vertex in \mathcal{X} (see Figure 5(a)).

Now let $i_L, i_R, i_T,$ and i_B be respectively the indices of the leftmost, rightmost, topmost, and bottommost vertices in Γ (some of these indices possibly equal). As in the previous setting, assume that the vertical length of Γ (i.e. the vertical distance between X_{i_T} and X_{i_B}) is at

least $\varepsilon r/\sqrt{2}$. Otherwise, the horizontal length of Γ has this property and we can rotate the descriptions in the argument. The upper halfcircle with centre X_{i_T} and the lower halfcircle with centre X_{i_B} are disjoint and are contained in \mathcal{S}' . If X_{i_R} is at greater vertical distance from X_{i_T} than from X_{i_B} , then consider the rectangle of height $\varepsilon r/(2\sqrt{2})$ and width $r - \varepsilon r/(2\sqrt{2})$ with one corner on X_{i_R} , and above and to the right of X_{i_R} . Otherwise, consider the same rectangle below and to the right of X_{i_R} . This rectangle is also contained in \mathcal{S}' and its interior does not intersect the previously described halfcircles. Analogously, we can find another rectangle of height $\varepsilon r/(2\sqrt{2})$ and width $r - \varepsilon r/(2\sqrt{2})$ to the left of X_{i_L} , and either above or below X_{i_L} , with the same properties. Hence, taking into account the fact that $\varepsilon \leq 10^{-18}$, we have

$$\text{Area}(\mathcal{S}') \geq \pi r^2 + 2\left(\frac{\varepsilon r}{2\sqrt{2}}\right)\left(r - \frac{\varepsilon r}{2\sqrt{2}}\right) > \left(1 + \frac{\varepsilon}{5}\right)\pi r^2.$$

Let \mathcal{S}^* be the union of all the cells in the tessellation which are fully contained in \mathcal{S}' . We lose a bit of area compared to \mathcal{S}' . However, if α was chosen small enough, we can guarantee that \mathcal{S}^* is topologically connected and has an area of $\text{Area}(\mathcal{S}^*) \geq (1 + \varepsilon/6)\pi r^2$. This α can be chosen to be the same for all components of type 3 (see Figure 5(b)).

Hence, we have shown that the event $(M_3 > 0)$ implies that some connected union of cells \mathcal{S}^* of area $\text{Area}(\mathcal{S}^*) \geq (1 + \varepsilon/6)\pi r^2$ contains no vertices. By removing some cells from \mathcal{S}^* , we can assume that $(1 + \varepsilon/6)\pi r^2 \leq \text{Area}(\mathcal{S}^*) < (1 + \varepsilon/6)\pi r^2 + \alpha^2 r^2$. Let \mathcal{S}^* be any union of cells with these properties. Note that there are $\Theta(1/r^2) = \Theta(n/\log n)$ many possible choices for \mathcal{S}^* . The probability that \mathcal{S}^* contains no vertices is

$$(1 - \text{Area}(\mathcal{S}^*))^n \leq \exp\left\{-\left(1 + \frac{\varepsilon}{6}\right)\pi r^2 n\right\} = \left(\frac{\mu}{n}\right)^{1+\varepsilon/6}.$$

Therefore, we can take the union bound over all the $\Theta(n/\log n)$ possible \mathcal{S}^* , and obtain an upper bound of the probability that there is some component of type 3:

$$\Pr[M_3 > 0] \leq \Theta\left(\frac{n}{\log n}\right)\left(\frac{\mu}{n}\right)^{1+\varepsilon/6} = \Theta\left(\frac{1}{n^{\varepsilon/6} \log n}\right).$$

Part 4. Consider all the possible components in $G(\mathcal{X}; r)$ which are not embeddable and not solitary either. Call them components of type 4, and let M_4 denote their number.

We tessellate the torus $[0, 1)^2$ into $\Theta(n/\log n)$ small square cells of side length αr , where $\alpha > 0$ is a sufficiently small positive constant.

