Connectivity of Metric Random Graphs

Santosh S. Venkatesh

Abstract

For any given $r \geq 0$, a random set of n points $X_1, \ldots, X_n$ in the unit disc in the Euclidean plane induces a metric random graph $\mathcal{G}(n,r)$ with vertex set $\{1,\ldots,n\}$ and edge set $\{\,\{i,j\} : |X_i - X_j| \leq r\,\}$ where $|X_i - X_j|$ denotes the Euclidean distance between the points X_i and X_j. Connectivity results analogous to the classical results of Erdös and Renyi for $G_{n,p}$ are shown to hold in this setting. In particular, if c is any real constant and $r = r_n = \sqrt{\frac{1}{n}\bigl(\log n + c + o(1)\bigr)}$ then the number of isolated vertices in the graph $\mathcal{G}(n,r)$ asymptotically satisfies a Poisson law with mean e^{-c} and, *a fortiori*, the probability that $\mathcal{G}(n,r)$ is connected tends to $e^{-e^{-c}}$. These results strengthen and expand on earlier results of Gupta and Kumar in this setting for the Euclidean case. It is also shown that these results continue to hold in general if the Euclidean norm is replaced by an ℓ^p-norm $(1 \leq p \leq \infty)$. Analogous results are shown in one dimension and extensions to higher dimensions sketched.

1 Metric Random Graphs

For $x \in \mathbb{R}^2$ write $|x|$ for the usual Euclidean norm of x. Let $\mathbb{S} = \{\, x \in \mathbb{R}^2 : |x| \leq 1 \,\}$ denote the unit radius disc in the Euclidean plane (centred at the origin for convenience). Suppose $X_1, \ldots, X_n$ is a random sequence of points drawn by independent sampling from the uniform distribution on $\mathbb{S}$. For a given $r > 0$ we drop edges between any pair of points X_i and X_j within a distance r of each other. The points X_i then induce a "metric" random graph $\mathcal{G}(n,r)$ with vertex set $\{1,\ldots,n\}$ indexed by the X_i and edge set $\{\,\{i,j\} : |X_i - X_j| \leq r\,\}$. If $\{i,j\}$ is an edge of the graph we say that the vertices i and j are *adjacent* or *communicating* in $\mathcal{G}(n,r)$.

A vertex i of the random graph $\mathcal{G}(n,r)$ is *isolated* if there are no vertices adjacent to it. Write N_0 for the number of isolated vertices. Our main result asserts a sharp limit theorem for N_0 in a suitable range.

THEOREM 1 (POISSON LAW FOR ISOLATED VERTICES) *Let c be any fixed real number and suppose $r = r_n$ varies with n such that $r_n = \sqrt{\frac{1}{n}\bigl(\log n + c + o(1)\bigr)}$ as $n \to \infty$. Then the number of vertices of $\mathcal{G}(n, r_n)$ that are isolated asymptotically has a Poisson distribution with mean e^{-c}. More precisely, for every fixed non-negative integer m, $\mathbf{P}\{N_0 = m\} \to e^{-e^{-c}}\bigl(e^{-c}\bigr)^m/m!$ as $n \to \infty$.*

The graph $\mathcal{G}(n,r)$ turns out to be almost surely comprised of isolated nodes and a single giant component. A sharp threshold function for connectivity analogous to the classical result of Erdös and Renyi [3] follows.

Key words and phrases. Random graphs, threshold functions, phase transitions, Poisson paradigm, inclusion-exclusion.

This version of the paper was printed on March 17, 2004.

THEOREM 2 (THRESHOLD FUNCTION FOR CONNECTIVITY) *As before, let* c *be any fixed real number and* $r_n = \sqrt{\frac{1}{n}\big(\log n + c + o(1)\big)}$. *Then the probability that* $\mathcal{G}(n, r_n)$ *is connected tends to* $e^{-e^{-c}}$ *as* $n \to \infty$.

Thus, for any choice of tiny $\epsilon > 0$:

(A) For all $c > -\log\log\frac{1}{1-\epsilon}$, the probability that $\mathcal{G}(n, r_n)$ is connected exceeds $1-\epsilon$ for large enough n. Note that when ϵ is small, c is large and positive.

(B) For all $c < -\log\log\frac{1}{\epsilon}$, the probability that $\mathcal{G}(n, r_n)$ is connected is less than ϵ, again for large enough n. Here, when ϵ is small, c is negative and large in absolute value. Note that this could have been deduced directly from Theorem 1 as the probability that $\mathcal{G}(n, r)$ is not connected is at least as large as the probability that there exist isolated nodes and this has probability $1 - e^{-e^{-c}} + o(1)$.

Thus, the critical radius $\sqrt{\frac{1}{n}\log n}$ is a bona fide threshold function for connectivity. In earlier work, Gupta and Kumar [5] show as their main result that if $r_n = \sqrt{\frac{1}{n}(\log n + c_n)}$, then it is necessary and sufficient that $c_n \to +\infty$ for $\mathcal{G}(n, r_n)$ to be almost surely connected. Our observation (A) constitutes a modest extension and tightening of their result. Also, a ready consequence of observation (B) is that $\mathcal{G}(n, r_n)$ is almost surely not connected if $c_n \to -\infty$. This was conjectured in Gupta and Kumar's paper. Theorems 1 and 2 provide a fine grain picture of the situation at the critical radius and are new to my knowledge.

In spite of the superficial similarity in results, the metric random graph $\mathcal{G}(n, r)$ is structurally quite different from the classical $G_{n,p}$ graph model where edges are dropped with probability p independently between vertex pairs. This is perhaps most clearly evident in the graph diameter: the diameter of $G_{n,p}$ when the graph is just connected is of the order of $\log n/\log\log n$ (see Bollobás [1]); the diameter of $\mathcal{G}(n, r)$ just when it is connected is, however, much larger, of the order of $2\sqrt{n}\,/\sqrt{\log n}$ as connectivity of $\mathcal{G}(n, r)$ will imply almost complete coverage of the unit disc by the points X_i. Unlike $G_{n,p}$, there are also pervasive statistical dependencies in the edges of $\mathcal{G}(n, r)$. Nonetheless, symmetries conferred by a weak dependency structure can be exploited to show $G_{n,p}$-like threshold functions for some graph properties, though the analogy can only be carried so far as the graph diameter illustrates.

The graph $\mathcal{G}(n, r)$ crops up in models of random deployments of wireless micro-sensor networks (see Gupta and Kumar [5] and Xue and Kumar [11], for instance); similar random structures driven by Poisson point processes have also been investigated in models of packet radio networks (Philips, Panwar, and Tantawi [10]), coverage processes (cf. Hall [7]), and in models of continuum percolation (cf. Häggström and Meester [6] and Booth, *et al* [2], for example).

The basic results of Theorems 1 and 2 can be extended in various directions.

In one direction, we can consider the effect of varying the norm. Fix any $1 \le p \le \infty$ and for $x \in \mathbb{R}^2$ write $|x|_p$ for the ℓ^p-norm of x. Suppose $X_1^{(p)}, \dots, X_n^{(p)}$ is a random sequence of points drawn by independent sampling from the unit ℓ^p-disc $\mathbb{S}^{(p)} = \{\, x : |x|_p \le 1 \,\}$. For any given $r > 0$, the random points $X_i^{(p)}$ induce a (p)-metric random graph $\mathcal{G}^{(p)}(n, r)$ with vertex set $\{1, \dots, n\}$ indexed by the $X_i^{(p)}$ and edge set $\big\{\, \{i, j\} : \big|X_i^{(p)} - X_j^{(p)}\big|_p \le r \,\big\}$.

THEOREM 3 (CONNECTIVITY IN ℓ^p-NORM) *Fix any* $1 \le p \le \infty$. *Then Theorems 1 and 2 continue to hold with* $\mathcal{G}(n, r_n) = \mathcal{G}^{(2)}(n, r_n)$ *replaced by* $\mathcal{G}^{(p)}(n, r_n)$ *in the theorem statements.*

In another direction, we may consider the effect of varying the dimension. Matters are simplest in one dimension. Suppose $X_1, \ldots, X_n$ are drawn by independent sampling from the uniform distribution on the interval $[-1, 1]$. Drop an edge between any two points X_i and X_j for which $|X_i - X_j| \leq r$. The points $X_1, \ldots, X_n$ then induce a random graph $\mathcal{G}_1(n, r)$ where we toss in the subscript to indicate that we are dealing with one dimension. The following is the analogue of Theorem 1 in one dimension.

THEOREM 4 *Let* c *be any fixed real number and* $r = r_n = \frac{1}{n}\bigl(\log n + c + o(1)\bigr)$. *Then the number of isolated nodes of* $\mathcal{G}_1(n, r_n)$ *tends asymptotically to a Poisson law with mean* e^{-c}.

The corresponding connectivity result is captured in

THEOREM 5 *Under the conditions of Theorem 4, the probability that* $\mathcal{G}_1(n, r_n)$ *is connected tends to* $e^{-e^{-c}}$ *as* $n \to \infty$.

Boundary effects that are sub-dominant in one and two dimensions become pronounced in dimensions greater than two signficantly complicating the analysis. In an alternative setting that sidesteps this issue we may consider random ensembles of points on a torus or on the surface of the unit ball in d-dimensions. There are still edge dependencies in these settings but boundary effects will now be abeyant. We briefly consider these issues in the final section of the paper and outline the nature of the results that may be obtained.

The basic idea behind our proofs is classical: we set up a Poisson paradigm for rare events via inclusion and exclusion. The technical difficulties are significant, however, and deal with handling weak but ubiquitous edge dependencies on the one hand and boundary effects on the other. Most of the effort will be expended in proving the Poisson law for isolated nodes in the unit disc (in two dimensions) that is the content of Theorem 1. The other results will follow quickly from the proof of this result. Before proceeding to the proofs, a word on

Notation: All logarithms are to base e, $\mathbf{P}$ stands for probability measure in the underlying probability space, and $\mathbf{E}$ stands for expectation. We use the following variants of the Landau order notation. Suppose $\{f_n\}$ and $\{g_n\}$ are real sequences. As $n \to \infty$ we say that:

- $f_n = \mathcal{O}(g_n)$ if $|f_n| \leq K|g_n|$ for some positive constant K.
- $f_n = o(g_n)$ if $|f_n|/|g_n| \to 0$.
- $f_n \sim g_n$ if $f_n/g_n \to 1$.

Note that the "big-O" and "small-o" variants in use here refer to absolute values. In consequence we will occasionally encounter an expression of the form $1+o(1)$ for a probability—the context makes it clear that the small-o order term must in fact be a negative, asymptotically vanishing quantity. The order terms we encounter will all be ultimately negligible and their signs will not matter. We will also commit the mild notational solecism of stringing together order relations: an expression of the form $f_n = \mathcal{O}(g_n) = o(h_n)$ means (a) that $f_n = \mathcal{O}(g_n)$ and (b) that $g_n = o(h_n)$. Of course, this also implies that $f_n = o(h_n)$ and we compact this syllogism into a single equation in the hope that there is no danger of confusion as to what is meant.

2 Poisson Law for Isolated Vertices

For the rest of this section and the next we will be concerned with metric random graphs $\mathcal{G}(n, r)$ induced by a random ensemble of points $X_1, \ldots, X_n$ drawn by independent sampling from the uniform distribution on the unit disc $\mathbb{S} = \{ x \in \mathbb{R}^2 : |x| \leq r \}$ in the plane. The norm $|x|$ is the ℓ^2 or Euclidean norm.

Let us first set up some notation. Write L_i for the event that vertex i is isolated and, in an extension of this notation, $L_{i_1,\dots,i_k} = L_{i_1} \cap \cdots \cap L_{i_k}$ for the event that vertices i_1, …, i_k are all isolated. It is clear that the events L_i $(1 \leq i \leq n)$ are exchangeable so that $\mathbf{P}(L_i)$ does not depend on the specific choice of i and, likewise, $\mathbf{P}(L_{i_1,\dots,i_k})$ do not depend on the choices of i_1, …, i_k (though, of course, these probabilities all implicitly depend upon n). We assume through this section and the next that c is any fixed real number and $r = r_n = \sqrt{\frac{1}{n}\bigl(\log n + c + o(1)\bigr)}$. In a nod to notational economy we will frequently write simply r with the dependence on n implicit.

The key first step in the proof of Theorem 1 is the demonstration that the expected number of isolated vertices tends to e^{-c}. Equivalently, we show

LEMMA 1 *The probability that any given vertex i is isolated satisfies $\mathbf{P}(L_i) \sim \frac{e^{-c}}{n}$ as $n \to \infty$.*

The second key step shows that the following weak "asymptotic independence" property holds for the events L_i.

LEMMA 2 *Let k be any fixed positive integer. Then the probability that any k distinct vertices i_1, …, i_k are all isolated satisfies $\mathbf{P}(L_{i_1,\dots,i_k}) \sim \frac{(e^{-c})^k}{n^k}$ as $n \to \infty$.*

With these two results in hand, a straightforward inclusion-exclusion argument serves to catch the aberrant vertices in a Poisson sieve.

2.1 Single Vertex Isolation

We begin with a proof of Lemma 1. This result holds the key to the proof of the theorem and the analysis will be in repeated use in subsequent stages. It will be convenient to write $P_0 = \mathbf{P}(L_i)$ as the probability of vertex isolation does not depend on the choice of vertex.

Vertex i communicates with vertices in the region defined by the intersection of the unit circle with the circle of radius r centred at X_i. We call this the *region of visibility* of vertex i. By symmetry it is clear that the area $A = A(X_i)$ of the region of visibility depends only on the distance of X_i from the origin.

