Discrete Applied Mathematics 160 (2012) 2037–2047

Contents lists available at SciVerse ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Small components in k-nearest neighbour graphs

Mark Walters *

School of Mathematical Sciences, Queen Mary, University of London, London E1 4NS, England, United Kingdom

ARTICLE INFO

Article history:
Received 12 January 2011
Received in revised form 8 January 2012
Accepted 27 March 2012
Available online 3 May 2012

Keywords:
Random geometric graph
Nearest neighbour graph
Connectivity

ABSTRACT

Let $G = G_{n,k}$ denote the graph formed by placing points in a square of area n according to a Poisson process of density 1 and joining each point to its k nearest neighbours. In Balister et al. (2005) [2] Balister et al. proved that if $k < 0.3043 \log n$ then the probability that G is connected tends to 0, whereas if $k > 0.5139 \log n$ then the probability that G is connected tends to 1.

We prove that, around the threshold for connectivity, all vertices near the boundary of the square are part of the (unique) giant component. This shows that arguments about the connectivity of G do not need to consider 'boundary' effects.

We also improve the upper bound for the threshold for connectivity of G to $0.4125 \log n$.

1. Introduction

Let S_n denote a $\sqrt{n} \times \sqrt{n}$ square and let $G_{n,k}$ denote the graph formed by placing points in S_n according to a Poisson process $\mathcal{P}$ of density 1 and joining each point to its k-nearest neighbours (in Euclidean distance) by an undirected edge.

One motivation for studying this graph comes from ad hoc wireless network theory. Imagine that we have a square region with wireless transmitters scattered at random in it, and that each transmitter adjusts its power and/or protocol so that it can talk its k nearest neighbours. This gives the network $G_{n,k}$. Here, unlike a model where all transmitters have the same range (e.g., the Gilbert disc model [8]), the power of transmitters in dense regions is lower and thus contention and interference are reduced (see e.g., [9]).

It is very natural to ask when this network is connected (allowing multiple hops). Of course for any fixed n we cannot *guarantee* that the graph is connected; instead we shall consider the asymptotic behaviour of this graph as $n \to \infty$. It is convenient to introduce one piece of notation. For a graph property Π and function $k = k(n)$ we say that $G_{n,k} = G_{n,k(n)}$ has *Π with high probability* (abbreviated to whp) if $\mathbb{P}(G_{n,k(n)}$ has $\Pi) \to 1$ as $n \to \infty$.

Xue and Kumar [10] proved that the threshold for connectivity is $\Theta(\log n)$; more precisely they showed that if $k = k(n) > 5.1774 \log n$ then $G_{n,k}$ is connected whp, and if $k = k(n) < 0.074 \log n$ then $G_{n,k}$ is whp not connected.

Subsequent work by Balister et al. [2] substantially improved the upper and lower bounds to $0.5139 \log n$ and $0.3043 \log n$ respectively. In their proof they also showed that for any $k = \Theta(\log n)$ the graph consists of a giant component containing a proportion $1 - o(1)$ of all vertices and (possibly) some other 'small' components of (Euclidean) diameter $O(\sqrt{\log n})$ (for a formal statement see Lemma 3).

Moreover, they showed that if $k > 0.311 \log n$ then G has no small component within distance $O(\sqrt{\log n})$ of the boundary of S_n. Unfortunately, there is a gap between this bound and the lower bound of 0.3043 mentioned above. This means that close to the threshold for connectivity the obstruction to connectivity could occur near the boundary of the square or it could occur in the centre (their methods did rule out the possibility that the obstruction occurs in the corner of the square). This has caused several problems in later papers (e.g., [3]) where the authors had to consider both cases in their proofs.

* Fax: +44 2078827684.
E-mail address: m.walters@qmul.ac.uk.

0166-218X/$ – see front matter
doi:10.1016/j.dam.2012.03.033

Our main result is the following theorem showing that, in fact, the obstruction must occur away from the boundary of S_n. This will simplify, and indeed already has simplified (e.g., [6,5]), subsequent work in the area as only central components need to be considered.

Theorem 1. *Suppose that $G = G_{n,k}$ for some $k > 0.272 \log n$. Then there is a constant $\varepsilon > 0$ such that the probability that there exists a vertex within distance $\log n$ of the boundary of S_n that is not contained in the giant component is $O(n^{-\varepsilon})$.*

Remark. The distance $\log n$ to the boundary is much larger than the typical edge length and (non-giant) component sizes which are $O(\sqrt{\log n})$. Moreover, the theorem would still be true with $\log n$ replaced by a small power of n.

Also, the improvement itself is only of minor interest, it is the fact that the new upper bound for the existence of components near the boundary is smaller than the general lower bound that is of importance.

Our second result is the following improvement on the upper bound for connectivity of G.

Theorem 2. *Suppose that $G = G_{n,k}$ for some $k > 0.4125 \log n$. Then whp G is connected.*

To illustrate Theorem 2 let D be a disc of radius r and consider the event that there are $k + 1$ points inside D and no points in $3D \setminus D$ (where $3D$ denotes the disc with same centre as D and three times the radius). If this event occurs then the k-nearest neighbours of any point in D also lie in D: in particular, there are no 'out'-edges from D to the rest of the graph. If we choose r such that $9\pi r^2 \approx k+1$ (to maximise the probability of this event) then the probability of a specific instance of this event is about $9^{-(k+1)}$. Since we can fit $\Theta(n/\log n)$ disjoint copies of this event into S_n we see that if $k < (\frac{1}{\log 9} - \varepsilon) \log n$ (for some $\varepsilon > 0$) then whp this event occurs somewhere in S_n and thus that G has a subgraph with no out-degree. Since $1/\log 9 \approx 0.455 > 0.4125$, Theorem 2 shows that there is a range of k for which the graph is connected whp but contains pieces with no out-degree. (The corresponding result for in-degree was proved in [2].)

The proofs of these two theorems are broadly similar: they use the ideas from [2] but also consider points which are near the small component but not contained in it. If one looks at the lower bound proved in [2] we see that the density of points near the small component is higher than average. This is an unlikely event and we incorporate it into our bounds. Indeed, the above observation that there are small pieces of the graph with no out-degree shows that any proof of Theorem 2 (or any stronger bound) *must* consider points outside of a potential small component and show that they send edges in.

The key step is to split into two regimes depending on whether there is a point 'close' to the small component. If there is no such point then the 'excluded area' from the small component is quite large (which is unlikely), whereas if there is such a point then it must have a small k-nearest neighbour radius (which is also unlikely).