Let Γ be a component of type 4. Let $\mathcal{S} = \mathcal{S}_\Gamma$ be the set of all points in the torus $[0, 1)^2$ which are a distance at most r from some vertex in Γ . Remove from \mathcal{S} the vertices of Γ and the edges (represented by straight segments), and denote by \mathcal{S}' the remaining set. By construction, \mathcal{S}' must contain no vertex in \mathcal{X} .

Suppose that there is a horizontal or a vertical band of width $2r$ in $[0, 1)^2$ which does not intersect the component Γ (assume without loss of generality that it is the topmost horizontal band consisting of all points with the y -coordinate in $[1 - 2r, 1)$). Let us divide the torus into vertical bands of width $2r$. All of them must contain at least one vertex of Γ , since otherwise Γ would be embeddable. Select any nine consecutive vertical bands and pick one vertex of Γ with maximal y -coordinate in each one. For each one of these nine vertices, we select the left-upper quartercircle centred at the vertex if the vertex is closer to the right side of the band or the right-upper quartercircle otherwise. These nine quartercircles we chose are disjoint and must contain no vertices by construction. Moreover, they belong to the same connected component

of the set \mathcal{S}' , which we denote by \mathcal{S}'' , and which has an area of $\text{Area}(\mathcal{S}'') \geq \frac{9}{4}\pi r^2$. Let \mathcal{S}^* be the union of all the cells in the tessellation of the torus which are completely contained in \mathcal{S}'' . We lose a bit of area compared to \mathcal{S}'' . However, as usual, by choosing α small enough we can guarantee that \mathcal{S}^* is connected and that it has an area of $\text{Area}(\mathcal{S}^*) \geq \frac{11}{5}\pi r^2$. Note that this α can be chosen to be the same for all components Γ of this kind.

Suppose otherwise that all horizontal and vertical bands of width $2r$ in $[0, 1)^2$ contain at least one vertex of Γ . Since Γ is not solitary, it must be possible that it coexists with some other nonembeddable component Γ' . Then all vertical bands or all horizontal bands of width $2r$ must also contain some vertex of Γ' (assume without loss of generality that the vertical bands do). Let us divide the torus into vertical bands of width $2r$. We can find a simple path Π with vertices in Γ' which passes through 11 consecutive bands. For each one of the nine internal bands, pick the uppermost vertex of Γ in the band below Π (in the torus sense). As before, each one of these vertices contributes with a disjoint quartercircle which must be empty of vertices, and by the same argument we obtain a connected union of cells of the tessellation, which we denote by \mathcal{S}^* , with $\text{Area}(\mathcal{S}^*) \geq \frac{11}{5}\pi r^2$ and containing no vertices.

Hence, we have shown that the event $(M_4 > 0)$ implies that some connected union of cells \mathcal{S}^* with $\text{Area}(\mathcal{S}^*) \geq \frac{11}{5}\pi r^2$ contains no vertices. By repeating the same argument we used for components of type 3 but replacing $(1 + \varepsilon/6)\pi r^2$ by $\frac{11}{5}\pi r^2$, we obtain

$$\Pr[M_4 > 0] = \Theta\left(\frac{1}{n^{6/5} \log n}\right).$$

For a random variable X and any $k \geq 1$, we denote by $E[X]_k$ the k th factorial moment of X , i.e. $E(X)_k = E(X(X - 1) \cdots (X - k + 1))$.

Lemma 3. *Let $\ell \geq 2$ be a fixed integer. Let $0 < \varepsilon < \frac{1}{2}$ be fixed. Assume that $\mu = \Theta(1)$. Then*

$$E(K'_{\varepsilon, \ell})_2 = O\left(\frac{1}{\log^{2\ell-2} n}\right).$$

Proof. As in the proof of Lemma 1, we assume that each component Γ which contributes to $K'_{\varepsilon, \ell}$ has a unique leftmost vertex X_i , and the vertex X_j in Γ at greatest distance from X_i is also unique. In fact, this happens with probability 1.