Given X_i, the conditional probability that any other vertex j is adjacent to i is given by $a(X_i) = A(X_i)/\pi$ so that the number of vertices adjacent to i corresponds to the number of successes in $n-1$ tosses of a bent coin with success probability $a(X_i)$. Given X_i, the conditional probability that vertex i is isolated satisfies

$$\mathbf{P}(L_i \mid X_i) = \bigl(1 - a(X_i)\bigr)^{n-1}.$$

Take expectation with respect to X_i to get rid of the conditioning and obtain

$$P_0 = \mathbf{P}(L_i) = \mathbf{E}\bigl\{\mathbf{P}(L_i \mid X_i)\bigr\} = \mathbf{E}\bigl\{\bigl(1 - a(X_i)\bigr)^{n-1}\bigr\}. \tag{1}$$

Some preliminary spadework helps put subsequent expressions into an analytically amenable framework. Observe that $A(X_i) \leq \pi r^2$ whence $a(X_i) \leq r^2$ with equality in the interior of the unit circle. As $\log(1-x) = -x + \mathcal{O}(x^2)$ and $e^x = 1 + \mathcal{O}(x)$ as $x \to 0$, we have

$$\begin{aligned}\bigl(1 - a(X_i)\bigr)^{n-1} &= e^{n \log(1 - a(X_i))}\bigl(1 - a(X_i)\bigr)^{-1} \\ &= e^{-na(X_i) + \mathcal{O}(nr^4)}\bigl[1 + \mathcal{O}(r^2)\bigr] = e^{-na(X_i)}\bigl[1 + \mathcal{O}\bigl(\tfrac{\log^2 n}{n}\bigr)\bigr]\end{aligned}$$

where the order bound is uniform in X_i. It follows that

$$P_0 = \left[1 + \mathcal{O}\left(\tfrac{\log^2 n}{n}\right)\right] \mathbf{E}\left(e^{-na(X_i)}\right) \qquad (n \to \infty) \tag{2}$$

and it suffices hence to evaluate

$$I(n) \triangleq \mathbf{E}\left(e^{-na(X_i)}\right) = \frac{1}{\pi}\int_{\mathbb{S}} e^{-na(x_i)}\, dx_i \tag{2$'$}$$

as we recall that X_i is drawn by sampling from the uniform distribution on the unit disc $\mathbb{S}$.

The observation that $A(X_i)$, hence also $a(X_i)$, depends only on the distance $|X_i|$ of vertex i from the origin helps further simplify the expression. The distance $|X_i|$ of vertex i from the field centre is a random variable with distribution function $\mathbf{P}\{|X_i| \leq \rho\} = \pi\rho^2/\pi = \rho^2$ for $0 \leq \rho \leq 1$. It follows that $|X_i|$ has density 2ρ $(0 \leq \rho \leq 1)$. In a slight abuse of notation, write $a(X_i) = a_{|X_i|}$ to emphasise the dependence of a on the length of X_i alone. We hence have

$$I(n) = \mathbf{E}\left(e^{-na_{|X_i|}}\right) = \int_0^1 2\rho e^{-na_\rho}\, d\rho.$$

It is expedient to partition the range of the integral and write $I(n) = I_{\text{interior}} + I_{\text{boundary}}$ where I_{interior} is the contribution to the integral from the interior of the unit circle $0 \leq \rho \leq 1-r$ and I_{boundary} is the contribution to the integral from the annulus $1 - r < \rho \leq 1$ at the boundary. We evaluate these contributions in turn.

The interior contribution. When $0 \leq \rho \leq 1 - r$, the point X_i is in the interior of the unit circle and the circle of radius r centred at X_i is contained wholly within the unit disc $\mathbb{S}$. It follows that $a_\rho = \pi r^2/\pi = r^2$ whence

$$I_{\text{interior}} = \int_0^{1-r} 2\rho e^{-na_\rho}\, d\rho = e^{-nr^2}(1-r)^2 = e^{-nr^2}\left[1 + \mathcal{O}\left(\tfrac{\log n}{\sqrt{n}}\right)\right] \sim \tfrac{e^{-c}}{n}$$

asymptotically as $n \to \infty$.

The boundary contribution. When $1 - r < \rho \leq 1$, the point X_i lies in an asymptotically vanishing annulus of width r at the boundary of the unit circle. The region of visibility is now the lens formed by the intersection of the disc of radius r centred at X_i with the unit disc $\mathbb{S}$ and it is clear that a_ρ decreases monotonically from r^2 to a value close to $r^2/2$ as ρ increases from $1 - r$ to 1.

The situation at the boundary is delicate. On the face of it, the probability that a vertex lands in the boundary is small so that other things being equal the boundary should contribute very little to the probability of isolation. However, in the boundary annulus the area of a vertex's region of visibility is about one-half of its area in the interior, in consequence fewer vertices are adjacent to it, and hence the chances of isolation increase as isolation now requires the extinction of fewer vertices than in the interior. The contribution to the probability of isolation from the boundary is dynamically balanced between these two opposing effects.

There are two geometrically distinct regimes in the boundary depending on whether X_i and the origin lie on the same side of the chord joining the intersections of the two circles or whether X_i and the origin are separated by the chord. Two typical situations are illustrated in Figure 1 where the small circle of radius r has been grossly exaggerated in size to make the details visible. Here X_i is located at point Q.

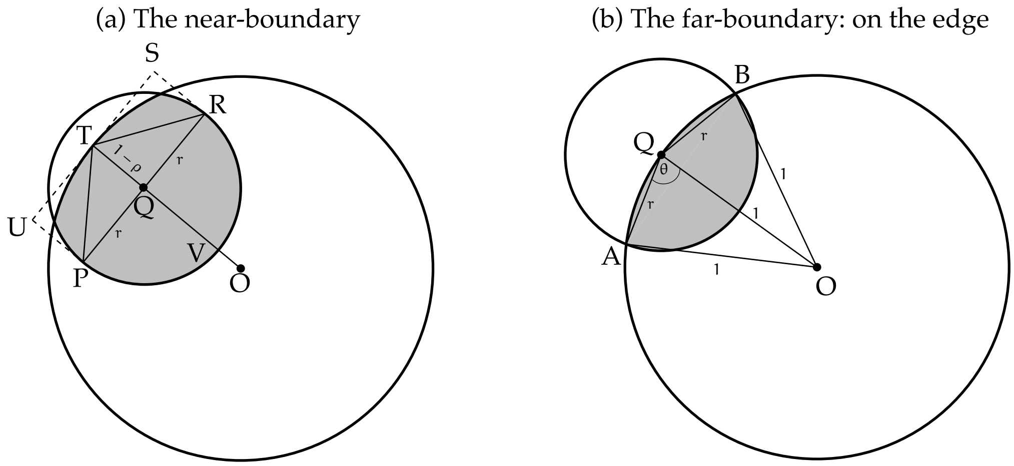

Figure 1: The boundary regime and the visibility lens (shown shaded) with X_i located at point Q. (a) The vertex is located in the near-boundary $1 - r < \rho \leq \sqrt{1-r^2}$. (b) The vertex is located at the very edge of the far-boundary $\sqrt{1-r^2} < \rho \leq 1$.

The contribution of the boundary to the probability integral may hence be further decomposed into $I_{\text{boundary}} = I^{(1)}_{\text{boundary}} + I^{(2)}_{\text{boundary}}$ where

$$I^{(1)}_{\text{boundary}} = \int_{1-r}^{\sqrt{1-r^2}} 2\rho e^{-na_\rho}\, d\rho \quad \text{and} \quad I^{(2)}_{\text{boundary}} = \int_{\sqrt{1-r^2}}^{1} 2\rho e^{-na_\rho}\, d\rho$$

are the near- and far-boundary contributions, respectively. We consider these in turn.

The near-boundary: $1 - r < \rho \leq \sqrt{1-r^2}$. This is the critical region where the boundary effects are most pronounced. Consider the typical situation illustrated in Figure 1(a). While the calculation of lens area is now routine, if tedious, we can finesse some of these calculations by bounding the dominant contribution though some care needs to be exercised as the asymptotics are delicate.

The conjoined area encompassed by the semicircle PQRVP and the isosceles triangle PTR is wholly contained within the lens. It follows that the area of the lens, A_ρ, is bounded from below by the sum of these conjoined areas. An easy calculation now shows that $A_\rho \geq \frac{1}{2}\pi r^2 + r(1-\rho)$ and, consequently, the probability $a_\rho = \frac{1}{\pi} A_\rho$ that a randomly placed vertex will land in the lenticular area at the near-boundary is bounded by

$$a_\rho \geq \tfrac{1}{2} r^2 + \tfrac{1}{\pi} r(1-\rho). \tag{3}$$

The lower bound for a_ρ may be immediately deployed in bounding the contribution of the near-boundary to the probability integral via

$$\begin{aligned} I^{(1)}_{\text{boundary}} &\overset{\text{(i)}}{\leq} e^{-nr^2/2} \int_{1-r}^{\sqrt{1-r^2}} 2\rho e^{-nr(1-\rho)/\pi}\, d\rho \overset{\text{(ii)}}{\leq} 2e^{-nr^2/2} \int_{1-r}^{\sqrt{1-r^2}} e^{-nr(1-\rho)/\pi}\, d\rho \\ &\overset{\text{(iii)}}{=} 2e^{-nr^2/2} \int_{1-\sqrt{1-r^2}}^{r} e^{-nru/\pi}\, du \leq 2e^{-nr^2/2} \int_0^\infty e^{-nru/\pi}\, du \\ &= \frac{2\pi e^{-nr^2/2}}{nr} \sim \frac{2\pi r\sqrt{\lambda}}{\sqrt{n}\,\log n} \qquad (n \to \infty) \end{aligned}$$

where (i) follows from the lower bound (3) for a_ρ, (ii) holds because ρ is trivially bounded above by $\sqrt{1-r^2} \leq 1$ inside the integral, and (iii) follows from a change of variable of integration from ρ to $u = 1 - \rho$. We hence obtain the asymptotic estimate

$$I^{(1)}_{\text{boundary}} = \mathcal{O}\left(\frac{r}{\sqrt{n}\,\log n}\right) = \mathcal{O}\left(\frac{1}{n\sqrt{\log n}}\right) = \mathfrak{o}\left(\frac{1}{n}\right) \qquad (n \to \infty).$$

It follows that the contribution from the near-boundary to the isolation probability is dominated asymptotically by the interior contribution $I_{\text{interior}} \sim \frac{\lambda}{n}$.

The far-boundary: $\sqrt{1-r^2} < \rho \leq 1$. Recall that a_ρ decreases monotonically as ρ increases so that $a_\rho \geq a_1 = \frac{1}{\pi} A_1$ where A_1 is the area of the visibility lens when X_i is situate on the circumference of the unit circle. The situation is illustrated in Figure 1(b) where the small circle of radius r has again been grossly exaggerated to make visible the geometric detail.

With notation as in Figure 1(b), $\theta = \angle AQO$ is the half-angle of the sector enclosed by the chords AQ and BQ and the interior rim of the circle of radius r centred at Q on the circumference of the unit circle. Write A_{sector} for the area of this sector. It follows that $A_{\text{sector}} = \frac{1}{2}(2\theta)r^2 = \theta r^2$. As AOQ is an isosceles triangle with side 1 and base r, we have $\theta = \arccos\left(\frac{r}{2}\right) = \frac{\pi}{2} - \arcsin\left(\frac{r}{2}\right)$. As $\arcsin x \leq 2x$ for all sufficiently small x, we obtain

$$a_1 = \tfrac{1}{\pi} A_1 \geq \tfrac{1}{\pi} A_{\text{sector}} = \tfrac{1}{2} r^2 - \tfrac{1}{\pi} r^2 \arcsin\left(\tfrac{r}{2}\right) \geq \tfrac{1}{2} r^2 - \tfrac{1}{\pi} r^3 \tag{4}$$

for all sufficiently large n (recall $r = r_n \to 0$). The far-boundary contribution to the probability integral may hence be bounded by

$$\begin{aligned} I^{(2)}_{\text{boundary}} &= \int_{\sqrt{1-r^2}}^{1} 2\rho e^{-n a_\rho}\, d\rho \leq \int_{\sqrt{1-r^2}}^{1} 2\rho e^{-n a_1}\, d\rho = r^2 e^{-n a_1} \\ &\leq r^2 e^{-\frac{1}{2} n r^2 + \frac{1}{\pi} n r^3} = r^2 e^{-n r^2/2}\left[1 + \mathfrak{o}\left(\tfrac{\log^2 n}{n}\right)\right] \sim r^2 \sqrt{\tfrac{\lambda}{n}} \qquad (n \to \infty) \end{aligned}$$

leading to the asymptotic estimate

$$I^{(2)}_{\text{boundary}} = \mathcal{O}\left(\frac{r^2}{\sqrt{n}}\right) = \mathcal{O}\left(\frac{\log n}{n^{3/2}}\right) = \mathfrak{o}\left(\frac{1}{n}\right).$$

From the near- and far-boundary estimates we obtain the entire boundary contribution to the probability integral as

$$I_{\text{boundary}} = I^{(1)}_{\text{boundary}} + I^{(2)}_{\text{boundary}} = \mathfrak{o}\left(n^{-1}\right).$$

Isolation probability. In accordance with naïve expectation, the boundary contributes an asymptotically negligible amount to the isolation probability for the given rate of variation of $r = r_n$ (though the conclusion is false for rates of variation $r = r_n$ larger than prescribed; if r_n has an asymptotic order higher than $\log n/\sqrt{n}$ then the contribution from the boundary is at least comparable to the interior contribution and the boundary comes into play). Combining the interior and boundary estimates for the integral we obtain

$$I(n) = I_{\text{interior}} + I_{\text{boundary}} = \tfrac{\lambda}{n}\left(1 + \mathfrak{o}(1)\right) + \mathfrak{o}\left(\tfrac{1}{n}\right).$$

From (2, 2′) it follows that, as advertised, the probability that any given vertex is isolated satisfies $P_0 \sim I(n) \sim \frac{\lambda}{n}$ as $n \to \infty$. In particular, the expected number of isolated vertices satisfies $\mathbf{E}(N) \to \lambda$ as $n \to \infty$. This concludes the proof of Lemma 1. We now proceed to refine this crude estimate by proving Lemma 2.