Related Models

Next we mention some related models and discuss their similarities and differences. There are many variants one could consider, although the model chosen in this paper is the most standard. The first is that we could work in some other region than the square such as the disc. More generally we could fix some convex region and consider taking larger and larger scalings of that region. These models would all exhibit exactly the same behaviour as in the square. Indeed, heuristically, since all edges are $O(\sqrt{\log n})$ in length the boundary would be almost flat (as the area of the region n tends to infinity) except at a finite number of 'corners'. Techniques similar to those in [2] would then yield the result. (We do not believe that anyone has written out the proof in detail however.)

Another variant is to consider the n points placed uniformly at random (rather than Poisson distributed with mean number of points n). The connectivity threshold for this model would be the same as for the model we consider. Indeed, conditioned on the number of points in a Poisson process, those points are placed uniformly at random. Moreover, since a Poisson random variable is heavily concentrated about its mean the two models are almost identical; see e.g. [4] for the analogous result for normal Erdős–Rényi random graphs. The nice feature about the Poisson process is that disjoint regions are independent: in the uniform case they are only 'nearly independent' and this makes the calculations more intricate.

Finally, we can define the directed graph $\vec{G}_{n,k}$ where each is point is joined by a directed edge to its k nearest neighbours. Thus $G_{n,k}$ is the graph formed by taking $\vec{G}_{n,k}$ and replacing each directed edge by an undirected edge. We could consider the connectivity of the directed graph $\vec{G}_{n,k}$ itself (as was done in [2]). Alternatively, we could form a different undirected graph $H_{n,k}$ from $\vec{G}_{n,k}$ by only placing an edge xy in H if *both* $\vec{xy}$ and $\vec{yx}$ are edges in $\vec{G}$ (we do not believe that the connectivity of this model has been considered although it is trivially $k = \Theta(\log n)$).

Both these models are significantly different from the model we consider: indeed in [2] Balister et al. proved that the threshold for connectivity in $\vec{G}_{n,k}$ is at least $0.7209 \log n$ and obviously the threshold for $H_{n,k}$ is at least as high.

Outline of paper

The layout of the paper is as follows. We start by introducing some notation. Then in Section 3 we make some preliminary observations about the problem and we recall the known results that we shall use later in the paper.

In Section 4 we prove Theorem 2, and then in Section 5 we prove Theorem 1. Although the two proofs are similar we prove Theorem 2 first as the absence of boundary effects makes the proof simpler and more natural in this case. We conclude with a discussion of some applications of this work and some open problems.

2. Notation

We start with some notation. For any two points $x, y \in \mathbb{R}^2$ we denote the (Euclidean) distance between them by $d(x, y)$. For any point $x \in \mathbb{R}^2$ and real number r let $D(x, r)$ denote the closed disc of radius r about x. We shall also use the term *half-disc of radius r based at x* to mean one of the four regions obtained by dividing the disc $D(x, r)$ in half vertically or horizontally.

For a set A in S_n let $|A|$ denote the (Lebesgue) measure of A, and $\#A$ denote the number of points of $\mathcal{P}$ in A. For any $x \in \mathbb{R}^2$ let $d(x, A)$ denote the distance from x to A, that is $d(x, A) = \inf_{y \in A} d(x, y)$. For any real number r let $A_{(r)}$ be the r-blowup of A defined by

$$A_{(r)} = \{x \in \mathbb{R}^2 : d(x, A) < r\}.$$

Note that we do allow $A_{(r)}$ to contain points outside of S_n.

We use the term *Poisson point-set* to mean a set of points of S_n chosen at random according to a Poisson process of density one. We shall use the term *point-set* to mean a specific instance of a Poisson point-set; however we shall ignore point-sets with infinitely many points (as the collection of such point-sets has measure zero). For any point-set $\mathcal{P}$ we let $G_k(\mathcal{P})$ denote the k-nearest neighbour graph formed from the point-set $\mathcal{P}$. More precisely for each point $x \in \mathcal{P}$ let r be the distance to its kth nearest neighbour and join x by an undirected edge to all the points of $\mathcal{P} \cap D(x, r)$.

Finally, whenever we use the term diameter we shall always mean the *Euclidean* diameter: we do not use graph diameter at any point in the paper.

3. Preliminaries

We shall wish to consider 'typical' point-sets and the 'typical' way the corresponding graphs are disconnected. In order to make this precise we start with some definitions.

Definition 1. Suppose that $\mathcal{P}$ is a point-set, that $G = G_k(\mathcal{P})$ is the corresponding graph, and that $c \geq 1$ is a real number. The point-set $\mathcal{P}$ is *c-good* if the following three conditions hold:

1. No two points of $\mathcal{P}$ at distance more than $c\sqrt{\log n}$ are joined in G,
2. Every pair of points of $\mathcal{P}$ at distance at most $\frac{1}{c}\sqrt{\log n}$ are joined in G,
3. Every half-disc of radius $c\sqrt{\log n}$ based at a point of $\mathcal{P}$ that is contained entirely in S_n contains a point of $\mathcal{P}$ other than the base-point.

Definition 2. Suppose that $\mathcal{P}$ is a point-set and $G = G_k(\mathcal{P})$ is the corresponding graph and that c is a real number.

1. A component of G is *c-small* if it has Euclidean diameter at most $c\sqrt{\log n}$.
2. A component of G is a *c-central component* if it contains no vertex within distance $c\sqrt{\log n}$ of the boundary of S_n.
3. A component of G is a *c-non-corner component* if at most one side of S_n is within distance $c\sqrt{\log n}$ of a vertex of the component.

We shall need a few results from the paper of Balister et al. [2]. Since our notation is slightly different we quote them here for convenience. The first is a slight variant of Lemma 6 of [2] which follows immediately from the proof given there (see also Lemma 1 of [3]). It says that we cannot have two components neither of which is small.

Lemma 3. *For fixed $c > 0$ and L, there exists $c_1 = c_1(c, L) > 0$, depending only on c and L, such that for any $k \geq c \log n$, the probability that $G_{n,k}$ contains two components each of (Euclidean) diameter at least $c_1\sqrt{\log n}$ is $o(n^{-L})$.*

The second bounds the probability of a small component near one side, or two sides of S_n; it is explicit in the proof of Theorem 7 of [2].

Lemma 4. *Suppose that $k = \Theta(\log n)$. The probability that there is a small component containing a vertex within $\log n$ of one boundary of S_n is $O(n^{\frac{1}{2}+o(1)}5^{-k})$ and the probability that there is a small component containing a vertex within $\log n$ of two sides of S_n is $O(n^{o(1)}3^{-k})$.*

(Note that Theorem 1 improves the first of the bounds in Lemma 4.)

The final result follows easily from concentration results for the Poisson distribution (see e.g. [1]) and most of it is implicit in Lemma 2 of [2].