Choose any two disjoint subsets of $\{1, \dots, n\}$ of size ℓ each, namely, J_1 and J_2 , with four distinguished elements $i_1, j_1 \in J_1$ and $i_2, j_2 \in J_2$. For $k \in \{1, 2\}$, denote by $\mathcal{Y}_k = \bigcup_{l \in J_k} X_l$ the set of random points in \mathcal{X} with indices in J_k . Let \mathcal{E} be the event that the following conditions hold for both $k = 1$ and $k = 2$. All vertices in \mathcal{Y}_k are a distance at most εr from X_{i_k} and to the right of X_{i_k} ; vertex X_{j_k} is the one in \mathcal{Y}_k with greatest distance from X_{i_k} ; and the vertices of \mathcal{Y}_k form a component Γ of $G(\mathcal{X}; r)$. If $\Pr[\mathcal{E}]$ is multiplied by the number of possible choices of i_k, j_k , and the remaining vertices of J_k , we obtain

$$E(K'_{\varepsilon, \ell})_2 = O(n^{2\ell}) \Pr[\mathcal{E}]. \tag{11}$$

In order to bound the probability of \mathcal{E} , we need some definitions. For each $k \in \{1, 2\}$, let $\rho_k = d(X_{i_k}, X_{j_k})$ and let \mathcal{S}_k be the set of all the points in the torus $[0, 1)^2$ which are a distance at most r from some vertex in \mathcal{Y}_k . Obviously, ρ_k and \mathcal{S}_k depend on the set of random points \mathcal{Y}_k . Also, define $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$.

Let \mathcal{F} be the event that $d(X_{i_1}, X_{i_2}) > 3r$, and let $\overline{\mathcal{F}}$ be the complementary event. This holds with probability $1 - O(r^2)$. In order to bound $\Pr[\mathcal{E} \mid \mathcal{F}]$, we apply a similar approach to

the one in the proof of Lemma 1. In fact, observe that if \mathcal{F} holds then $\mathcal{S}_1 \cap \mathcal{S}_2 = \emptyset$. Therefore, in view of (5) we can write

$$\pi r^2 \left(2 + \frac{\rho_1 + \rho_2}{6r} \right) < \text{Area}(\mathcal{S}) < \frac{18\pi}{4} r^2, \tag{12}$$

and using the same techniques that gave us (6), we obtain

$$(1 - \text{Area}(\mathcal{S}))^{n-2\ell} < \left(\frac{\mu}{n} \right)^{2+(\rho_1+\rho_2)/(6r)} \frac{1}{(1 - 18\pi r^2/4)^{2\ell}}. \tag{13}$$

Observe that \mathcal{E} can also be described as follows. For each $k \in \{1, 2\}$, there is some nonnegative real $\rho_k \leq \varepsilon r$ such that X_{j_k} is placed at distance ρ_k from X_{i_k} and to the right of X_{i_k} ; all the remaining vertices in \mathcal{Y}_k are inside the halfcircle of centre X_{i_k} and radius ρ_k ; and the $n - \ell$ vertices not in \mathcal{Y}_k lie outside \mathcal{S}_k . In fact, rather than this last condition, we only require for our bound that all vertices in $\mathcal{X} \setminus (\mathcal{Y}_1 \cup \mathcal{Y}_2)$ are placed outside \mathcal{S} , which has probability $(1 - \text{Area}(\mathcal{S}))^{n-2\ell}$. Then, from (13) and following an analogous argument to the one that leads to (7), we obtain the bound

$$\begin{aligned} \Pr[\mathcal{E} \mid \mathcal{F}] &\leq \Theta(1) \int_0^{\varepsilon r} \int_0^{\varepsilon r} \pi \rho_1 \left(\frac{\pi}{2} \rho_1^2 \right)^{\ell-2} \pi \rho_2 \left(\frac{\pi}{2} \rho_2^2 \right)^{\ell-2} \frac{1}{n^{2+(\rho_1+\rho_2)/(6r)}} d\rho_1 d\rho_2 \\ &= \Theta(1) I \left(\frac{1}{6} \right)^2, \end{aligned}$$

where $I(\frac{1}{6})$ is defined in (8). Thus, from (9) we conclude that

$$\Pr[\mathcal{E} \wedge \mathcal{F}] \leq \Theta(1) P(\mathcal{F}) I \left(\frac{1}{6} \right)^2 = O \left(\frac{1}{n^{2\ell} \log^{2\ell-2} n} \right). \tag{14}$$