2.2 Conjunctions of Vertex Isolations

The new features that are encountered in moving from one vertex to several vertices are statistical dependencies that arise due to overlaps of visibility regions as well as slightly more complicated boundary interactions. We will keep the burgeoning complexity under bounds by reducing considerations of dependencies that accrue for groups of vertices to repeated considerations of vertex pairs.

Let us pause here to introduce some new notation. Write $\mathbb{D}(z, x)$ for the closed disc of radius z centred at the point x in the plane. In this notation, $\mathbb{S} = \mathbb{D}(1, 0)$ is the unit disc centred at the origin. We also write $\mathbb{V}(x) = \mathbb{D}(r, x) \cap \mathbb{S}$ for the region of visibility of a vertex located at point x. For brevity we say simply that $\mathbb{V}(x)$ is the region of visibility of the point x. We also define the *overlap region* of a vertex located at point x by $\mathbb{O}(x) = \mathbb{D}(2r, x) \cap \mathbb{S}$. The motivation arises from the observation that if $x_2 \in \mathbb{O}(x_1)$ then $\mathbb{V}(x_1) \cap \mathbb{V}(x_2) \neq \emptyset$ and the regions of visibility of the two points overlap. Overlap regions will be important in the characterisation of dependencies. Finally, in a slight but hopefully natural modification of an earlier notation, write $A_k(x_1, \dots, x_k) = \int_{\mathbb{V}(x_1)\cup\cdots\cup\mathbb{V}(x_k)} dx$ for the area of the conjoined visibility region $\mathbb{V}(x_1)\cup\cdots\cup\mathbb{V}(x_k)$ of the points $x_1, \dots, x_k$. Specialising to $k = 1$ and 2, $A_1(x)$ is the area of the visibility region of point x (replacing the earlier notations $A(x)$ and $A_{|x|}$) and $A_2(x_1, x_2)$ is the area of $\mathbb{V}(x_1) \cup \mathbb{V}(x_2)$. Finally write $a_k(x_1, \dots, x_k) = \frac{1}{\pi} A_k(x_1, \dots, x_k)$ for the probability that a randomly selected point in the unit disc $\mathbb{S}$ falls in the region of visibility of at least one of the k points $x_1, \dots, x_k$. In particular, $a_1(x)$ is the probability that a randomly selected point falls in $\mathbb{V}(x)$ (replacing the earlier notations $a(x)$ and $a_{|x|}$) and $a_2(x_1, x_2)$ is the probability that a randomly selected point falls in $\mathbb{V}(x_1) \cup \mathbb{V}(x_2)$.

Fix any positive integer $k \geq 2$ and let $i_1, \dots, i_k$ be any collection of k distinct vertices. We now proceed to estimate the probability of the event $L_{i_1,\dots,i_k} = L_{i_1} \cap \cdots \cap L_{i_k}$ that each of the k vertices $i_1, \dots, i_k$ is isolated. By exchangeability, $\mathbf{P}(L_{i_1,\dots,i_k}) = \mathbf{P}(L_{1,\dots,k})$ and we may without loss of generality focus on vertices 1 through k.

Given the random k-tuple of points $(X_1, \dots, X_k)$, the conditional probability that vertices 1 through k are isolated is a random variable $\mathbf{P}(L_{1,\dots,k} \mid X_1, \dots, X_k)$ that takes the value $\mathbf{P}(L_{1,\dots,k} \mid X_1 = x_1, \dots, X_k = x_k)$ at sample points in the space characterised by the joint occurrence of the events $\{X_1 = x_1\}, \dots, \{X_k = x_k\}$. As $X_1, \dots, X_n$ are drawn by independent sampling from the uniform distribution on the unit disc we have

$$\mathbf{P}(L_{1,\dots,k}) = \mathbf{E}\big[\mathbf{P}(L_{1,\dots,k} \mid X_1, \dots, X_k)\big] = \frac{1}{\pi^k} \int\cdots\int_{\mathbb{S}^k} \mathbf{P}(L_{1,\dots,k} \mid X_1 = x_1, \dots, X_k = x_k)\, dx_k \cdots dx_1. \quad (5)$$

It will again be expedient to partition the range of the integral. But first, one more definition.

The overlap graph. With every choice of points $x_1, \dots, x_k$ in $\mathbb{S}$ we may associate the *overlap graph* $\mathcal{G}_k(x_1, \dots, x_k)$ on the set of vertices $\{1, \dots, k\}$ whose edges are the vertex pairs (i, j) for which $|x_i - x_j| \leq 2r$, that is to say, $\mathbb{V}(x_i) \cap \mathbb{V}(x_j) \neq \emptyset$. Each overlap graph $\mathcal{G}_k(x_1, \dots, x_k)$ may be partitioned into $c = c(x_1, \dots, x_k)$ *components* $\{\mathcal{G}_k^{(1)}, \dots, \mathcal{G}_k^{(c)}\}$ where $1 \leq c \leq k$ and each component $\mathcal{G}_k^{(l)}$ is a connected sub-graph with no out-going edges. (More formally, if $\mathcal{G}_k^{(l)}$ and $\mathcal{G}_k^{(m)}$ are two distinct components then $\mathcal{G}_k$ exhibits no edges (i, j) where i is a vertex of $\mathcal{G}_k^{(l)}$ and j is a vertex of $\mathcal{G}_k^{(m)}$.)

Now any graph $\mathcal{F}$ on the vertices $\{1, \dots, k\}$ induces an equivalence class $\mathbb{C}(\mathcal{F})$ of k-tuples of points $(x_1, \dots, x_k)$ in $\mathbb{S}^k$ whose overlap graph $\mathcal{G}_k(x_1, \dots, x_k)$ coincides with $\mathcal{F}$.

As each k-tuple of points $(x_1, \dots, x_k)$ in $\mathbb{S}^k$ belongs to one and only one equivalence class $\mathbb{C}(\mathcal{F})$, as $\mathcal{F}$ varies over all graphs on k vertices the sets $\mathbb{C}(\mathcal{F})$ partition $\mathbb{S}^k$ into $2^{\binom{k}{2}}$ non-overlapping regions. We may hence partition the integral on the right-hand side of (5) and write

$$\mathbf{P}(L_{1,\dots,k}) = \sum_{\mathcal{F}} \mathbf{E}\bigl[1_{\mathbb{C}(\mathcal{F})}(X_1, \dots, X_k)\, \mathbf{P}(L_{1,\dots,k} \mid X_1, \dots, X_k)\bigr] \tag{6}$$

where the sum ranges over all $2^{\binom{k}{2}}$ graphs $\mathcal{F}$ on k vertices and we use the generic indicator function notation

$$1_{\mathcal{A}}(\zeta) = \begin{cases} 1 & \text{if } \zeta \in \mathcal{A}, \\ 0 & \text{if } \zeta \notin \mathcal{A}. \end{cases}$$

Of particular interest is the graph $\mathcal{F}_*$ obtained when the vertices $\{1, \dots, k\}$ form an independent set, that is to say, the graph has no edges. Clearly $\mathcal{F}_*$ is the unique graph with $c = k$ components, each component consisting of course of a singleton vertex. The corresponding equivalence class $\mathbb{C}(\mathcal{F}_*)$ consists of k-tuples of points $(x_1, \dots, x_k)$ where the x_i have mutually non-overlapping regions of visibility, that is to say, $\mathbb{V}(x_i) \cap \mathbb{V}(x_j) = \emptyset$ if $i \neq j$. The remaining graphs $\mathcal{F} \neq \mathcal{F}_*$ contain $1 \leq c \leq k - 1$ components and correspond to overlap graphs induced by points $x_1, \dots, x_k$ for which at least two regions of visibility overlap. We separate these cases for consideration and write

$$\begin{aligned} J_{\text{non-overlap}} &= \mathbf{E}\bigl[1_{\mathbb{C}(\mathcal{F}_*)}(X_1, \dots, X_k)\, \mathbf{P}(L_{1,\dots,k} \mid X_1, \dots, X_k)\bigr], \\ J_{\text{overlap}} &= \sum_{\mathcal{F} \neq \mathcal{F}_*} \mathbf{E}\bigl[1_{\mathbb{C}(\mathcal{F})}(X_1, \dots, X_k)\, \mathbf{P}(L_{1,\dots,k} \mid X_1, \dots, X_k)\bigr]. \end{aligned} \tag{7}$$

Then $\mathbf{P}(L_{1,\dots,k}) = J_{\text{non-overlap}} + J_{\text{overlap}}$ and we proceed to evaluate these contributions in turn.

Non-overlapping regions of visibility; independent sets and the overlap graph $\mathcal{F}_*$. The set of points $(x_1, \dots, x_k)$ forming the equivalence class $\mathbb{C}(\mathcal{F}_*)$ may be systematically specified by the following recursive procedure.

BASE: The point x_1 is allowed to range over all points in the unit disc $\mathbb{S}$.

RECURRENCE: As j ranges from 2 to k, for every selection of $x_1, \dots, x_{j-1}$, the point x_j is allowed to range over the set of points $\mathbb{S} \setminus \bigcup_{i=1}^{j-1} \mathbb{O}(x_i)$ consisting of the unit disc with the overlap regions of the points $x_1, \dots, x_{j-1}$ excised.

It is clear from the recursive construction that the points $x_1, \dots, x_k$ specified thus have non-overlapping regions of visibility and, furthermore, a little thought shows that indeed this process covers all the points in $\mathbb{C}(\mathcal{F}_*)$. Now suppose $(x_1, \dots, x_k) \in \mathbb{C}(\mathcal{F}_*)$ so that the x_i have mutually non-overlapping regions of visibility. Then $a_k(x_1, \dots, x_k) = a_1(x_1) + \cdots + a_1(x_k)$ where each $a_1(x_i) \leq r^2$ (with equality iff x_i is an interior point). Following the line of argument leading up to (1, 2) we then obtain

$$\begin{aligned} \mathbf{P}(L_{1,\dots,k} \mid X_1 = x_1, \dots, X_k = x_k) &= \bigl(1 - a_k(x_1, \dots, x_k)\bigr)^{n-k} \\ &= \bigl[1 + \mathcal{O}\bigl(\tfrac{\log^2 n}{n}\bigr)\bigr] e^{-n a_k(x_1, \dots, x_k)} = \bigl[1 + \mathcal{O}\bigl(\tfrac{\log^2 n}{n}\bigr)\bigr] e^{-n a_1(x_1)} \cdots e^{-n a_1(x_k)} \end{aligned}$$

where the order term is uniform over all choices of $x_1, \dots, x_k$. By systematically conditioning on X_1 first, then X_2, and so on, and finally conditioning on X_k, we obtain

$$\begin{aligned} J_{\text{non-overlap}} &= \left[1 + \mathcal{O}\left(\tfrac{\log^2 n}{n}\right)\right] \mathbf{E}\left(1_{\mathbb{C}(\mathcal{F}_*)}(X_1, \dots, X_k) e^{-na_1(X_1)} \cdots e^{-na_1(X_k)}\right) \\ &= \left[1 + \mathcal{O}\left(\tfrac{\log^2 n}{n}\right)\right] \frac{1}{\pi} \int_{\mathbb{S}} dx_1 e^{-na_1(x_1)} \frac{1}{\pi} \int_{\mathbb{S} \setminus \mathbb{O}(x_1)} dx_2 e^{-na_1(x_2)} \\ &\qquad \cdots \frac{1}{\pi} \int_{\mathbb{S} \setminus \bigcup_{i=1}^{k-1} \mathbb{O}(x_i)} dx_k e^{-na_1(x_k)}. \end{aligned}$$

A typical nested integral on the right-hand side is of the form

$$\frac{1}{\pi} \int_{\mathbb{S} \setminus \bigcup_{i=1}^{j-1} \mathbb{O}(x_i)} dx_j e^{-na_1(x_j)} = \frac{1}{\pi} \int_{\mathbb{S}} dx_j e^{-na_1(x_j)} - \frac{1}{\pi} \int_{\bigcup_{i=1}^{j-1} \mathbb{O}(x_i)} dx_j e^{-na_1(x_j)}.$$

The first integral on the right is, up to a multiplicative factor of $1 + \mathfrak{o}(1)$, just the probability that vertex j is isolated (see (2, 2′)) whence

$$\frac{1}{\pi} \int_{\mathbb{S}} dx_j e^{-na_1(x_j)} \sim P_0 \sim \frac{e^{-c}}{n} \qquad (n \to \infty)$$

from our considerations for a single vertex. As $a_1(x_j) \geq \frac{1}{2}r^2 - \frac{1}{\pi}r^3$ for large enough n (see (4)), the second of the two integrals on the right is bounded by

$$\begin{aligned} &\frac{1}{\pi} \int_{\bigcup_{i=1}^{j-1} \mathbb{O}(x_i)} dx_j e^{-na_1(x_j)} \\ &\qquad \leq e^{-\frac{1}{2}nr^2 + \frac{1}{\pi}nr^3} \left(\frac{1}{\pi} \int_{\bigcup_{i=1}^{j-1} \mathbb{O}(x_i)} dx_j \right) \leq (j-1)4r^2 e^{-\frac{1}{2}nr^2 + \frac{1}{\pi}nr^3} \\ &\qquad\quad = (j-1)4r^2 \sqrt{\tfrac{e^{-c}}{n}} \left(1 + \mathfrak{o}(1)\right)\left(1 + \mathcal{O}(r \log n)\right) = \mathcal{O}\left(\frac{r^2}{\sqrt{n}}\right) = \mathcal{O}\left(\frac{\log n}{n^{3/2}}\right) \end{aligned}$$

where the order term is uniform with respect to the x_i. Consequently,

$$\frac{1}{\pi} \int_{\mathbb{S} \setminus \bigcup_{i=1}^{j-1} \mathbb{O}(x_i)} dx_j e^{-na_1(x_j)} = \frac{e^{-c}}{n} \left[1 + \mathfrak{o}(1) + \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right)\right] \tag{8}$$

for each j and we obtain the estimate

$$J_{\text{non-overlap}} = \frac{(e^{-c})^k}{k!} \left[1 + \mathfrak{o}(1) + \mathcal{O}\left(\frac{\log n}{\sqrt{n}}\right) + \mathcal{O}\left(\frac{\log^2 n}{n}\right)\right] \sim \frac{(e^{-c})^k}{n^k} \tag{9}$$

asymptotically as $n \to \infty$.