Lemma 5. *For any fixed c and L there is a constant $c_2(c, L)$ such that for any k with $c \log n < k < \log n$ the probability that a Poisson point-set $\mathcal{P}$ is not c_2-good is $o(n^{-L})$.*

We shall also use the isoperimetric inequality in the plane in the following form (this is a precise statement of the fact that a disc is the shape of a given area with the smallest boundary; alternatively it can be viewed a trivial consequence of the Brunn–Minkowski inequality; see e.g., [7])

Lemma 6. *For any $\lambda > 0$ the subset A of the plane of area λ that minimises the area of the r-blowup $A_{(r)}$ is the disc of area λ.*

A simple reflection argument shows that a similar result holds in the half-plane $E_+ = \{(x, y) \in \mathbb{R}^2 : y \geq 0\}$.

Lemma 7. *For any $\lambda > 0$ a subset A of the half plane E_+ of area λ that minimises the area of the intersection of the r-blowup $A_{(r)}$ and E_+ is the half-disc of area λ centred at the origin.*

Next we prove a simple but technical lemma, which we shall use several times in the proofs of Theorems 1 and 2. It bounds the probability of certain (collections of) point-sets.

Lemma 8. *Suppose that A, B, C are three sets in S_n with $|A| \leq |C|$ and $|B| \leq |C|$ then*

$$\mathbb{P}(\#A \geq k,\ \#B \geq k,\ \#(A \cap B) = 0 \text{ and } \#C = 0) \leq \left(\frac{4|A|\,|B|}{(|A| + |B| + |C|)^2}\right)^k.$$

Proof. Suppose the event occurs. Let $A' = (A \setminus B) \setminus C$, $B' = (B \setminus A) \setminus C$, $C' = C \cup (A \cap B)$, and $U = A \cup B \cup C = A' \cup B' \cup C$. We see that A', B' and C' are pairwise disjoint so $|U| = |A'| + |B'| + |C'|$ and, since $\#(A \cap B) = 0$, that $\#A' \geq k$, $\#B' \geq k$. We have

$$\begin{aligned}
\mathbb{P}(\#A \geq k,\ \#B \geq k,\ \#(A \cap B) = 0 \text{ and } \#C = 0) &= \mathbb{P}(\#A' \geq k,\ \#B' \geq k \text{ and } \#C' = 0) \\
&= \sum_{l \geq k, m \geq k} \mathbb{P}(\#A' = l,\ \#B' = m \text{ and } \#U = l + m) \\
&= \sum_{l \geq k, m \geq k} \mathbb{P}(\#A' = l,\ \#B' = m \mid \#U = l + m)\mathbb{P}(\#U = l + m) \\
&\leq \max_{l \geq k, m \geq k} \mathbb{P}(\#A' = l,\ \#B' = m \mid \#U = l + m)
\end{aligned}$$

(the final line follows since $\sum_{l \geq k, m \geq k} \mathbb{P}(\#U = l + m) \leq 1$).

We have $|A'| \leq |A| \leq |C| \leq |C'|$ so $|A'| \leq \frac{1}{2}|U|$ and similarly $B' \leq \frac{1}{2}|U|$. Hence, for $l, m \geq k$,

$$\begin{aligned}
\mathbb{P}(\#A' = l,\ \#B' = m \mid \#U = l + m) &= \binom{l+m}{l}\left(\frac{|A'|}{|U|}\right)^l\left(\frac{|B'|}{|U|}\right)^m \\
&\leq 2^{l+m}\left(\frac{|A'|}{|U|}\right)^l\left(\frac{|B'|}{|U|}\right)^m \\
&\leq 2^{2k}\left(\frac{|A'|}{|U|}\right)^k\left(\frac{|B'|}{|U|}\right)^k \\
&= \left(\frac{4|A'|\,|B'|}{|U|^2}\right)^k \\
&= \left(\frac{4|A'|\,|B'|}{(|A'| + |B'| + |C'|)^2}\right)^k.
\end{aligned}$$

Finally, observe that $|A'| \leq |A| \leq |C| \leq |C'|$ and $|B'| \leq |B| \leq |C| \leq |C'|$ imply that

$$\begin{aligned}
\frac{4|A'|\,|B'|}{(|A'| + |B'| + |C'|)^2} &\leq \frac{4|A|\,|B'|}{(|A| + |B'| + |C'|)^2} \\
&\leq \frac{4|A|\,|B|}{(|A| + |B| + |C'|)^2} \\
&\leq \frac{4|A|\,|B|}{(|A| + |B| + |C|)^2}
\end{aligned}$$

which completes the proof. □

We conclude this section by introducing the idea of a tessellation. We will be considering tessellations of S_n with square tiles of some side length s, say. In order to relate a graph $G_k(\mathcal{P})$ to the tessellation we make the following definition.

Definition 3. Suppose that A is a set of tiles and that $\mathcal{P}$ is a point-set. We say that $\mathcal{P}$ is *disconnected on A* if $G_k(\mathcal{P})$ contains a component F such that

1. every vertex of F is in a tile of A
2. every tile of A contains a vertex of F.

Now, we may choose the size of the tiles in the tessellations (i.e. s) as we wish. Obviously the smaller the tiles are the more information they give us about $\mathcal{P}$. However, as s decreases we get more and more possible sets A (in the above definition). It is natural to choose s to be of the same scale as the events that we are interested in: namely a small constant times $\sqrt{\log n}$. Thus we consider $s = \varepsilon\sqrt{\log n}$ for some fixed $\varepsilon > 0$.

There are two reasons that this is a good size. First, in a c-good point-set a point in one tile can only be joined to points in at most a constant number of other tiles (e.g., in at most $4c^2/\varepsilon^2$ other tiles). Second, by making ε sufficiently small we can make the size of a tile much smaller than the k-nearest neighbour distance (which for a c-good point-set is between $c\sqrt{\log n}$ and $\frac{1}{c}\sqrt{\log n}$).

In the proofs of both Theorems 2 and 1 we shall use a tessellation with $\varepsilon = \frac{1}{20\,000c}$. The exact value is not important, anything smaller than this would also be fine, it was just chosen as being sufficiently small to obtain sufficient accuracy in the bounds (reducing ε further would give a negligible improvement to our bounds).

4. Central components

In this section we prove Theorem 2. Our initial aim is to prove that, for some constant c, the probability that $\mathcal{P}$ is c-good and that $G_k(\mathcal{P})$ contains a small central component tends to zero.

Lemma 9. *Suppose that $c > 0$ and that S_n is tessellated with square tiles of side length $s < \frac{1}{20\,000c}\sqrt{\log n}$. Further suppose that A is a collection of tiles of S_n containing no tile within $c\sqrt{\log n}$ of the boundary of S_n. The probability that a Poisson point-set $\mathcal{P}$ is c-good and that $\mathcal{P}$ is disconnected on A is $O(11.3^{-k})$.*

Proof. Suppose that $\mathcal{P}$ is a point-set that is c-good and that F is a component of G meeting every tile in A.