Otherwise, suppose that \mathcal{F} does not hold (i.e. $d(X_{i_1}, X_{i_2}) \leq 3r$). Observe that \mathcal{E} implies that $d(X_{i_1}, X_{i_2}) > r$, since X_{i_1} and X_{i_2} must belong to different components. Hence, the circles with centres on X_{i_1} and X_{i_2} and radius r have an intersection of area less than $(\pi/2)r^2$. These two circles are contained in \mathcal{S} , and then we can write $\text{Area}(\mathcal{S}) \geq \frac{3}{2}\pi r^2$. Note that \mathcal{E} implies that all vertices in $\mathcal{X} \setminus (\mathcal{Y}_1 \cup \mathcal{Y}_2)$ are placed outside \mathcal{S} and that, for each $k \in \{1, 2\}$, all the vertices in $\mathcal{Y}_k \setminus \{X_{i_k}\}$ are a distance at most εr and to the right of X_{i_k} . This gives us the following rough bound:

$$\Pr[\mathcal{E} \mid \overline{\mathcal{F}}] \leq \left(\frac{\pi}{2} (\varepsilon r)^2 \right)^{2\ell-2} \left(1 - \frac{3\pi}{2} r^2 \right)^{n-2\ell} = O(1) \left(\frac{\log n}{n} \right)^{2\ell-2} \left(\frac{\mu}{n} \right)^{3/2}.$$

Multiplying this by $\Pr[\overline{\mathcal{F}}] = O(r^2) = O(\log n/n)$ we obtain

$$\Pr[\mathcal{E} \wedge \overline{\mathcal{F}}] = O \left(\frac{\log^{2\ell-1} n}{n^{2\ell+1/2}} \right), \tag{15}$$

which is negligible compared to (14). The statement follows from (11), (14), and (15).

Our main theorem now follows easily. From Corollary 1.12 of [1], we have

$$\mathbb{E}(K'_{\varepsilon,\ell}) - \frac{1}{2} \mathbb{E}(K'_{\varepsilon,\ell})_2 \leq \Pr[K'_{\varepsilon,\ell} > 0] \leq \mathbb{E}(K'_{\varepsilon,\ell}),$$

and, therefore, by Lemmas 1 and 3, we obtain

$$\Pr[K'_{\varepsilon,\ell} > 0] = \Theta\left(\frac{1}{\log^{\ell-1} n}\right).$$

Combining this with Lemma 2 yields the statement.

4. Proof of Corollary 1

Before proving Corollary 1, we give a proof of Proposition 1, since we will make use of the arguments used in the proof of this proposition.

Proof of Proposition 1. Recall that $\mu = n \exp\{-\pi r^2 n\}$ and $r = \sqrt{(\log n - \log \mu)/\pi n}$. Observe that $r \in [0, +\infty)$ is monotonically decreasing with respect to $\mu \in (0, n]$. Hence, the probability that $G(\mathcal{X}; r)$ is connected is also decreasing with respect to μ .

Suppose first that $\mu = \Theta(1)$. From (1) and since $O(r^4 n) = o(1)$, we have $E(K_1) \sim \mu$. We will compute the factorial moments of K_1 and show that $E(K_1)_k \sim \mu^k$ for each fixed k . As in Lemma 1, for $k \geq 2$, we fix an arbitrary set of indices $J \subset \{1, \dots, n\}$ of size $|J| = k$. Denote by $\mathcal{Y} = \bigcup_{k \in J} X_k$ the set of random points in \mathcal{X} with indices in J . Let \mathcal{E} be the event that all vertices in \mathcal{Y} are isolated, and denote by \mathcal{S} the set of points in $[0, 1]^2$ that are a distance at most r from some vertex in \mathcal{Y} . We have $E(K_1)_k \sim n^k \Pr[\mathcal{E}]$. Note that in order for the event \mathcal{E} to happen, we must have $\mathcal{S} \cap (\mathcal{X} \setminus \mathcal{Y}) = \emptyset$. To compute $\Pr[\mathcal{E}]$, we distinguish two cases.