Overlapping regions of visibility preliminaries; connected overlap graphs. We may systematically group graphs $\mathcal{F}$ on k vertices according to (i) the number of components $c = c(\mathcal{F})$ of the graph and (ii) the sizes of the components $k_1, \dots, k_c$. Define

$$J_{k_1,\dots,k_c} \triangleq \sum_{\mathcal{F}_{k_1,\dots,k_c}} \mathbf{E}\left[1_{\mathbb{C}(\mathcal{F}_{k_1,\dots,k_c})}(X_1, \dots, X_k)\, \mathbf{P}\left(L_{1,\dots,k} \mid X_1, \dots, X_k\right)\right] \tag{10}$$

where, in an obvious notation, the sum on the right ranges over all graphs $\mathcal{F}_{k_1,\dots,k_c}$ with c components of sizes $k_1, \dots, k_c$.[1] As the family of graphs $\mathcal{F} \neq \mathcal{F}_*$ have a number of components varying between $1 \leq c \leq k-1$ we accordingly obtain

$$J_{\text{overlap}} = \sum_{c=1}^{k-1} \sum_{\substack{k_1 \geq \cdots \geq k_c \geq 1 \\ k_1 + \cdots + k_c = k}} J_{k_1,\dots,k_c}. \tag{10$'$}$$

Begin by considering the cases when the overlap graphs consist of a single component, $c = 1$. Of course, in these cases, the graphs themselves are connected and the single component has size k. As we shall see, this case informs all the others.

Specialising to the case $c = 1$, (10) becomes

$$J_k = \sum_{\mathcal{F}_k} \mathbf{E}\big[\mathbf{1}_{\mathbb{C}(\mathcal{F}_k)}(X_1, \dots, X_k)\, \mathbf{P}\big(L_{1,\dots,k} \mid X_1, \dots, X_k\big)\big] \tag{11}$$

where $\mathcal{F}_k$ now ranges over all *connected overlap graphs* on k vertices. Consider any connected overlap graph $\mathcal{F}_k$ and the associated equivalence class of k-tuples $\mathbb{C}(\mathcal{F}_k)$. As to any vertex i there is at least one vertex j such that (i, j) is an edge of $\mathcal{F}_k$, it follows that for any k-tuple $(x_1, \dots, x_k) \in \mathbb{C}(\mathcal{F}_k)$ to every point x_i there is at least one other point x_j such that the regions of visibility of x_i and x_j overlap, that is, $|x_i - x_j| \leq 2r$. If the proximity of x_i and x_j is such that indeed $|x_i - x_j| \leq r$ then vertices i and j will be adjacent to each other (i.e., communicating) in the original graph. It follows that conditioned on $\{X_1 = x_1, \dots, X_k = x_k\}$, the occurrence of the event $L_{1,\dots,k}$ will require not only that any of the vertices $k+1, \dots, n$ that fall in the conjoined visibility region $\mathbb{V}(x_1) \cup \cdots \cup \mathbb{V}(x_k)$ be extinguished but may also enjoin the extinction of two or more of the vertices 1 through k themselves. But in all cases we may bound

$$\begin{aligned}\mathbf{P}\big(L_{1,\dots,k} \mid X_1 = x_1, \dots, X_k = x_k\big) &\overset{(v)}{\leq} \big(1 - a_k(x_1, \dots, x_k)\big)^{n-k} \\ &\overset{(vi)}{=} \big[1 + \mathcal{O}\big(\tfrac{\log^2 n}{n}\big)\big] e^{-n a_k(x_1,\dots,x_k)}\end{aligned} \tag{12}$$

where in (v) the inequality takes cognizance of the fact that some of the vertices from 1 through k may be mutually adjacent (communicating) and (vi) follows from the by-now standard process (1, 2). It follows that

$$\begin{aligned}J_k &\leq \big[1 + \mathcal{O}\big(\tfrac{\log^2 n}{n}\big)\big] \sum_{\mathcal{F}_k} \mathbf{E}\big[\mathbf{1}_{\mathbb{C}(\mathcal{F}_k)}(X_1, \dots, X_k) e^{-n a_k(X_1,\dots,X_k)}\big] \\ &= \big[1 + \mathcal{O}\big(\tfrac{\log^2 n}{n}\big)\big]\, \mathbf{E}\Big[\sum_{\mathcal{F}_k} \mathbf{1}_{\mathbb{C}(\mathcal{F}_k)}(X_1, \dots, X_k) e^{-n a_k(X_1,\dots,X_k)}\Big].\end{aligned}$$

We now begin the process of consolidating all connected overlap graphs under one rubric. Observe that $\sum_{\mathcal{F}_k} \mathbf{1}_{\mathbb{C}(\mathcal{F}_k)}(X_1, \dots, X_k)$ is itself the indicator for all k-tuples $(X_1, \dots, X_k)$ for which the corresponding overlap graph is connected. Accordingly, write $\mathbb{C}_k$ for the subset of k-tuples $(x_1, \dots, x_k)$ in $\mathbb{S}^k$ for which $\mathcal{G}_k(x_1, \dots, x_k)$ is connected. It then follows that $\sum_{\mathcal{F}_k} \mathbf{1}_{\mathbb{C}(\mathcal{F}_k)}(X_1, \dots, X_k) = \mathbf{1}_{\mathbb{C}_k}(X_1, \dots, X_k)$ whence

$$J_k \leq \big[1 + \mathcal{O}\big(\tfrac{\log^2 n}{n}\big)\big]\, \mathbf{E}\big[\mathbf{1}_{\mathbb{C}_k}(X_1, \dots, X_k) e^{-n a_k(X_1,\dots,X_k)}\big]. \tag{13}$$

[1] In this new notation the graph $\mathcal{F}_*$ may be identified with the unique graph $\mathcal{F}_{1,\dots,1}$ but we will not be pedagogically fussy here.

We may range over the k-tuples $(x_1, \ldots, x_k)$ in $\mathbb{C}_k$ by allowing x_1 to range over the unit disc $\mathbb{S}$ and, for each x_1, allowing the $(k-1)$-tuple $(x_2, \ldots, x_k)$ to range over the subset $\mathbb{C}_{k-1}(x_1)$ of $\mathbb{S}^{k-1}$ for which the overlap graph $\mathcal{G}_k(x_1, \ldots, x_k)$ remains connected. Now connectivity enjoins that the maximum distance from x_1 to any of the remaining points $x_2, \ldots, x_k$ can be no larger than $2(k-1)r$. It follows that the $\mathbb{C}_{k-1}(x_1)$ is contained in the set of $(k-1)$-tuples $(x_2, \ldots, x_k)$ for which $\max_{2\le i\le k} |x_i - x_1| \le (2k-2)r$.

These considerations suggest that we introduce a new random variable Z defined by $Z = \max_{2\le i\le k} |X_i - X_1|$ and representing the girth of the overlap graph as viewed from X_1. The game plan is to exploit the fact that large girth connected graphs have relatively large footprints while small girth connected graphs are relatively unlikely to occur while retaining bounded footprints. The latter effect is similar in flavour to the balancing act we encountered at the boundary in the analysis for a single vertex. To proceed, we now have

$$1_{\mathbb{C}_k}(X_1, \ldots, X_k) = 1_{\mathbb{S}}(X_1) 1_{\mathbb{C}_{k-1}(X_1)}(X_2, \ldots, X_k) \le 1_{\mathbb{S}}(X_1) 1_{[0,(2k-2)r]}(Z)$$

so that the expectation on the right-hand side of (13) may be bounded by

$$\begin{aligned} \mathbf{E}\bigl[1_{\mathbb{C}_k}(X_1, \ldots, X_k) e^{-n a_k(X_1,\ldots,X_k)}\bigr] &\\ = \mathbf{E}\bigl[1_{\mathbb{S}}(X_1)\, \mathbf{E}\bigl(1_{\mathbb{C}_{k-1}(X_1)}(X_2, \ldots, X_k) e^{-n a_k(X_1,\ldots,X_k)} \mid X_1\bigr)\bigr] &\\ \le \mathbf{E}\bigl[1_{\mathbb{S}}(X_1)\, \mathbf{E}\bigl(1_{[0,(2k-2)r]}(Z) e^{-n a_k(X_1,\ldots,X_k)} \mid X_1\bigr)\bigr]. \end{aligned}$$

Now set $\jmath^* = \arg\max_{2\le i\le k} |X_i - X_1|$ and let $X_{\jmath^*}$ be any point at maximal distance Z from X_1. It is clear that the conjoined area of visibility of the points X_1, …, X_k certainly includes the conjoined areas of visibility of the points X_1 and $X_{\jmath^*}$ alone so that $a_k(X_1, \ldots, X_n) \ge a_2(X_1, X_{\jmath^*})$. Substituting on the right-hand side above we may continue to bound the expectation to obtain

$$\begin{aligned} \mathbf{E}\bigl[1_{\mathbb{C}_k}(X_1, \ldots, X_k) e^{-n a_k(X_1,\ldots,X_k)}\bigr] &\\ \le \mathbf{E}\bigl[1_{\mathbb{S}}(X_1)\, \mathbf{E}\bigl(1_{[0,(2k-2)r]}(Z) e^{-n a_2(X_1,X_{\jmath^*})} \mid X_1\bigr)\bigr] \triangleq K(n). \end{aligned} \tag{14}$$

The messy problem of estimating $a_k(X_1, \ldots, X_k)$ for a random k-tuple is now reduced to that of estimating $a_2(X_1, X_{\jmath^*})$ for a pair of vertices, albeit with a special statistical structure.

Following the mode of analysis for a single vertex we now partition $\mathbb{S}$ into an interior region $\mathbb{I} = \{ x : |x| \le 1 - (2k-1)r \}$ and an (expanded) boundary annulus $\mathbb{S} \setminus \mathbb{I} = \{ x : 1 - (2k-1)r < |x| \le 1 \}$. Then $1_{\mathbb{S}}(X_1) = 1_{\mathbb{I}}(X_1) + 1_{\mathbb{S}\setminus\mathbb{I}}(X_1)$ and we may partition the bound $K(n)$ on the right-hand side of (14) further into the sum of the two contributions, $K(n) = K_{\text{interior}} + K_{\text{boundary}}$, where

$$\begin{aligned} K_{\text{interior}} &= \mathbf{E}\bigl[1_{\mathbb{I}}(X_1)\, \mathbf{E}\bigl(1_{[0,(2k-2)r]}(Z) e^{-n a_2(X_1,X_{\jmath^*})} \mid X_1\bigr)\bigr] \\ K_{\text{boundary}} &= \mathbf{E}\bigl[1_{\mathbb{S}\setminus\mathbb{I}}(X_1)\, \mathbf{E}\bigl(1_{[0,(2k-2)r]}(Z) e^{-n a_2(X_1,X_{\jmath^*})} \mid X_1\bigr)\bigr]. \end{aligned} \tag{15}$$

We consider these in turn.

The interior overlap contribution: If $X_1 \in \mathbb{I}$ then the discs of radius r centred at X_1 and $X_{\jmath^*}$ are both wholly contained within the unit disc. As illustrated in Figure 2 there are now two possible cases depending on whether the regions of visibility of X_1 and $X_{\jmath^*}$ intersect.