The proof of this lemma naturally divides into three steps. In the first step we define some regions based on the component F some of which must contain many points and some which must be empty. In the second step we bound the areas of these regions. In the final step we bound the probability that these regions do indeed contain the required number of points.

Step 1: *Defining the regions.* We use the following hexagonal construction which was introduced by Balister et al. in [2]. Let H be the circumscribed hexagon of the points of F obtained by taking the six tangents to the convex hull of F at angles 0 and $\pm 60°$ to the horizontal, and let $H_1, \ldots, H_6$ be the regions bounded by the exterior angle bisectors of H as in Fig. 1. Let $P_1, \ldots, P_6$ be the points of F on these tangents, and let $D_1, \ldots, D_6$ denote the k-nearest neighbour discs of $P_1, \ldots, P_6$. For $1 \le i \le 6$ let $A_i = D_i \cap H_i$. Let A_0 be the set $D_i \cap H$ with the smallest area. We see that for each $1 \le i \le 6$ the set A_i contains no points of $\mathcal{P}$. Also A_0 contains $k+1$ points all of which must be in F. Since $F \subset A$, it follows that $A' = A_0 \cap A$ contains at least $k+1$ points of $\mathcal{P}$.

We also wish to take account of points near to but not contained in F. Let $P \in F$ and $Q \in G \setminus F$ be vertices minimising the distance between F and $G \setminus F$. Let $r_0 = d(P, Q)$ and $r = r_0 - \sqrt{2}s$. Since, we are assuming that every square of A contains a point in F we see that $A_{(r)} \setminus A$ contains no point of $\mathcal{P}$. Indeed, suppose there is a point of $\mathcal{P}$ in $A_{(r)} \setminus A$. Then this point is in $G \setminus F$ and is within r_0 of some point of F which contradicts the definition of r_0.

Obviously the points Q and P are not joined so, in particular, the k points nearest to Q must all be nearer to Q than P is. Moreover, since Q is the point closest to F, we see that these k points must all be further away from P than Q is. Combining these we see that these k points lie in the set $B = D(Q, r_0) \setminus D(P, r_0)$: see Fig. 2.

Summarising all of the above, we see that A' and B each contain at least k points and $A_{(r)} \setminus A$ and $\bigcup_{i=1}^{6} A_i$ are both empty. Observe that the intersection $A' \cap B$ contains no points. Indeed, suppose that $x \in \mathcal{P}$ is in $A' \cap B$. Since x is in A' it would be in F (recall A' is part of the k-nearest neighbour disc of one of the P_i) and, since $x \in B$, we have $d(x, Q) < r_0$ contradicting the definition of the pair P, Q. However, $A_{(r)}$ and $\bigcup_{i=1}^{6} A_i$ will overlap significantly. Thus we will use Lemma 8 to form two separate bounds, one based on $A_{(r)} \setminus A$ being empty and one based on $\bigcup_{i=1}^{6} A_i$ being empty.

Step 2: *Bounding the area of the regions.*

First we bound $|\bigcup_{i=1}^{6} A_i|$. Since the point-set is c-good each disc D_i has radius at most $c\sqrt{\log n}$. By hypothesis every point P_i is in a tile of A and so is at least $c\sqrt{\log n}$ from the boundary of S_n. In particular D_i is contained in S_n for each i. Now, since the reflection of the region H_1 about the tangent through P_1 contains H, we see that $|D_1 \cap H_1| \ge |D_1 \cap H|$. A similar argument applied to each of the regions $D_i \cap H_i$ shows that $|D_i \cap H_i| \ge |D_i \cap H|$ for each $1 \le i \le 6$. Hence, $|A_i| \ge |A_0|$ for $1 \le i \le 6$. Since the H_i and therefore the A_i are disjoint, we have

$$\left|\bigcup_{i=1}^{6} A_i\right| = \sum_{i=1}^{6} |A_i| \ge 6|A_0| \ge 6|A'|.$$

The sets B and $A_{(r)}$ both depend on r so it is convenient to write r in terms of $|A'|$ by letting $x = r/(\sqrt{|A'|/\pi})$.

Since $B = D(Q, r_0) \setminus D(P, r_0)$ a simple calculation shows that $|B| = \left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right) r_0^2$. Since $\mathcal{P}$ is c-good, $r_0 > \frac{1}{c}\sqrt{\log n}$ and hence, by hypothesis, $r_0 > 20\,000s$. Thus

$$r = r_0 - \sqrt{2}s > r_0(1 - 10^{-4}).$$

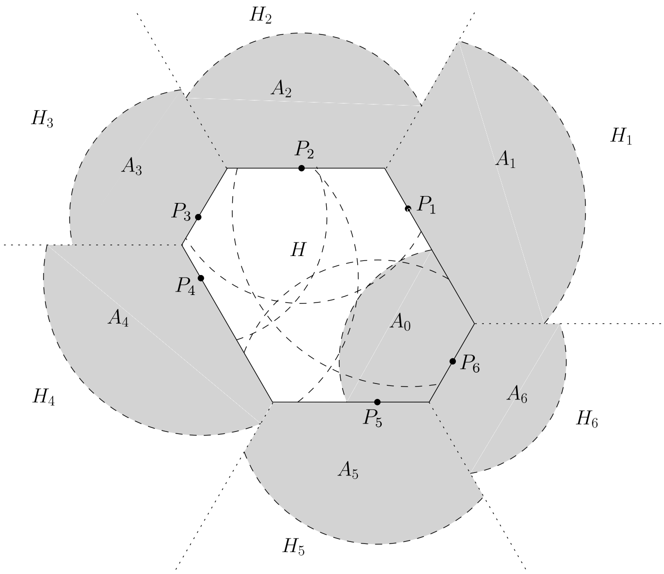

Fig. 1. The circumscribed hexagon H and associated regions.

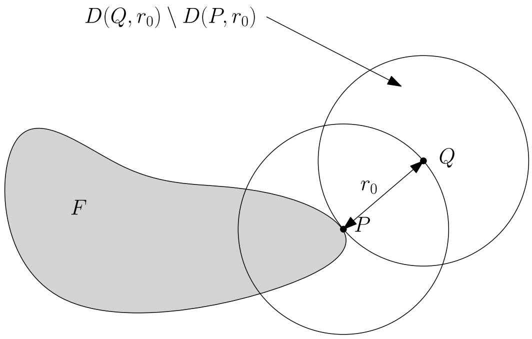

Fig. 2. The component F and the points P and Q. The k nearest neighbours of Q must all lie in the region $D(Q, r_0) \setminus D(P, r_0)$.