Case 1. Let \mathcal{J}_0 be the event that, for all $i \neq j \in J$, $d(X_i, X_j) > 4r$, which has probability $1 - O(r^2) = 1 - o(1)$. Note that if \mathcal{J}_0 holds then $\text{Area}(\mathcal{S}) = kr2\pi$, and, thus, the contribution to $\Pr[\mathcal{E}]$ is

$$\Pr[\mathcal{E} \mid \mathcal{J}_0] \Pr[\mathcal{J}_0] \sim (1 - kr2\pi)^{n-k} \sim e^{-kr2\pi n}.$$

Case 2. Otherwise there exists $i \neq j \in J$ such that $d(X_i, X_j) \leq 4r$. Define

$$J' = \{j \in J \mid \text{there exists } i \in J, i < j, d(X_i, X_j) \leq 4r\},$$

and let $\ell = |J'|$. Note that $1 \leq \ell \leq k - 1$. Let j' be the smallest element of J' , and let $i' < j'$ be the (smallest) element of J with $d(X_{i'}, X_{j'}) \leq 4r$. Denote by $C_{i'}$ the circle of radius r centred at $X_{i'}$, and consider the halfcircle of radius r centred at $X_{j'}$ delimited by the line going through $X_{j'}$, perpendicular to the line connecting $X_{i'}$ with $X_{j'}$, and which does not intersect $C_{i'}$ (note that $d(X_{i'}, X_{j'}) > r$, so this halfcircle exists). This circle and halfcircle contribute to $\text{Area}(\mathcal{S})$ by $\frac{3}{2}r^2\pi$, and, thus, in total, $\text{Area}(\mathcal{S}) \geq (k - \ell + \frac{1}{2})r^2\pi$. Moreover, the probability that any $j \in J$ belongs to J' is at most $\Theta(r^2)$. Hence, if we denote by \mathcal{J}_ℓ the event that such a set J' with $|J'| = \ell$ exists, we have, for any $1 \leq \ell \leq k - 1$,

$$\Pr[\mathcal{E} \mid \mathcal{J}_\ell] \Pr[\mathcal{J}_\ell] \leq \left(1 - \left(k - \ell + \frac{1}{2}\right)r^2\pi n\right)^{r^2\pi n} \Theta(r^2)^\ell = o(\exp\{-kr^2\pi n\}).$$

Then, the main contribution to $\Pr[\mathcal{E}]$ comes from case 1, and, therefore,

$$E(K_1) \sim n^k \exp\{-kr^2\pi n\} = \mu^k,$$

so the random variable K_1 is asymptotically Poisson with parameter μ (see Theorem 1.22 of [1]). By Theorem 1, a.a.s., $G(\mathcal{X}; r)$ consists only of isolated vertices and a solitary component, and the second statement in the result is proven.

The first and third statements follow directly from the fact that, for any $\mu = \Theta(1)$,

$$\Pr[G(\mathcal{X}; r) \text{ is connected}] \sim e^{-\mu},$$

combined with the decreasing monotonicity of this probability with respect to μ .

Proof of Corollary 1. For any $\varepsilon > 0$, we can find a large enough constant $\kappa = \kappa(\varepsilon)$ such that $\exp\{-e^\kappa\} < \varepsilon/2$ and $1 - \exp\{-e^{-\kappa}\} < \varepsilon/2$. Let

$$r_\ell = \sqrt{\frac{\log n - \kappa}{\pi n}} \quad \text{and} \quad r_u = \sqrt{\frac{\log n + \kappa}{\pi n}}.$$