CASE 1: When $0 \le Z \le 2r$ the regions of visibility of X_1 and $X_{\jmath^*}$ intersect as shown in Figures 2(a) and (b). Let A_{lens} be the area of the shaded lenticular overlap region shown in the figures. Then $A_2(X_1, X_{\jmath^*}) = A_1(X_1) + A_2(X_2) - A_{\text{lens}} = 2\pi r^2 - A_{\text{lens}}$. Now consider the sector enclosed by the radial lines QP and QR on the one hand and the arc PSR on the

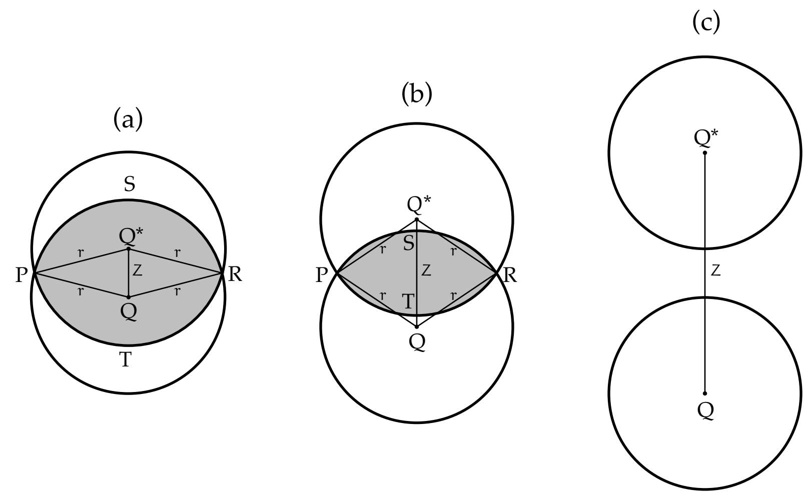

Figure 2: Point X_1 in the interior $\mathbb{I}$ of the unit circle is located at point Q while point X_{j^*} is located at point Q^*. (a, b) Two cases where the regions of visibility of X_1 and X_{j^*} overlap. The overlap lens is shown shaded. (c) The regions of visibility of X_1 and X_{j^*} are disjoint.

other. This sector has half-angle given by $\angle PQQ^* = \arccos\left(\frac{Z}{2r}\right) = \frac{\pi}{2} - \arcsin\left(\frac{Z}{2r}\right)$. It follows that the sector has area $A_{\text{sector}} = r^2\angle PQQ^* = \frac{\pi}{2}r^2 - r^2\arcsin\left(\frac{Z}{2r}\right)$. By symmetry, the sector enclosed by the radial lines Q^*P and Q^*R and the arc PTR also has area A_{sector}. The sum of the areas of these two sectors exceeds the area of the overlap lens exactly by the area of the rhombus $PQRQ^*$. While an exact calculation is not difficult we can afford to be a little cavalier here with the bound $A_{\text{lens}} \le 2A_{\text{sector}} = \pi r^2 - 2r^2\arcsin\left(\frac{Z}{2r}\right)$. We hence obtain

$$A_2(X_1, X_{j^*}) \ge \pi r^2 + 2r^2\arcsin\left(\tfrac{Z}{2r}\right) \ge \pi r^2 + rZ$$

by virtue of the bound $\arcsin x \ge x$ (Lemma **??**). In consequence,

$$a_2(X_1, X_{j^*}) = \tfrac{1}{\pi}A_2(X_1, X_{j^*}) \ge r^2 + \tfrac{1}{\pi}rZ.$$

CASE 2: When $2r < Z \le (2k-2)r$ the regions of visibility of X_1 and X_{j^*} are disjoint (Figure 2(c)) and $A_2(X_1, X_{j^*}) = A_1(X_1) + A_1(X_{j^*}) = 2\pi r^2$ whence $a_2(X_1, X_{j^*}) = 2r^2$.

Write $F(z \mid X_1)$ and $f(z \mid X_1)$ for the conditional distribution function and density, respectively, of Z given X_1. Then for all $0 \le z \le 1 - |X_1|$, the circle of radius z centred at X_1 is wholly contained within the unit circle so that the conditional distribution satisfies $F(z \mid X_1) = \mathbf{P}\{Z \le z \mid X_1\} = \mathbf{P}\{|X_2 - X_1| \le z, \dots, |X_k - X_1| \le z \mid X_1\} = (z^2)^{k-1}$ as the distances $|X_i - X_1|$ $(2 \le i \le k)$ are conditionally independent given X_1; the corresponding conditional density is $f(z \mid X_1) = (2k-2)z^{2k-3}$. In particular, for all $X_1 \in \mathbb{I}$, the conditional distribution and density of Z satisfy $F(z \mid X_1) = z^{2k-2}$ and $f(z \mid X_1) = (2k-2)z^{2k-3}$ for $0 \le z \le (2k-2)r$. Putting the pieces together, for $X_1 \in \mathbb{I}$,

$$\begin{aligned}
&\mathbf{E}\left(\mathbf{1}_{[0,(2k-2)r]}(Z)e^{-na_2(X_1,X_{j^*})} \mid X_1\right) \\
&\quad= \mathbf{E}\left(\mathbf{1}_{[0,2r]}(Z)e^{-na_2(X_1,X_{j^*})} \mid X_1\right) + \mathbf{E}\left(\mathbf{1}_{(2r,(2k-2)r]}(Z)e^{-na_2(X_1,X_{j^*})} \mid X_1\right) \\
&\quad\le e^{-nr^2G}\int_0^{2r}(2k-2)z^{2k-3}e^{-nrzG/\pi}\,dz + e^{-2nr^2G}\,\mathbf{P}\{2r < Z \le (2k-2)r \mid X_1\} \\
&\quad\le (2k-2)e^{-nr^2G}\int_0^{\infty} z^{2k-3}e^{-nrzG/\pi}\,dz + e^{-2nr^2G}\,\mathbf{P}\{Z \le (2k-2)r \mid X_1\}
\end{aligned}$$

$$= \pi^{2k-2}(2k-2)!\,\frac{r^{2k-2}e^{-nr^2G}}{(nr^2G)^{2k-2}} + (2k-2)^{2k-2}r^{2k-2}e^{-2nr^2G}$$
$$= \mathcal{O}\left(\frac{r^{2k-2}}{n\log^{2k-2}n} + \frac{r^{2k-2}}{n^2}\right) = \mathcal{O}\left(\frac{r^{2k-2}}{n\log^{2k-2}n}\right)$$

where the order bounds are again uniform in X_1. As $\mathbf{E}\big(\mathbf{1}_{\mathbb{I}}(X_1)\big) = \mathbf{P}\big\{|X_1| \le 1-(2k-1)r\big\} = \big[1-(2k-1)r\big]^2 = 1+\mathcal{O}(r)$, substitution in (15) yields

$$K_{\text{interior}} = \mathcal{O}\left(\frac{r^{2k-2}}{n\log^{2k-2}n}\right) = \mathfrak{o}\left(\frac{1}{n^k}\right) \tag{16}$$

which is sub-dominant, if barely, compared to the non-overlap contribution $J_{\text{non-overlap}} \sim \lambda^k/n^k$ asymptotically as $n \to \infty$.

The boundary overlap contribution: As was the case for a single vertex, the calculations in the near-boundary are delicate and a careful case-by-case analysis is unavoidable. Write $\mathbf{1}_{[0,(2k-2)r]}(Z) = \mathbf{1}_{[0,2r]}(Z) + \mathbf{1}_{(2r,(2k-2)r]}(Z)$ so that the expression for K_{boundary} in (15) can be written as the sum of two terms, one for *proximate* $Z \le 2r$ when the visibility region of X_1 overlaps with that of each of X_2 through X_k and the other for *well-separated* $Z > 2r$ when the two points X_1 and X_{j^*} are not adjacent in the overlap graph. We begin by disposing of the latter case first.

Well-separated Z: If $Z > 2r$ then $a_2(X_1, X_{j^*}) = a_1(X_1) + a_1(X_{j^*})$ as the visibility regions of X_1 and X_{j^*} do not overlap (even though there is a chain of overlaps connecting them in the overlap graph). The lower bound for a_1 in (4) shows then that $a_2(X_1, X_{j^*}) \ge r^2 - \frac{2}{\pi}r^3$ for all sufficiently large n while, as seen above, $\mathbf{P}\big\{2r < Z \le (2k-2)r \mid X_1\big\} \le \big[(2k-2)r\big]^{2(k-1)}$ and $\mathbf{E}\big(\mathbf{1}_{\mathbb{S}\setminus\mathbb{I}}(X_1)\big) = 1-\mathbf{P}\big\{|X_1| \le 1-(2k-1)r\big\} \le (4k-2)r$. We hence obtain the asymptotic estimate

$$\begin{aligned}\mathbf{E}\big[\mathbf{1}_{\mathbb{S}\setminus\mathbb{I}}(X_1)\,\mathbf{E}\big(\mathbf{1}_{(2r,(2k-2)r]}(Z)e^{-na_2(X_1,X_{j^*})} \mid X_1\big)\big]&\\ \le (4k-2)(2k-2)^{2k-2}r^{2k-1}e^{-nr^2G+2nr^3G/\pi}&\\ = \mathcal{O}\big(\tfrac{r^{2k-1}}{n}\big) = \mathfrak{o}\big(\tfrac{\log^{2k-1}n}{n^{k+1/2}}\big) = \mathfrak{o}\big(\tfrac{1}{n^k}\big)&\end{aligned} \tag{17}$$

which, like the interior overlap contribution, is sub-dominant.

Proximate Z: We are left with the principal boundary overlap contributions arising from the regime $X_i \in \mathbb{S}\setminus\mathbb{I}$ and $Z \le 2r$. As X_1 varies over the boundary $\mathbb{S}\setminus\mathbb{I}$, three regions may be identified: (i) the *very-near-boundary*, $1-(2k-1)r < |X_1| \le 1-r$, when the disc of radius r centred at X_1 is completely contained within the unit disc; (ii) the *near-boundary*, $1-r < |X_1| \le \sqrt{1-r^2}$, as seen for a single vertex when part of the disc $\mathbb{D}(r, X_1)$ is lost to the visibility region due to intersection with the boundary; and (iii) the *far-boundary*, $\sqrt{1-r^2} < |X_1| \le 1$, also as seen in the analysis for a single vertex when the area of the visibility region of X_1 is approximately $\frac{1}{2}\pi r^2$. Accordingly, write

$$\mathbf{1}_{\mathbb{S}\setminus\mathbb{I}}(X_1) = \mathbf{1}_{(1-(2k-1)r,1-r]}(|X_1|) + \mathbf{1}_{(1-r,\sqrt{1-r^2}]}(|X_1|) + \mathbf{1}_{(\sqrt{1-r^2},1]}(|X_1|)$$

and partition the range of X_1 to obtain

$$\mathbf{E}\big[\mathbf{1}_{\mathbb{S}\setminus\mathbb{I}}(X_1)\,\mathbf{E}\big(\mathbf{1}_{[0,2r]}(Z)e^{-na_2(X_1,X_{j^*})} \mid X_1\big)\big] = B^{(i)}_{\text{boundary}} + B^{(ii)}_{\text{boundary}} + B^{(iii)}_{\text{boundary}} \tag{18}$$

where $B^{(i)}_{\text{boundary}}$ is the contribution from the very-near-boundary, $B^{(ii)}_{\text{boundary}}$ is the contribution from the near-boundary, and $B^{(iii)}_{\text{boundary}}$ is the contribution from the far-boundary.

Before proceeding further it will be useful to make two key observations at this point that will help to substantially simplify the calculations.

KEY OBSERVATION 1: *The conditional density of* Z *given* X_1 *has the uniform envelope* $f(z \mid X_1) \leq (2k-2)z^{2k-3}$ *for all* $z \geq 0$. Indeed, for any $i \neq 1$, the conditional distribution of $|X_i - X_1|$ given X_1 is bounded by

$$\mathbf{P}\bigl\{|X_i - X_1| \leq z \mid X_1\bigr\} = \frac{1}{\pi}\int_{\mathbb{D}(z,X_1)\cap\mathbb{S}} dx \leq \frac{1}{\pi}\int_{\mathbb{D}(z,X_1)} dx = z^2$$

for all $z \geq 0$ and it follows similarly that, for any $\epsilon > 0$,

$$\begin{aligned}\mathbf{P}\bigl\{z < |X_i - X_1| \leq z + \epsilon \mid X_1\bigr\} &= \frac{1}{\pi}\int_{\mathbb{D}(z+\epsilon,X_1)\setminus\mathbb{D}(z,X_1)\cap\mathbb{S}} dx \\ &\leq \frac{1}{\pi}\int_{\mathbb{D}(z+\epsilon,X_1)\setminus\mathbb{D}(z,X_1)} dx = (z+\epsilon)^2 - z^2 = 2z\epsilon + \epsilon^2\end{aligned}$$

as the intersection with the unit circle of the annulus of width ϵ at the boundary of the circle of radius z centred at X_1 has area certainly no larger than that of the annulus itself. For convenience introduce the notational shorthand $F_2 = \mathbf{P}\bigl\{|X_i - X_1| \leq z\bigr\}$ and $\Delta F_2 = \mathbf{P}\bigl\{z < |X_i - X_1| \leq z + \epsilon \mid X_1\bigr\}$. As the event $\{z < Z \leq z + \epsilon\}$ occurs if, and only if, for some $j \geq 1$, exactly j of the X_i $(2 \leq i \leq k)$ satisfy $z < |X_i - X_1| \leq z + \epsilon$ with the remaining $k - 1 - j$ of the X_i satisfying $|X_i - X_j| \leq z$ it follows that

$$\begin{aligned}\frac{1}{\epsilon}\bigl[F(z+\epsilon \mid X_1) - F(z \mid X_1)\bigr] &= \frac{1}{\epsilon}\,\mathbf{P}\{z < Z \leq z+\epsilon \mid X_1\} = \frac{1}{\epsilon}\sum_{j=1}^{k-1}\binom{k-1}{j}F_2^{k-1-j}\bigl[\Delta F_2\bigr]^j \\ &\leq \frac{1}{\epsilon}\sum_{j=1}^{k-1}\binom{k-1}{j}z^{2k-2-2j}\bigl(2z\epsilon + \epsilon^2\bigr)^j = \sum_{j=1}^{k-1}\binom{k-1}{j}z^{2k-2-2j}\epsilon^{j-1}(2z+\epsilon)^j.\end{aligned}$$

Allowing ϵ to tend to zero on both sides of the bound we obtain $f(z \mid X_1) \leq (2k-2)z^{2k-3}$ as claimed.