Hence,

$$|B| = \left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right) r_0^2 \le \left(\frac{\pi}{3} + \frac{\sqrt{3}}{2}\right) \frac{x^2|A'|}{\pi(1 - 10^{-4})^2} < 0.61x^2|A'|.$$

Finally we bound $A_{(r)}$. Let D and D' be balls of area $|A|$ and $|A'|$ respectively. Since A contains no tile within $c\sqrt{\log n}$ of the boundary of S_n, the half-disc of radius $c\sqrt{\log n}$ about the right-most point of F is contained entirely in S_n and, since the configuration is c-good, must contain a point of $\mathcal{P}$. In particular $r < r_0 \le c\sqrt{\log n}$, and so $A_{(r)}$ is contained in S_n. By the isoperimetric inequality (stated as Lemma 6) in the plane

$$|A_{(r)} \setminus A| \ge |D_{(r)} \setminus D|,$$

and it easy to see that $|D_{(r)} \setminus D| \ge |D'_{(r)} \setminus D'|$. Since D' is a ball of radius $\sqrt{|A'|/\pi}$, $D'_{(r)}$ is a ball of radius $\sqrt{|A'|/\pi} + r = (1 + x)\sqrt{|A'|/\pi}$, and we have

$$|D'_{(r)} \setminus D'| = ((x + 1)^2 - 1)|A'|.$$

Step 3: *Bounding the probability of such a configuration.* We have seen that if $\mathcal{P}$ contains such a component F then there exist regions as defined in Step 1. These regions are determined by 14 (not necessarily distinct) points: the six points defining the sides of the hexagonal hull, their six kth nearest neighbour points and the points P and Q; that is, if there is such a component F then there are 14 points of $\mathcal{P}$ defining regions A', B, $A_1, \ldots, A_6$ and $A_{(r)}$ with $\#A' \ge k$, $\#B \ge k$, $\#(A' \cap B) = 0$, and both $\#\bigcup_{i=1}^{6} A_i = 0$ and $\#(A_{(r)} \setminus A) = 0$. Moreover, if the configuration is c-good all of these points must lie within $c\sqrt{\log n}$ of A.

Let Z be the event that there are 14 points of $\mathcal{P}$ all within $c\sqrt{\log n}$ of A defining regions with the above properties. We have

$$\begin{aligned}\mathbb{P}(\text{there exists } F \text{ and the configuration is } c\text{-good}) &\leq \mathbb{P}(Z \text{ and the configuration is } c\text{-good})\\ &\leq \mathbb{P}(Z).\end{aligned}$$

We bound the probability that Z occurs (note that we are no longer assuming that the configuration is c-good). Fix a particular collection of 14 points of $\mathcal{P}$ and let Z' be the event that these particular points witness Z. Note, since we are assuming these 14 points all lie with $c\sqrt{\log n}$ of A, the corresponding regions all lie entirely within S_n.

We apply Lemma 8 to the sets A', B together with each of $\bigcup_{i=1}^{6} A_i$ and $A_{(r)} \setminus A$.

First we form the bound based on $\#\bigcup_{i=1}^{6} A_i = 0$. We have $|A'| \leq |\bigcup_{i=1}^{6} A_i|$ and, provided $x < 3.13$, we have $|B| \leq 0.61x^2|A'| < 6|A'| \leq |\bigcup_{i=1}^{6} A_i|$ so Lemma 8 applies. Thus we see that

$$\mathbb{P}(Z') \leq \left(\frac{4|A'|\,|B|}{\left(|A'| + \left|\left(\bigcup_{i=1}^{6} A_i\right)\right| + |B|\right)^2} \right)^k \leq \left(\frac{4 \cdot 0.61x^2}{(7 + 0.61x^2)^2} \right)^k.$$

Second we form a bound based on $\#(A_{(r)} \setminus A) = 0$. This time $|B| \leq 0.61x^2|A'| \leq ((x+1)^2 - 1)|A'| \leq |A_{(r)} \setminus A|$ and, provided that $x > \sqrt{2} - 1$, we have $|A_{(r)} \setminus A| \geq |A'|$ so the conditions of Lemma 8 are satisfied. Thus

$$\mathbb{P}(Z') \leq \left(\frac{4|A'|\,|B|}{(|A'| + |A_{(r)} \setminus A| + |B|)^2} \right)^k \leq \left(\frac{4 \cdot 0.61x^2}{((x+1)^2 + 0.61x^2)^2} \right)^k.$$

It is easy to check that the maximum of the minimum of these two bounds occurs when they are equal, i.e., when $x = \sqrt{7} - 1$; at this point they are α^{-k} for some $\alpha > 11.3$. Therefore $\mathbb{P}(Z') \leq \alpha^{-k}$.

Since all 14 points must lie within $c\sqrt{\log n}$ of A there are $O((\log n)^{14})$ ways of choosing them. Hence, the expected number of 14 point sets for which Z' occurs is $O((\log n)^{14}\alpha^{-k}) = O(11.3^{-k})$. Thus $\mathbb{P}(Z) = O(11.3^{-k})$ and the proof of the lemma is complete. □

Lemma 10. *Suppose that $c > 1$ is a constant and that $k = k(n) > 0.4125 \log n$. Then the probability that a Poisson point-set $\mathcal{P}$ is c-good and that $G_k(\mathcal{P})$ contains a c-small $2c$-central component is $o(1)$.*

Proof. Let $\mathcal{E}$ denote the collection of point-sets $\mathcal{P}$ which are c-good and for which $G_k(\mathcal{P})$ contains a c-small $2c$-central component.

Tile the square S_n with squares of side length $\frac{1}{20\,000c}\sqrt{\log n}$. Suppose that $\mathcal{P} \in \mathcal{E}$. Let F be a c-small $2c$-central component of $G_k(\mathcal{P})$ and let A be the collection of tiles that contain a point of F. By definition $\mathcal{P}$ is disconnected on A.

Since F is a $2c$-central component, A contains no tile within $c\sqrt{\log n}$ of the boundary of S_n. Moreover, since F is c-small it has diameter at most $c\sqrt{\log n}$ so all tiles of A lie within $c\sqrt{\log n}$ of each other.

Thus, each point-set $\mathcal{P} \in \mathcal{E}$ is disconnected on some set of tiles A where A contains no tile within $c\sqrt{\log n}$ of the boundary of S_n and all tiles of A lie within $c\sqrt{\log n}$ of each other. Let $\mathcal{A}$ be the collection of such sets of tiles A.

Next we bound the number of such sets A (i.e. we bound $|\mathcal{A}|$). There are $O(n)$ choices for the first tile in A and then all other tiles lie within distance $c\sqrt{\log n}$ of that first tile. Hence, there are only $4(c^2 \log n/s^2)$ other tiles that can be in A. Since $4(c^2 \log n/s^2)$ is just a (large) constant the total number of such sets A is $O(n)$; i.e., $|\mathcal{A}| = O(n)$.