By Proposition 1, K_1 is asymptotically Poisson in $G(\mathcal{X}; r_\ell)$ and $G(\mathcal{X}; r_u)$ with parameter $\mu = e^\kappa$ and $\mu = e^{-\kappa}$, respectively. Therefore, in $G(\mathcal{X}; r_\ell)$ we have

$$\Pr[K_1 = 0] \sim \exp\{-e^\kappa\} < \frac{\varepsilon}{2},$$

and in $G(\mathcal{X}; r_u)$ we have

$$\Pr[K_1 > 0] \sim 1 - \exp\{-e^{-\kappa}\} < \frac{\varepsilon}{2}.$$

Moreover, by Theorem 1, a.a.s., both $G(\mathcal{X}; r_\ell)$ and $G(\mathcal{X}; r_u)$ consist only of isolated vertices and a giant solitary component. Hence, with probability at least $1 - \varepsilon$, the random process $(G(\mathcal{X}; r))_{r \in \mathbb{R}^+}$ has the following evolution: for $r \leq r_\ell$, the graph stays disconnected; at $r = r_\ell$, there are only a few isolated vertices and a giant component; for r between r_ℓ and r_u , all isolated vertices merge together or with other components; finally, for $r \geq r_u$, the graph is connected. For this particular evolution of the process, $r_c = r_i$ unless for an r with $r_\ell < r < r_u$ some isolated vertices merge together and create a small component before being absorbed by the giant component. Then, it is sufficient for our purposes to show that, a.a.s., any two isolated vertices in $G(\mathcal{X}; r_\ell)$ are at a distance greater than r_u .

Define Z to be the random variable that counts the pairs of vertices i and j , which are both isolated in $G(\mathcal{X}; r_\ell)$ and such that $d(X_i, X_j) \leq r_u$. By the same argument as in the proof of Proposition 1, setting \mathcal{B} to be the set of points in $[0, 1]^2$ at distance at most r_ℓ from either X_i or X_j , we obtain $\text{Area}(\mathcal{B}) \geq \frac{3}{2}r_\ell^2\pi$. Moreover, since $r_\ell < d(X_i, X_j) \leq r_u$, X_j must lie in an annulus of area $\Theta(1/n)$ around X_i , which occurs with probability $\Theta(1/n)$. Taking a union bound over all pairs of vertices i and j ,

$$\Pr[Z > 0] \leq n(n - 1) \left(1 - \frac{3}{2}r_\ell^2\pi\right)^{n-2} \Theta\left(\frac{1}{n}\right) = \Theta(n^{-1/2}).$$

Therefore, when gradually increasing r from r_ℓ to r_u , a.a.s. no pair of isolated vertices in $G(\mathcal{X}; r_\ell)$ gets connected before joining the solitary component, and, thus, no component of size 2 or larger (except for the solitary component) appears in this part of the process. Hence, with probability at least $1 - \varepsilon$, we have $r_c = r_i$, and the statement follows, since ε can be chosen to be arbitrarily small.

Acknowledgements

This research was partially supported by the Spanish CYCIT: TIN2007-66523 (FORMALISM). The first author was also supported by La distinció per a la promoció de la recerca de la Generalitat de Catalunya, 2002.

We thank an anonymous referee for suggesting the application of Theorem 1 to obtain Corollary 1.

References

- [1] BOLLOBÁS, B. (2001). *Random Graphs* (Camb. Stud. Adv. Math. **73**), 2nd edn. Cambridge University Press.
- [2] DÍAZ, J., PETIT, J. AND SERNA, M. (2001). A guide to concentration bounds. In *Handbook of Randomized Computing*, Vol. II, eds. S. Rajasekaran *et al.*, Kluwer, Dordrecht, pp. 457–507.
- [3] GUPTA, P. AND KUMAR, P. R. (1999). Critical power for asymptotic connectivity in wireless networks. In *Stochastic Analysis, Control, Optimization and Applications*, Birkhäuser, Boston, MA, pp. 547–566.
- [4] HEKMAT, R. (2006). *Ad-hoc Networks: Fundamental Properties and Network Topologies*, Springer, Dordrecht.
- [5] PENROSE, M. (1997). The longest edge of the random minimal spanning tree. *Ann. Appl. Prob.* **7**, 340–361.
- [6] PENROSE, M. (2003). *Random Geometric Graphs* (Oxford Stud. Prob. **5**). Oxford University Press.