KEY OBSERVATION 2: *For given* X_1 *and* Z, $a_2(X_1, X_{j^*})$ *attains its minimum value when* X_{j^*} *is situated as closely as possible to the periphery of the unit disc.* In particular, if $Z \leq 1 - |X_1|$ this is occasioned when X_{j^*} lies on the ray from the origin through the point X_1 and athwart X_1 and the circumference of the unit circle. And if $Z > 1 - |X_1|$ this is occasioned when X_{j^*} is one of the two points on the circumference of the unit circle at distance Z from X_1. The validity of the observation may be seen as a consequence of the convexity of the circle whence $\mathbb{D}(r, X_{j^*}) \setminus \mathbb{D}(r, X_1)$ has its minimum area of overlap with the unit disc when X_{j^*} is situate on the point(s) on the circumference of the circle of radius Z centred at X_1 which is closest to the periphery of the unit disc. This observation is key as it allows us to repeatedly leverage our analysis for a single vertex.

The very-near-boundary: $1 - (2k-1)r < |X_1| \leq 1 - r$. The simplest bounds suffice here. It is trivial that $A_2(X_1, X_{j^*}) \geq A_1(X_1) = \pi r^2$ so that $a_2(X_1, X_{j^*}) \geq r^2$. We hence obtain

$$B^{(i)}_{\text{boundary}} \leq e^{-nr^2 G}\,\mathbf{E}\bigl[\mathbf{1}_{(1-(2k-1)r,1-r]}(|X_1|)\,\mathbf{P}\bigl\{Z \leq 2r \mid X_1\bigr\}\bigr]$$

$$\le e^{-nr^2G}(4r^2)^{k-1}\,\mathbf{P}\big\{1-(2k-1)r < |X_1| \le 1-r\big\}$$
$$= e^{-nr^2G}4^{k-1}r^{2k-2}\big[(1-r)^2-\big(1-(2k-1)r\big)^2\big]$$
$$\le 4^k(k-1)r^{2k-1}e^{-nr^2G} = \mathcal{O}\big(\tfrac{r^{2k-1}}{n}\big) = o\big(\tfrac{\log^{2k-1}n}{n^{k+1/2}}\big) = o\big(\tfrac{1}{n^k}\big) \tag{19}$$

which is easily sub-dominant with respect to the main contribution from the interior.

The near-boundary: $1-r < |X_1| \le \sqrt{1-r^2}$. As in the case of a single vertex, the situation is most delicate here and requires careful estimates. As Z varies from 0 to $2r$ two regions may be identified: (i) $0 \le Z \le 1-|X_1|$ where the smallest value of $a_2(X_1, X_{j^*})$ occurs when X_{j^*} lies on the ray outward from the origin through the point X_1, and (ii) $1-|X_1| < Z \le 2r$ when the smallest value of $a_2(X_1, X_{j^*})$ occurs when X_{j^*} is on the circumference of the unit circle at a distance Z from X_1. The key to the analysis of both cases is identifying suitable bounds for $a_2(X_1, X_{j^*})$ *explicitly as functions of* $1-|X_1|$ *and* Z.

(i) X_{j^*} *is collinear with* X_1 *and the origin*. In the worst case X_{j^*} is pinched into the narrow region between X_1 and the circumference of the unit disc and the worst-case area of the conjoined visibility region of X_1 and X_{j^*} will in consequence not be much larger than that of X_1 alone. But this is sufficient as we see next.

Leveraging our results for a single vertex we obtain

$$a_2(X_1, X_{j^*}) \ge a_1(X_1) \ge \tfrac{1}{2}r^2 + \tfrac{1}{\pi}r(1-|X_1|)$$

by an appeal to (3). Consequently,

$$\mathbf{E}\big[\mathbf{1}_{(1-r,\sqrt{1-r^2}]}(|X_1|)\,\mathbf{E}\big(\mathbf{1}_{[0,1-|X_1|]}(Z)e^{-na_2(X_1,X_{j^*})} \mid X_1\big)\big]$$
$$\overset{\text{(vii)}}{\le} 2(2k-2)e^{-nr^2G/2}\int_{1-r}^{\sqrt{1-r^2}} d\rho\,\rho e^{-nr(1-\rho)/\pi}\int_0^{1-\rho} dz\,z^{2k-3}$$
$$\overset{\text{(viii)}}{\le} 2\,e^{-nr^2G/2}\int_{1-r}^{\sqrt{1-r^2}} (1-\rho)^{2k-2}e^{-nr(1-\rho)/\pi}\,d\rho$$
$$\overset{\text{(ix)}}{=} \frac{2\pi^{2k-1}e^{-nr^2G/2}}{(nrG)^{2k-1}}\int_{nrG(1-\sqrt{1-r^2})/\pi}^{nr^2G/\pi} u^{2k-2}e^{-u}\,du$$
$$\le \frac{2\pi^{2k-1}e^{-nr^2G/2}}{(nrG)^{2k-1}}\int_0^\infty u^{2k-2}e^{-u}\,du$$
$$= \frac{2\pi^{2k-1}(2k-2)!r^{2k-1}e^{-nr^2G/2}}{(nr^2G)^{2k-1}} = \mathcal{O}\left(\frac{r^{2k-1}}{\sqrt{n}\,\log^{2k-1}n}\right) = o\left(\frac{1}{n^k}\right) \tag{20}$$

where (vii) follows from the key observation that the conditional density of Z has an envelope, in (viii) we use the fact that $\rho \le \sqrt{1-r^2} < 1$ inside the integral, and (ix) follows from the change of variable of integration to $u = nr(1-\rho)/\pi$.

(ii) X_{j^*} *is on the periphery*. The situation is somewhat more complex here as, even in the worst-case, X_{j^*} situated on the circumference of the unit disc starts contributing more significantly to the conjoined region of visibility. Two cartoon figures illustrating this situation are shown in Figure 3. An arc of the circumference of the unit disc is shown passing through X_{j^*} situated at the point Q^* in Figure 3(a). On the scale of the radius r of the region of visibility, the intersection of the circumference of the unit disc with the circles of radius r at Q and Q^* is almost a straight line as shown in (a). In order to make the critical region more visible the curvature of the sensor field vis à vis the regions of visibility has

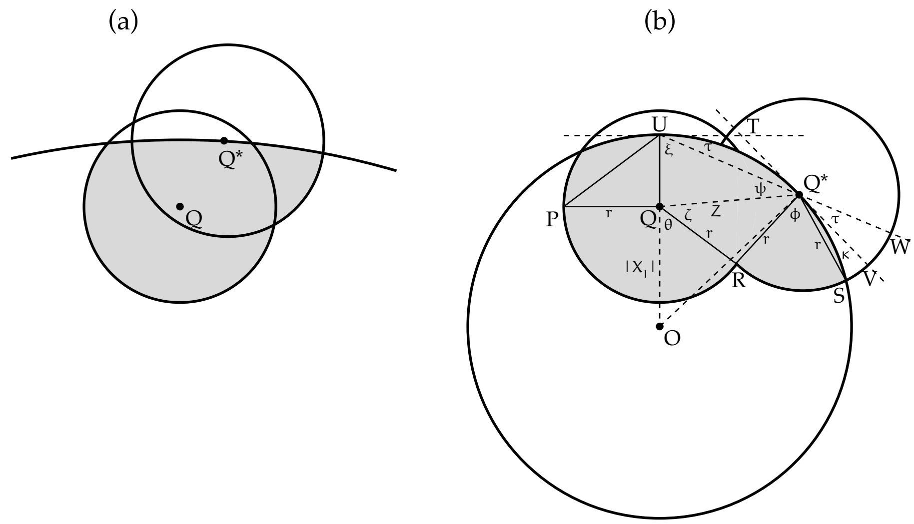

Figure 3: The near-boundary regime $1 - r < |X_1| \leq \sqrt{1 - r^2}$ for proximate Z satisfying $1 - |X_1| < Z \leq 2r$. The point X_1 is situate at Q with X_{J^*} at Q^* on the periphery of the sensor field. The conjoined region of visibility is shown shaded.

been much exaggerated in Figure 3(b) where X_1 is situated at point Q and X_{J^*} at a distance $Z > 1 - |X_1|$ from X_1 is located on the circumference of the sensor field at point Q^*. We take our notation from this figure.

The right-angle triangle PQU whose base PQ has length r and height QU has length $1 - |X_1|$, the sector PQR of angle $\pi/2 + \theta$ of the circle of radius r at Q, and the sector RQ^*S of angle ϕ of the circle of radius r at Q^* are non-overlapping and contained wholly within the conjoined visibility region $\mathbb{V}(X_1) \cup \mathbb{V}(X_{J^*})$. It follows that

$$A_2(X_1, X_{J^*}) \geq \tfrac{1}{2} r(1 - |X_1|) + \tfrac{1}{2} r^2 \left(\tfrac{\pi}{2} + \theta\right) + \tfrac{1}{2} r^2 \phi.$$

Elementary calculations now serve to determine the various angles of interest.

1. Let Z' denotes the length of the line segment UQ^*. As two sides of a triangle are larger than the third, a consideration of triangle QUQ^* shows that $Z' \leq Z + (1 - |X_1|) \leq 3r$ as $Z \leq 2r$ and $1 - |X_1| < r$.

2. As UOQ^* is an isosceles triangle with sides of length 1 and base of length Z', we have $\xi = \angle OUQ^* = \arccos\left(\frac{Z'}{2}\right) = \frac{\pi}{2} - \arcsin\left(\frac{Z'}{2}\right)$.

3. As UT is the tangent line at U, $\tau = \angle TUQ^* = \frac{\pi}{2} - \xi = \arcsin\left(\frac{Z'}{2}\right) \leq Z' \leq 3r$ for large enough n by virtue of the upper bound for the arcsine function from Lemma **??**. With TQ^*V the tangent line at Q^*, note that τ may also be identified with $\angle VQ^*W$.

4. As QRQ^* is an isosceles triangle with sides r and base Z, we have $\zeta = \angle RQQ^* = \angle RQ^*Q = \frac{\pi}{2} - \arcsin\left(\frac{Z}{2r}\right) \leq \frac{\pi}{2} - \frac{Z}{2r}$, by virtue of the lower bound for the arcsine function this time, again from Lemma **??**.

5. With $\kappa = \angle SQ^*V$, the development leading up to (4) in the analysis of the far-boundary for a single vertex shows that $r^2\left(\frac{\pi}{2} - \kappa\right) \geq \frac{\pi}{2} r^2 - r^3$ so that $\kappa \leq r$.

6. With $\psi = \angle QQ^*U$, it follows that $\phi = \pi - \psi - \zeta - \kappa - \tau \geq \frac{\pi}{2} - \psi + \frac{Z}{2r} - 4r$.

7. And finally, an examination of the exterior angle of the triangle UQQ^* shows that $\theta + \zeta = \xi + \psi$ so that $\theta = \frac{\pi}{2} - \tau + \psi - \zeta \geq \psi + \frac{Z}{2r} - 3r$.

Proceeding with bounding the area of the conjoined region of visibility, we have

$$A_2(X_1, X_{J^*}) \geq \tfrac{1}{2}r(1 - |X_1|) + \tfrac{1}{2}r^2\big(\tfrac{\pi}{2} + \theta + \phi\big) \geq \tfrac{1}{2}\pi r^2 + \tfrac{1}{2}r(1 - |X_1|) + \tfrac{1}{2}rZ - \tfrac{7}{2}r^3,$$

and, in consequence,

$$a_2(X_1, X_{J^*}) \geq \tfrac{1}{2}r^2 + \tfrac{1}{2\pi}r(1 - |X_1|) + \tfrac{1}{2\pi}rZ - \tfrac{7}{2\pi}r^3.$$

Again exploiting the fact that the conditional density of Z has a uniform envelope, we obtain

$$\begin{aligned}
&\mathbf{E}\big[\mathbf{1}_{(1-r,\sqrt{1-r^2}]}(|X_1|)\,\mathbf{E}\big(\mathbf{1}_{(1-|X_1|,2r]}(Z)e^{-na_2(X_1,X_{J^*})} \mid X_1\big)\big] \\
&\qquad \leq 2(2k-2)e^{-\frac{1}{2}nr^2G + \frac{7}{2\pi}nr^3G}\int_{1-r}^{\sqrt{1-r^2}} d\rho\,\rho e^{-nr(1-\rho)/2\pi}\int_{1-\rho}^{2r} dz\, z^{2k-3}e^{-nrzG/2\pi}.
\end{aligned}$$

Finish off by bounding the inner integral first,

$$\int_{1-\rho}^{2r} z^{2k-3}e^{-nrzG/2\pi}\,dz \leq \int_0^\infty z^{2k-3}e^{-nrzG/2\pi}\,dz = \frac{(2k-3)!(2\pi)^{2k-2}}{(nrG)^{2k-2}},$$

then the outer integral,

$$\begin{aligned}
\int_{1-r}^{\sqrt{1-r^2}} 2\rho e^{-nr(1-\rho)/2\pi}\,d\rho &\leq 2\int_{1-\sqrt{1-r^2}}^{r} e^{-nruG/2\pi}\,du \\
&\leq 2\int_0^\infty e^{-nruG/2\pi}\,du = \frac{4\pi}{nrG},
\end{aligned}$$

via the change of variable of integration $u \leftarrow 1 - \rho$, to finally obtain the bound

$$\begin{aligned}
&\mathbf{E}\big[\mathbf{1}_{(1-r,\sqrt{1-r^2}]}(|X_1|)\,\mathbf{E}\big(\mathbf{1}_{(1-|X_1|,2r]}(Z)e^{-na_2(X_1,X_{J^*})} \mid X_1\big)\big] \\
&\qquad \leq 2^{2k}\pi^{2k-1}(2k-2)!\,\frac{r^{2k-1}e^{-\frac{1}{2}nr^2G+\frac{7}{2\pi}nr^3G}}{(nr^2G)^{2k-1}} = \mathcal{O}\left(\frac{r^{2k-1}}{\sqrt{n}\,\log^{2k-1} n}\right) = o\left(\frac{1}{n^k}\right). \qquad (21)
\end{aligned}$$

Mop up by putting (20) and (21) together to get

$$\begin{aligned}
B^{(ii)}_{\text{boundary}} &= \mathbf{E}\big[\mathbf{1}_{(1-r,\sqrt{1-r^2}]}(|X_1|)\,\mathbf{E}\big(\mathbf{1}_{[0,2r]}(Z)e^{-na_2(X_1,X_{J^*})} \mid X_1\big)\big] \\
&= \mathcal{O}\big(\tfrac{r^{2k-1}}{\sqrt{n}\,\log^{2k-1} n}\big) = o\big(\tfrac{1}{n^k}\big) \qquad (22)
\end{aligned}$$

which is again of the requisite asymptotic order, though barely so for the given rate of $r = r_n$.