Finally, for any fixed set $A \in \mathcal{A}$ Lemma 9 says that the probability $\mathcal{P}$ is disconnected on A is $O(11.3^{-k})$. Thus,

$$\begin{aligned}\mathbb{P}(\mathcal{P} \in \mathcal{E}) &\leq \sum_{A \in \mathcal{A}} \mathbb{P}(\mathcal{P} \text{ is disconnected on } A)\\ &= |\mathcal{A}|O(11.3^{-k})\\ &= O(n11.3^{-k})\\ &= o(1)\end{aligned}$$

where the final line follows since $k > 0.4125 \log n$. □

Proof of Theorem 2. By hypothesis we have $k > 0.4125 \log n$. Also, we may assume that $k < 0.6 \log n$ since we already know that $G_{n,k}$ is connected whp if $k \geq 0.6 \log n$. Let $c = \max\{c_1(0.4125, 0), c_2(0.4125, 0), 1\}$ where c_1 and c_2 are as given by Lemmas 3 and 5 respectively.

If $G = G_k(\mathcal{P})$ is not connected then (at least) one of the following must occur:

1. G is not c-good,
2. G contains (at least) two components neither of which is c-small,
3. G contains a c-small component that is not $2c$-central,
4. G is c-good and contains a c-small component that is $2c$-central.

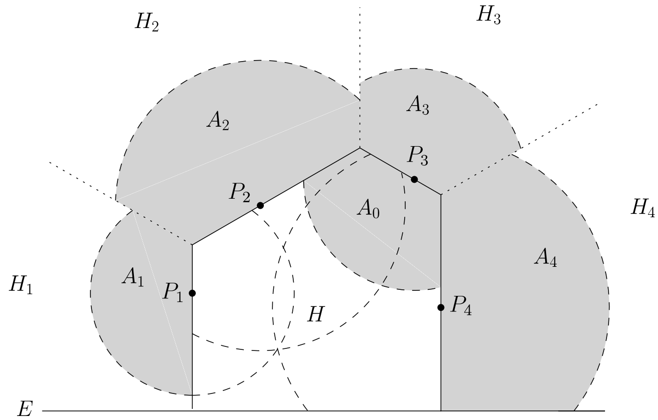

Fig. 3. The circumscribing set H and associated regions.

Since $c > c_2(0.4125, 0)$, Lemma 5 implies that the probability of Condition 1 is $o(1)$. Similarly, since $c > c_1(0.4125, 0)$, Lemma 3 shows that the probability of Condition 2 is $o(1)$. Since $k > 0.4125 \log n > \frac{1}{\log 25} \log n$, Lemma 4 implies the probability of Condition 3 is $o(1)$. Finally, since $k > 0.4125 \log n$ Lemma 10 implies that the probability of Condition 4 is $o(1)$.

Combining all these we see that the probability that G is not connected is $o(1)$; i.e., G is connected whp. $\square$

5. Boundary components

In this section we prove Theorem 1. Much of this is the same as the proof of Theorem 2 so we shall concentrate on the differences. This time our initial aim is to prove that the probability that $\mathcal{P}$ is c-good (for some c) and contains a small edge non-corner component tends to zero.

Lemma 11. *Suppose that $c > 0$ and that S_n is tessellated with square tiles of side length $s < \frac{1}{20\,000c}\sqrt{\log n}$. Further suppose that A is a collection of tiles of S_n such that at most one side of S_n is within distance $c\sqrt{\log n}$ of a tile in A. The probability a Poisson point-set $\mathcal{P}$ is c-good and that $\mathcal{P}$ is disconnected on A is $O((6.3)^{-k})$.*

Remark. Obviously this lemma is only of interest for sets A near the boundary of S_n, since otherwise Lemma 9 is stronger.

Proof. The proof divides into the same three steps as Lemma 9.

Step 1: *Defining the regions.* As before suppose that F is a component of G meeting every tile in A. Let E be the (almost surely unique) side of S_n closest to F.

This time let H be the region bounded by the four interior sides of the circumscribed hexagon of the points of F obtained by taking four of the tangents to the convex hull of F at angles 90° and ±30° to E, together with E as in Fig. 3. Let $H_1, \ldots, H_4$ be the regions bounded by the exterior angle bisectors of H together with the line E. Let $P_1, \ldots, P_4$ be the points of F on these tangents, and let $D_1, \ldots, D_4$ denote the k-nearest neighbour discs of $P_1, \ldots, P_4$. For $1 \le i \le 4$ let $A_i = D_i \cap H_i$. Let A_0 be the set $D_i \cap H$ with the smallest area and write A' for the set $A_0 \cap A$. Exactly as before we see that for $1 \le i \le 4$ the set A_i contains no points of $\mathcal{P}$, and that A' must contain at least $k + 1$ points of $\mathcal{P}$.

As before let $P \in F$ and $Q \in G \backslash F$ be vertices minimising the distance between F and $G \backslash F$, $r_0 = d(P, Q)$ and $r = r_0 - \sqrt{2}s$. Again, since F meets every tile of A we see that $A_{(r)} \setminus A$ must be empty. Also, as before, the set $B = (D(Q, r_0) \setminus D(P, r_0)) \cap S_n$ must contain at least k points.

Step 2: *Bounding the area of the regions.* First we bound $|\bigcup_{i=1}^4 A_i|$. Since the point-set $\mathcal{P}$ is c-good, each disc D_i has radius at most $c\sqrt{\log n}$ so meets no side of S_n apart from possibly E. Thus, we have $|D_i \cap H_i| \ge |D_i \cap H|$ for each $1 \le i \le 4$, so we see that $|A_i| \ge |A_0|$. As before the H_i and therefore the A_i are disjoint so

$$\left|\bigcup_{i=1}^{4} A_i\right| \ge 4|A_0| \ge 4|A'|.$$

As before let $x = r/\sqrt{|A'|/\pi}$ and exactly as in the proof of Lemma 9 we have $|B| < 0.61x^2|A'|$.

Finally we bound $A_{(r)} \setminus A$. Consider the point of F furthest from E and the half-disc of radius $c\sqrt{\log n}$ about that point facing away from E. Since no point of F is within $c\sqrt{\log n}$ of any side of S_n apart from E, this half-disc is entirely inside S_n. Thus, since the point-set is c-good this half-disc must contain a point of $\mathcal{P}$ (which is obviously not in F). Therefore, as before, $r < r_0 \le c\sqrt{\log n}$. Thus $A_{(r)} \cap S_n = A_{(r)} \cap E_+$ where E_+ denotes the half-plane bounded by E that contains S_n.