The far-boundary: $\sqrt{1-r^2} < |X_1| \leq 1$. Borrowing again from the analysis for a single vertex, we have from (4) that $a_2(X_1, X_{J^*}) \geq a_1(X_1) \geq \frac{1}{2}r^2 - \frac{1}{\pi}r^3$ for sufficiently large n. It follows that

$$B^{(iii)}_{\text{boundary}} = \mathbf{E}\big[\mathbf{1}_{(\sqrt{1-r^2},1]}(|X_1|)\,\mathbf{E}\big(\mathbf{1}_{[0,2r]}(Z)e^{-na_2(X_1,X_{J^*})} \mid X_1\big)\big]$$

$$\begin{aligned}
&\le e^{-\frac{1}{2}nr^2G+\frac{1}{\pi}nr^3G}\,\mathbf{E}\big[\mathbf{1}_{(\sqrt{1-r^2},1]}(|X_1|)\,\mathbf{P}\{Z\le 2r\mid X_1\}\big]\\
&\le e^{-\frac{1}{2}nr^2G+\frac{1}{\pi}nr^3G}(4r^2)^{k-1}\,\mathbf{P}\big\{\sqrt{1-r^2}<|X_1|\le 1\big\}\\
&\le e^{-\frac{1}{2}nr^2G+\frac{1}{\pi}nr^3G}(4r^2)^{k-1}\big[1-(1-r^2)\big]\\
&=\mathcal{O}\Big(\frac{r^{2k}}{\sqrt{n}}\Big)=\mathfrak{o}\Big(\frac{\log^{2k}n}{n^{k+1/2}}\Big)=\mathfrak{o}\Big(\frac{1}{n^k}\Big)
\end{aligned}\tag{23}$$

which is also asymptotically sub-dominant.

Finale. We can now trace our way back to the starting point. Stitching the boundary contributions for proximate Z together in (18) we have from (19), (22), and (23) that

$$\mathbf{E}\big[\mathbf{1}_{\mathbb{S}\setminus\mathbb{I}}(X_1)\,\mathbf{E}\big(\mathbf{1}_{[0,2r]}(Z)e^{-na_2(X_1,X_{j^*})}\mid X_1\big)\big]=\mathfrak{o}\big(n^{-k}\big)$$

which together with the estimate (17) when Z is well-separated implies in (15) that $K_{\text{boundary}}=\mathfrak{o}\big(n^{-k}\big)$, which, in turn, in conjunction with the corresponding estimate (16) for K_{interior} yields $K(n)=K_{\text{interior}}+K_{\text{boundary}}=\mathfrak{o}\big(n^{-k}\big)$, which then implies in (14) that

$$\mathbf{E}\big[\mathbf{1}_{\mathbb{C}_k}(X_1,\dots,X_k)e^{-na_k(X_1,\dots,X_k)}\big]=\mathfrak{o}\big(n^{-k}\big).\tag{24}$$

At long last the weariest river winds its way safely to sea. The last estimate returns us finally to (13) and (11) from which we conclude that $J_k=\mathfrak{o}\big(n^{-k}\big)$ and the contribution due to connected overlap graphs is asymptotically sub-dominant.

Overlapping regions of visibility, conclusion. Return to the system of equations (10, 10′) reproduced here for convenience. The contribution from all overlap graphs $\mathcal{F}\ne\mathcal{F}_*$ to the probability of the event that vertices 1 through k are all isolated is

$$J_{\text{overlap}}=\sum_{c=1}^{k-1}\sum_{\substack{k_1\ge\cdots\ge k_c\ge 1\\ k_1+\cdots+k_c=k}}J_{k_1,\dots,k_c}$$

where $J_{k_1,\dots,k_c}$ is the total contribution from all overlap graphs $\mathcal{F}_{k_1,\dots,k_c}$ with exactly c components of sizes $k_1,\dots,k_c$ and is given by

$$J_{k_1,\dots,k_c}=\sum_{\mathcal{F}_{k_1,\dots,k_c}}\mathbf{E}\big[\mathbf{1}_{\mathbb{C}(\mathcal{F}_{k_1,\dots,k_c})}(X_1,\dots,X_k)\,\mathbf{P}\big(L_{1,\dots,k}\mid X_1,\dots,X_k\big)\big].$$

Some algebraic simplification in the expression for $J_{k_1,\dots,k_c}$ can be achieved by observing that as the points $X_1,\dots,X_n$ are generated by independent sampling from the uniform distribution on the unit disc, it follows by symmetry that any two overlap graphs $\mathcal{F}_{k_1,\dots,k_c}$ and $\mathcal{F}'_{k_1,\dots,k_c}$ that differ only in a permutation of the vertices will contribute exactly the same amount to the sum for $J_{k_1,\dots,k_c}$. As there are exactly $k!/(k_1!\cdots k_c!)$ allocations of the vertices from 1 to k into c cells with respective occupancies $k_1,\dots,k_c$, it follows that

$$J_{k_1,\dots,k_c}=\frac{k!}{k_1!\cdots k_c!}\sum_{\mathcal{F}_{k_1,\dots,k_c}}{}'\;\mathbf{E}\big[\mathbf{1}_{\mathbb{C}(\mathcal{F}_{k_1,\dots,k_c})}(X_1,\dots,X_k)\,\mathbf{P}\big(L_{1,\dots,k}\mid X_1,\dots,X_k\big)\big]$$

where the sum $\sum'$ is now restricted to overlap graphs $\mathcal{F}_{k_1,\dots,k_c}$ for which the first component consists of the first k_1 vertices $\{1,\dots,k_1\}$, the second component consists of the next k_2 vertices $\{k_1+1,\dots,k_1+k_2\}$, and so on, with the cth component consisting of the final k_c vertices $\{k-k_c+1,\dots,k\}$. As in the case of the connected overlap graphs we now proceed to further consolidate these graphs.

Write $\mathbb{C}_{k_1,\dots,k_c}$ for the subset of k-tuples $(X_1, \dots, X_k)$ for which the corresponding overlap graph is divided into c components with the first component comprised of the vertices $\{1, \dots, k_1\}$, the second component of the vertices $\{k_1 + 1, \dots, k_1 + k_2\}$, and so on, with the cth component consisting of the vertices $\{k - k_c + 1, \dots, k\}$. It is then clear that the sum $\sum'_{\mathcal{F}_{k_1,\dots,k_c}} \mathbf{1}_{\mathbb{C}(\mathcal{F}_{k_1,\dots,k_c})}(X_1, \dots, X_k)$ is itself the indicator for the set $\mathbb{C}_{k_1,\dots,k_c}$ and consequently

$$J_{k_1,\dots,k_c} = \frac{k!}{k_1! \cdots k_c!} \mathbf{E}\bigl[\mathbf{1}_{\mathbb{C}_{k_1,\dots,k_c}}(X_1, \dots, X_k)\, \mathbf{P}\bigl(L_{1,\dots,k} \mid X_1, \dots, X_k\bigr)\bigr].$$

One final piece of notation to help compact expressions: for $j = 1, \dots, c$, introduce the nonce notation $x^{(j)} = (x_{k_1+\dots+k_{j-1}+1}, \dots, x_{k_1+\dots+k_j})$ to indicate the jth subsequence of length k_j and, mirroring the notation introduced for connected overlap graphs, let $\mathbb{C}_{k_j}$ denote the set of k_j-tuples $x^{(j)}$ in $\mathbb{S}^{k_j}$ for which the overlap graph $\mathcal{G}_{k_j}\bigl(x^{(j)}\bigr)$ is connected. We may now range through the k-tuples $(x_1, \dots, x_k)$ comprising the set $\mathbb{C}_{k_1,\dots,k_c}$ recursively as follows:

BASE: The k_1-tuple $x^{(1)}$ is allowed to range over the subset $\mathbb{C}^{(1)}_{k_1} = \mathbb{C}_{k_1}$ of $\mathbb{S}^{k_1}$ for which the overlap graph $\mathcal{G}_{k_1}\bigl(x^{(1)}\bigr)$ is connected.

RECURRENCE: As j ranges from 2 to c, for every selection of $x_1, \dots, x_{k_1+\dots+k_{j-1}}$, the k_j-tuple $x^{(j)}$ is allowed to range over the subset $\mathbb{C}^{(j)}_{k_j} = \mathbb{C}_{k_j} \setminus \bigl(\bigcup_{i=1}^{k_1+\dots+k_{j-1}} \mathbb{O}(x_i)\bigr)^{k_j}$ of $\mathbb{S}^{k_j}$ for which the overlap graph $\mathcal{G}_{k_j}\bigl(x^{(j)}\bigr)$ is connected while avoiding the overlap regions of the points $x_1, \dots, x_{k_1+\dots+k_{j-1}}$.

Write $dx^{(j)} = dx_{k_1+\dots+k_{j-1}+1} \cdots dx_{k_1+\dots+k_j}$ in the natural extension of the notation to the differential. It then follows that

$$J_{k_1,\dots,k_c} = \frac{k!}{k_1! \cdots k_c!} \frac{1}{\pi^k} \int_{\mathbb{C}^{(1)}_{k_1}} dx^{(1)} \int_{\mathbb{C}^{(2)}_{k_2}} dx^{(2)} \cdots \\ \cdots \int_{\mathbb{C}^{(c)}_{k_c}} dx^{(c)}\, \mathbf{P}\bigl(L_{1,\dots,k} \mid X_1 = x_1, \dots, X_k = x_k\bigr)$$

Now arguing as in (12), we obtain the bound

$$\mathbf{P}\bigl(L_{1,\dots,k} \mid X_1 = x_1, \dots, X_k = x_k\bigr) \le \bigl[1 + \mathcal{O}\bigl(\tfrac{\log^2 n}{n}\bigr)\bigr] e^{-n a_k(x_1,\dots,x_k)}$$

where, for each k-tuple $(x_1, \dots, x_k)$ in $\mathbb{C}_{k_1,\dots,k_c}$, we have

$$a_k(x_1, \dots, x_k) = a_{k_1}\bigl(x^{(1)}\bigr) + a_{k_2}\bigl(x^{(2)}\bigr) + a_{k_3}\bigl(x^{(3)}\bigr) + \dots + a_{k_c}\bigl(x^{(c)}\bigr)$$

as the visibility regions across components do not overlap. It then follows that

$$J_{k_1,\dots,k_c} \le \bigl[1 + \mathcal{O}\bigl(\tfrac{\log^2 n}{n}\bigr)\bigr] \frac{k!}{k_1! \cdots k_c!} \frac{1}{\pi} \int_{\mathbb{C}^{(1)}_{k_1}} dx^{(1)} e^{-n a_{k_1}(x^{(1)})} \times \\ \times \frac{1}{\pi} \int_{\mathbb{C}^{(2)}_{k_2}} dx^{(2)} e^{-n a_{k_2}(x^{(2)})} \cdots \frac{1}{\pi} \int_{\mathbb{C}^{(c)}_{k_c}} dx^{(c)} e^{-n a_{k_c}(x^{(c)})}. \quad (25)$$

A typical nested integral on the right-hand side is of the form

$$\frac{1}{\pi}\int_{\mathbb{C}^{(j)}_{k_j}} dx^{(j)} e^{-na_{k_j}(x^{(j)})} = \mathbf{E}\big[\mathbf{1}_{\mathbb{C}^{(j)}_{k_j}}\big(X^{(j)}\big)e^{-na_{k_j}(X^{(j)})} \mid X_1 = x_1, \dots, X_{k_1+\dots+k_{j-1}} = x_{k_1+\dots+k_{j-1}}\big] \quad (26)$$

and exhibits one of two distinct behaviours depending on the value of k_j. If $k_j = 1$ (that is the jth component consists of a singleton point) then as per our earlier calculations in (8) the integral differs from λ/n only by a multiplicative term $1 + o(1)$ as $\mathbb{C}^{(j)}_{k_j}$ differs from $\mathbb{S}^{k_j} = \mathbb{S}$ only in a region $\bigcup_{i=1}^{k_1+\dots+k_{j-1}} \mathbb{O}(x_i)$ whose area is of order $\mathcal{O}(r^2)$. If $k_j \geq 2$ on the other hand then the integral is bounded above by

$$\mathbf{E}\big[\mathbf{1}_{\mathbb{C}_{k_j}}\big(X^{(j)}\big)e^{-na_{k_j}(X^{(j)})}\big]$$

as $\mathbb{C}^{(j)}_{k_j} \subset \mathbb{C}_{k_j}$ and the set $\mathbb{C}_{k_j}$ is independent of $X_1, \dots, X_{k_1+\dots+k_{j-1}}$. But the last equation differs only notationally in the replacement of k by k_j from the expectation on the right-hand side of the bound (13) for J_k. The results of our analysis for connected overlap graphs hence carries over *in toto* and, as per the estimate (24), the integral (26) is asymptotically $o\big(n^{-k_j}\big)$.