This time let D and D' be half-discs of area $|A|$ and $|A'|$ respectively centred on E. Then, by the isoperimetric inequality in the half-plane E_+ (Lemma 7)

$$|(A_{(r)} \cap E_+) \setminus A| \geq |(D_{(r)} \cap E_+) \setminus D| \geq |(D'_{(r)} \cap E_+) \setminus D'|.$$

Now D' is a half-disc of radius $\sqrt{2}\sqrt{|A'|/\pi}$ and $D'_{(r)} \cap E_+$ is a half-disc of radius $\sqrt{2}\sqrt{|A'|/\pi} + r = (1 + x/\sqrt{2})\sqrt{2|A'|/\pi}$, so this time we have

$$|(D'_{(r)} \cap E_+) \setminus D'| = ((1 + x/\sqrt{2})^2 - 1)|A'|.$$

Step 3: *Bounding the probability of such a configuration.*

We have seen that if there is such a component F then there exist regions as defined above. These regions are determined by 10 points: the four points defining the four tangents, their four kth nearest neighbour points and the points P and Q; that is, if there is such a component F then there are 10 points of $\mathcal{P}$ defining regions $A', B, A_1, \ldots, A_4$ and $A_{(r)}$ with $\#A' \geq k$, $\#B \geq k$, $\#(A' \cap B) = 0$, and both $\# \bigcup_{i=1}^4 A_i = 0$ and $\#((A_{(r)} \cap S_n) \setminus A) = 0$. Again, if the configuration is c-good, all of these 10 points must lie within $c\sqrt{\log n}$ of A.

Similarly to before, let Z be the event that there are 10 points of $\mathcal{P}$ all within $c\sqrt{\log n}$ of A defining regions with the above properties. Again

$$\begin{aligned} \mathbb{P}(\text{there exists } F \text{ and the configuration is } c\text{-good}) &\leq \mathbb{P}(Z \text{ and the configuration is } c\text{-good}) \\ &\leq \mathbb{P}(Z) \end{aligned}$$

so, as before, we bound $\mathbb{P}(Z)$.

Fix a particular collection of 10 points and let Z' be the event that these 10 points witness Z. Note, since we are assuming these 10 points all lie with $c\sqrt{\log n}$ of A, the regions $A', A_1, \ldots, A_4$ all lie entirely within S_n. By definition, B and $(A_{(r)} \cap S_n) \setminus A$ also lie in S_n.

Again we apply Lemma 8 to the sets A', B together with each of $\bigcup_{i=1}^4 A_i$ and $(A_{(r)} \cap S_n) \setminus A$. This time, however, neither bound will be valid for large x so we form a third bound based just on the two sets A' and $(A_{(r)} \cap S_n) \setminus A$.

As before we base the first bound on $\# \bigcup_{i=1}^4 A_i = 0$. We have $|A'| \leq |\bigcup_{i=1}^4 A_i|$ and, provided $x < 2.56$, we have $|B| \leq 0.61x^2|A'| < 4|A'| \leq |\bigcup_{i=1}^4 A_i|$ so Lemma 8 implies

$$\mathbb{P}(Z') \leq \left(\frac{4|A'|\,|B|}{\left(|A'| + \left|\left(\bigcup_{i=1}^4 A_i\right)\right| + |B|\right)^2} \right)^k \leq \left(\frac{4 \cdot 0.61x^2}{(5 + 0.61x^2)^2} \right)^k.$$

The second bound based on $\#((A_{(r)} \cap S_n) \setminus A) = 0$ is also very similar to before. However, this time the middle inequality in

$$|B| \leq 0.61x^2|A'| \leq ((1 + x/\sqrt{2})^2 - 1)|A'| \leq |(A_{(r)} \cap S_n) \setminus A|$$

is not valid for all x, but it is valid for all $x < 12$. Also provided that $x > 2 - \sqrt{2}$, we have $|(A_{(r)} \cap S_n) \setminus A| \geq |A'|$ so for $2 - \sqrt{2} < x < 12$ the conditions of Lemma 8 are satisfied. Thus

$$\mathbb{P}(Z') \leq \left(\frac{4|A'|\,|B|}{(|A'| + |(A_{(r)} \cap S_n) \setminus A| + |B|)^2} \right)^k \leq \left(\frac{4 \cdot 0.61x^2}{((1 + x/\sqrt{2})^2 + 0.61x^2)^2} \right)^k.$$

Since neither bound applies for large x we form a third bound based on the two sets A' and $(A_{(r)} \cap S_n) \setminus A$. We know A' contains at least k points and $(A_{(r)} \cap S_n) \setminus A$ is empty. This has probability at most

$$\mathbb{P}(Z') \leq \left(\frac{|A'|}{|A'| + |(A_{(r)} \cap S_n) \setminus A|} \right)^k \leq \frac{1}{(1 + x/\sqrt{2})^{2k}}$$

which is less than 80^{-k} for all $x \geq 12$.

As before the maximum of the minimum of the first two bounds occurs when they are equal at $x = \sqrt{2}(\sqrt{5} - 1)$; at this point they are α^{-k} for some $\alpha > 6.3$. Moreover the third bound is tiny in comparison. Thus, in all cases, $\mathbb{P}(Z') \leq \alpha^{-k}$ for some $\alpha > 6.3$.

Since all 10 points must lie within $c\sqrt{\log n}$ of A there are $O((\log n)^{10})$ ways of choosing them. Hence, similarly to before, the expected number of 10 point sets for which Z' occurs is $O((\log n)^{10}\alpha^{-k}) = O(6.3^{-k})$. Hence $\mathbb{P}(Z) = O(6.3^{-k})$ and the proof of the lemma is complete. □

Lemma 12. *Suppose that $c > 1$ is a constant and that $k = k(n) > 0.272 \log n$. Then the probability that a Poisson point-set $\mathcal{P}$ is c-good and that $G_k(\mathcal{P})$ contains a c-small $2c$-non-corner component that contains a point within $\log n$ of a side of S_n is $o(n^{-\varepsilon})$ for some $\varepsilon > 0$.*

Proof. Let $\mathcal{E}$ denote the collection of point-sets $\mathcal{P}$ which are c-good and for which $G_k(\mathcal{P})$ contains a c-small $2c$-non-corner component that contains a point within $\log n$ of a side of S_n.

As before we tile the square S_n with squares of side length $\frac{1}{20\,000c}\sqrt{\log n}$. Suppose that $\mathcal{P} \in \mathcal{E}$. Let F be a c-small $2c$-non-corner component of $G_k(\mathcal{P})$ that contains a point within $\log n$ of a side of S_n and let A be the collection of tiles that contain a point of F. By definition $\mathcal{P}$ is disconnected on A.

Since F is a $2c$-non-corner component, at most one side of S_n is within $c\sqrt{\log n}$ of a tile of A. Moreover, since F contains a point within $\log n$ of a side of S_n, A contains a tile within $\log n$ of a side of S_n. As before, since F is c-small it has diameter at most $c\sqrt{\log n}$ so all tiles of A lie within $c\sqrt{\log n}$ of each other.

Thus, each point-set $\mathcal{P} \in \mathcal{E}$ is disconnected on some set of tiles A which satisfies:

1. A includes a tile within $\log n$ of a side of S_n,
2. at most one side of S_n is within distance $c\sqrt{\log n}$ of A,
3. all tiles of A lie within $c\sqrt{\log n}$ of each other.

Let $\mathcal{A}$ be the collection of such sets of tiles A.

As before we bound the number of such sets A. Since the area of the region in S_n within $\log n$ of a side of S_n is $O((\log n)\sqrt{n})$ and each tile has area $\Theta(\log n)$ there are $O(\sqrt{n})$ choices for the tile in A that is within $\log n$ of a side of S_n. Again all other tiles lie within distance $c\sqrt{\log n}$ of that first tile and so there are only $4(c^2 \log n/s^2)$ other tiles that can be in A. Since $4(c^2 \log n/s^2)$ is just a (large) constant the total number of such sets A is $O(\sqrt{n})$; i.e., $|\mathcal{A}| = O(\sqrt{n})$.

Finally, for any fixed set $A \in \mathcal{A}$ Lemma 11 says that the probability $\mathcal{P}$ is disconnected on A is $O(6.3^{-k})$. Thus,

$$\begin{aligned}\mathbb{P}(\mathcal{P} \in \mathcal{E}) &\le \sum_{A\in\mathcal{A}} \mathbb{P}(\mathcal{P} \text{ is disconnected on } A)\\ &= |\mathcal{A}|O(6.3^{-k})\\ &= O(\sqrt{n}\,6.3^{-k})\\ &= o(n^{-\varepsilon})\end{aligned}$$

for some $\varepsilon > 0$ (where the final step follows since $k > 0.272 \log n$). $\square$

Proof of Theorem 1. By hypothesis we have $k > 0.272 \log n$. Also, we may assume that $k < 0.6 \log n$ since we already know that $G_{n,k}$ is connected whp if $k \ge 0.6 \log n$. Let $c = \max\{c_1(0.25, 1), c_2(0.25, 1), 1\}$ be as given by Lemmas 3 and 5.

Suppose that $G = G_k(\mathcal{P})$ has a vertex within $\log n$ of the boundary of S_n that is not in the unique giant component. Then (at least) one of the following must occur:

1. G is not c-good,
2. G contains (at least) two components neither of which is c-small,
3. G contains a c-small component that is not $2c$-non-corner,
4. G is c-good and contains a c-small component that is $2c$-non-corner.

Since $c > c_2(0.25, 1)$, Lemma 5 implies that the probability of Condition 1 is $o(n^{-1})$. Similarly, since $c > c_1(0.25, 1)$ Lemma 3 shows that the probability of Condition 2 is $o(n^{-1})$. Since $k > 0.272 \log n = \Omega(\log n)$, Lemma 4 implies the probability of Condition 3 is $o(n^{-0.25})$. Finally, since $k > 0.272 \log n$ Lemma 12 implies that the probability of Condition 4 is $o(n^{-\varepsilon})$ for some $\varepsilon > 0$.

Combining all these we see that the probability that G has a vertex within $\log n$ of the boundary of S_n that is not contained in the giant component is $o(n^{-\varepsilon'})$ for some $\varepsilon' > 0$ as required. $\square$

Conclusions

In this paper we have proved two results about the behaviour of the small components in the graph $G_{n,k}$, and we have shown that bounds for this type of question can be substantially improved by considering edges coming *in* to the potential small component in addition to those leaving it. Moreover, we have given a reasonably general method (i.e., using isoperimetric inequalities) for proving bounds using these 'in'-edges.

Extensive numerical simulations were done in [2] and these provide strong evidence both for the fact that boundary components are not the obstruction to connectivity (i.e., Theorem 1) and that the threshold for connectivity is close to the lower bound of $0.3043 \log n$ (Theorem 2 is a step towards this lower bound).

As we have seen the obstruction to connectivity is the existence of a small central component. There are several natural questions one can ask about these components. For example, how are small components located: are they typically well-separated? The following theorem of Falgas-Ravry [5] has answered this question in the affirmative provided that the probability that G is connected is not too small.

Theorem 13. *Suppose that $\gamma > 0$ and $k = k(n)$ is such that*

$$\mathbb{P}(G \textit{ is connected}) = \Omega(n^{-\gamma}).$$

Then whp, there do not exist two small components within distance of $O(\sqrt{\log n})$ of each other.

We remark that he used Theorem 1 of this paper to ignore components near the boundary of S_n.

There are several other questions about these small components that remain open.

Question 1. *How many vertices do small components contain?*

It is immediate from Lemma 6 of [2] (quoted as Lemma 3 of this paper) that all small components contain $O(k)$ vertices. If the lower bound construction of Balister et al. in [2] is extremal then, as the authors remark there, all small components would contain $k + O(1)$ vertices.

Question 2. *Are all the small components convex in the sense that all points of $\mathcal{P}$ within the convex hull of a small component are actually part of the small component?*

Finally, in the introduction we mentioned the related model of the graph $H_{n,k}$ (where edges are present only if both directed edges are).

Question 3. *What is the threshold for connectivity in the graph model $H_{n,k}$?*

As mentioned before it is clear that the threshold is $k = \Theta(\log n)$ but we believe no non-trivial bound has been proved.

References

[1] N. Alon, J.H. Spencer, The Probabilistic Method, third ed., in: Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2008.
[2] P. Balister, B. Bollobás, A. Sarkar, M. Walters, Connectivity of random k-nearest-neighbour graphs, Adv. Appl. Probab. 37 (1) (2005) 1–24.
[3] P. Balister, B. Bollobás, A. Sarkar, M. Walters, A critical constant for the k-nearest-neighbour model, Adv. Appl. Probab. 41 (1) (2009) 1–12.
[4] B. Bollobás, Random Graphs, second edition, in: Cambridge Studies in Advanced Mathematics, vol. 73, Cambridge University Press, Cambridge, 2001.
[5] V. Falgas-Ravry, On the distribution of small components in the k-nearest neighbours random geometric graph model, Preprint.
[6] V. Falgas-Ravry, M. Walters, Sharpness in the k-nearest neighbours random geometric graph model, Adv. Appl. Probab. (in press). ArXiv e-print 1101.3083 math.PR, January 2011.
[7] R.J. Gardner, The Brunn–Minkowski inequality, Bull. Amer. Math. Soc. (NS) 39 (3) (2002) 355–405.
[8] E.N. Gilbert, Random plane networks, J. Soc. Ind. Appl. Math. 9 (1961) 533–543.
[9] T.-C. Hou, V. Li, Transmission range control in multihop packet radio networks, IEEE Trans. Commun. 34 (1) (1986) 38–44.
[10] F. Xue, P.R. Kumar, The number of neighbors needed for connectivity of wireless networks, Wirel. Netw. 10 (2004) 169–181.