Suppose now that there are exactly j components, say $i_1, \dots, i_j$, each of which consists of two or more vertices with the remaining $c - j$ components being singletons. Bear in mind that $j \geq 1$ as $c < k$ so that there is at least one non-singleton component. Clearly, also, $k_{i_1} + \dots + k_{i_j} = k - c + j$. Each of the $c - j$ singleton components contributes a multiplicative factor of $\frac{\lambda}{n}\big(1 + o(1)\big)$ to the right-hand side of (25) so that the cumulative multiplicative contribution of the singleton components to the expression for $J_{k_1,\dots,k_c}$ is $\frac{\lambda^{c-j}}{n^{c-j}}\big(1 + o(1)\big)$. On the other hand, the $j \geq 1$ non-singleton components contribute multiplicative factors of $o\big(n^{-k_{i_1}}\big), \dots, o\big(n^{-k_{i_j}}\big)$, respectively, for a cumulative multiplicative factor of $o\big(n^{-(k_{i_1}+\dots+k_{i_j})}\big)$. Putting both terms together we obtain the asymptotic estimate

$$J_{k_1,\dots,k_c} = o\left(\frac{1}{n^{k_{i_1}+\dots+k_{i_j}}} \times \frac{1}{n^{c-j}}\right) \qquad (n \to \infty).$$

Thus each of the summands in the expression for J_{overlap} is $o\big(n^{-k}\big)$ and as the sum ranges over only a finite number of terms we conclude that $J_{\text{overlap}} = o\big(n^{-k}\big)$ as well. Together with our hard-won estimate $J_{\text{non-overlap}} \sim \lambda^k/n^k$ in (9), this completes our evaluation of the system of equations (7).

Joint isolation probability. Returning finally to (6) we obtain $\mathbf{P}\big(L_{1,\dots,k}\big) = J_{\text{non-overlap}} + J_{\text{overlap}} = \frac{\lambda^k}{n^k}\big(1 + o(1)\big) + o\big(\frac{1}{n^k}\big) \sim \frac{\lambda^k}{n^k}$. It follows that, as claimed, *for any fixed positive integer* k *and any collection of vertices* $i_1, \dots, i_k$, $\mathbf{P}\big(L_{i_1,\dots,i_k}\big) \sim \frac{\lambda^k}{n^k}$ *as* $n \to \infty$. Recall that our analysis for a single vertex revealed that $\mathbf{P}\big(L_i\big) = P_0 \sim \frac{\lambda}{n}$. Thus, for any group of k vertices we have shown that $\mathbf{P}\big(L_{i_1} \cap \dots \cap L_{i_k}\big) \sim \big[P_0\big]^k$ and we may paraphrase this succinctly, if somewhat imprecisely, by saying that the events L_i are weakly asymptotically independent.

2.3 A Poisson Sieve

The stage is set for an application of Bonferroni's inequalities (Lemma **??**). Let $S_k(n)$ denote the sum of all the probabilities of k-fold conjunctions of the events L_i $(1 \leq i \leq n)$. As

$n \to \infty$ we obtain

$$S_k(n) = \sum_{1 \le i_1 < \cdots < i_k \le n} \mathbf{P}\big(L_{i_1} \cap \cdots \cap L_{i_k}\big) = \binom{n}{k} \mathbf{P}\big(L_{1,\dots,k}\big) \sim \binom{n}{k} \frac{\lambda^k}{n^k}.$$

Now $\binom{n}{k} = \frac{1}{k!} n(n-1)\cdots(n-k+1) = \frac{n^k}{k!}\big[1 + \mathcal{O}\big(\frac{1}{n}\big)\big] \sim \frac{n^k}{k!}$ where, of course, k is fixed and n tends to infinity. It follows that, for every positive integer k, $S_k(n) \to \frac{\lambda^k}{k!}$ as $n \to \infty$. Hence, for every fixed K,

$$\sum_{k=0}^{K} (-1)^k \binom{m+k}{m} S_{m+k}(n) \to \sum_{k=0}^{K} \frac{(-1)^k \lambda^{m+k}}{(m+k)!} \frac{(m+k)!}{m!k!} = \frac{\lambda^m}{m!} \sum_{k=0}^{K} \frac{(-\lambda)^k}{k!}$$

as $n \to \infty$. We recognise the truncated Taylor series for $e^{-\lambda}$ on the right-hand side. As the exponential series $e^x = \sum_{k=0}^{\infty} x^k/k!$ converges absolutely (and uniformly over every closed and bounded interval) it follows that for every $\epsilon > 0$ we can select $K = K(\epsilon)$ so that

$$\left| \sum_{k=0}^{K} \frac{(-\lambda)^k}{k!} - e^{-\lambda} \right| < \epsilon.$$

Let N_0 denote the number of isolated vertices and let m be any nonnegative integer. Then the event $\{N_0 = m\}$ occurs if, and only if, precisely m of the events $\{\, L_i, 1 \le i \le n \,\}$ occur. By Lemma **??**, for every $\epsilon > 0$ we may select a sufficiently large value of K so that

$$(e^{-\lambda} - \epsilon)\frac{\lambda^m}{m!} < \sum_{k=0}^{2K-1} (-1)^k \binom{m+k}{m} S_{m+k}(n) \le \mathbf{P}\{N_0 = m\}$$
$$\le \sum_{k=0}^{2K} (-1)^k \binom{m+k}{m} S_{m+k}(n) < (e^{-\lambda} + \epsilon)\frac{\lambda^m}{m!}$$

with the book-end inequalities on either side holding for all sufficiently large n. As the tiny ϵ is arbitrary, it follows that $\mathbf{P}\{N_0 = m\} \to e^{-\lambda}\lambda^m/m!$ as $n \to \infty$, as advertised. This completes the proof of the theorem.

3 Threshold Function for Connectivity

We will show that the random graph $\mathcal{G}(n, r_n)$ almost surely consists only of isolated vertices and a single giant component. Our results from the previous section will then quickly yield a threshold function for connectivity.

Begin by observing that the graph $\mathcal{G}(n, r_n)$ is disconnected if, and only if, it contains a component of order k for some $1 \le k \le n/2$. A component of order $k = 1$ is an isolated vertex and, as we've already seen, the probability that there are no isolated vertices is $e^{-e^{-c}} + o(1)$. The proof shows additionally that the probability there exists a component of order $k \ge 2$ is $o(1)$ for every fixed k. This follows via an obvious modification of the argument estimating J_k for overlap graphs—simply strike out the odd reference to $2r$ in the analysis of the connected overlap graph $\mathcal{F}_k$ of order k and replace it by r to get the corresponding considerations for a component of order k in the original graph $\mathcal{G}(n, r_n)$. As is easy to verify, none of the estimates are materially affected. A tightening of the (slightly over-zealous) bounds for the truncated gamma function in the estimates for J_k allows us to extend the result to components of order k increasing with n though this will entail a messy retracing of our path through all the boundary effects.

A more direct path to showing that there are almost surely no components of order $2 \le k \le n/2$ leverages results from continuum percolation and in the interests of space we present this approach here. These results were also exploited by Gupta and Kumar [5] though we will need to tighten the arguments to get the claimed sharp asymptotics.

In a slight and temporary abuse of notation, suppose X is a Poisson point process with intensity λ in the Euclidean plane $\mathbb{R}^2$. The *Poisson random-connection model* (X, λ, r) is a random graph in the plane where there is an edge between any two points x_1 and x_2 of the point process X if, and only if, $|x_1 - x_2| \le r$. The model has two key properties:

1. *There is at most one unbounded compenent almost surely.*
2. *The probability that any given point of the process* X *belongs to a bounded component is asymptotic to the probability that the point is isolated. More precisely, to foreshadow notation to come, write* $q(\lambda, r)$ *for the probability that a given point of the process* X *lies in an unbounded component and* $q_0(\lambda, r)$ *for the probability that the point is isolated. Then* $1 - q(\lambda, r) \sim q_0(\lambda, r) = e^{-\lambda r^2}$ *as* $\lambda \to \infty$.

These results are specialisations of Theorems 6.3 and 6.4, respectively, in Meester and Roy [9].

Now let $\mathcal{P}(\lambda, r)$ be the the random graph obtained from the restriction of the Poisson random-connection model (X, λ, r) to the unit disc $\mathbb{S}$. We are interested in the parametrisation $\lambda = \lambda_n$ and $r = r_n$ where $\lambda_n = \frac{1}{\pi}\bigl(n - \beta\sqrt{n}\bigr)$ for a suitably large choice of $\beta > 0$ to be specified shortly and, as before, $r_n = \sqrt{\frac{1}{n}\bigl(\log n + c + o(1)\bigr)}$. Let N be the (random) number of vertices in $\mathcal{P}(\lambda_n, r_n)$. Then N has a Poisson distribution with mean $\pi\lambda_n = n - \beta\sqrt{n} \sim n$. Now fix any $\epsilon > 0$. Chebyshev's inequality then readily shows that we may select $\beta = \beta(\epsilon)$ sufficiently large so that

$$\mathbf{P}\bigl\{n - 2\beta\sqrt{n} \le N \le n\bigr\} = \sum_{n-2\beta\sqrt{n} \le m \le n} e^{-\lambda_n} \frac{\lambda_n^m}{m!} \ge 1 - \epsilon,$$

or, equivalently,

$$\sum_{m \notin [n-2\beta\sqrt{n}, n]} e^{-\lambda_n} \frac{\lambda_n^m}{m!} \le \epsilon.$$

It follows that N is concentrated at n asymptotically. Now, conditioned on the event $\{N = m\}$, the points $X_1, \dots, X_m$ comprising the vertices of $\mathcal{P}(\lambda_n, r_n)$ are independent and identically distributed uniformly in the unit disc. With high probability the graphs $\mathcal{G}(n, r_n)$ and $\mathcal{P}(\lambda_n, r_n)$ hence differ only in a sub-dominant number of vertices of the order of $\mathcal{O}\bigl(\sqrt{n}\bigr)$ and it should be possible to deduce results for one from considerations of the other. We formalise the inter-dependence between these graphs next.

Write $Q(\lambda, r)$ for the probability that $\mathcal{P}(\lambda, r)$ is not connected and $Q_0(\lambda, r)$ for the probability that $\mathcal{P}(\lambda, r)$ contains isolated vertices. Likewise, write $P(m, r)$ for the probability that the random graph $\mathcal{G}(m, r)$ is not connected and $P_0(m, r)$ for the probability that $\mathcal{G}(m, r)$ contains isolated vertices.

First consider isolated vertices. By conditioning on the number of vertices in $\mathcal{P}(\lambda_n, r_n)$ we obtain

$$Q_0(\lambda_n, r_n) = \sum_{m=1}^{\infty} P_0(m, r_n) e^{-\lambda_n} \frac{\lambda_n^m}{m!} = \sum_{n-2\beta\sqrt{n} \le m \le n} P_0(m, r_n) e^{-\lambda_n} \frac{\lambda_n^m}{m!} + \zeta_1$$

where $0 \le \zeta_1 \le \epsilon$. In the range $n - 2\beta\sqrt{n} \le m \le n$ we have $n = m + \mathcal{O}(\sqrt{m})$ whence

$$r_n = \sqrt{\tfrac{1}{n}\big(\log n + c + o(1)\big)} = \sqrt{\tfrac{1}{m+\mathcal{O}(\sqrt{m})}\big\{\log\big(m + \mathcal{O}(\sqrt{m})\big) + c + o(1)\big\}}$$
$$= \sqrt{\tfrac{1}{m}\big(\log m + c + \mathcal{O}(\tfrac{\log m}{\sqrt{m}}) + o(1)\big)} = \sqrt{\tfrac{1}{m}\big(\log m + c + o(1)\big)}$$

as $n \to \infty$. Our estimates for vertex isolation in $\mathcal{G}(n, r_n)$ hence show that, for $m = n + \mathcal{O}(\sqrt{n})$,

$$P_0(m, r_n) = P_0(m, r_m) + o(1) = 1 - e^{-e^{-c}} + o(1)$$

as $n \to \infty$. It follows that

$$Q_0(\lambda_n, r_n) = \big(1 - e^{-e^{-c}}\big) \sum_{n-2\beta\sqrt{n} \le m \le n} e^{-\lambda_n} \frac{\lambda_n^m}{m!} + \zeta_1 + o(1)$$
$$= 1 - e^{-e^{-c}} + \zeta_1 + \zeta_2 + o(1)$$

where $|\zeta_2| \le \epsilon$.

With the natural convention that a graph without vertices is connected, consider the restricted graph $\mathcal{P}(\lambda_n, r_n)$. If it contains isolated vertices then surely the graph is not connected. If, on the other hand, there are no isolated vertices and the graph $\mathcal{P}(\lambda_n, r_n)$ is not connected then at least one of the vertices of $\mathcal{P}(\lambda_n, r_n)$ a.s. belongs to a bounded component of order $k \ge 2$ in the Poisson random-connection model (X, λ_n, r_n). Now the probability that a given vertex belongs to a bounded component of order $k \ge 2$ is $o\big(q_0(\lambda_n, r_n)\big)$ where $q_0(\lambda_n, r_n) = e^{-\pi\lambda_n r_n^2} \sim e^{-c}/n$ is the probability that a given point of the process X is isolated in (X, λ_n, r_n). It follows that the probability that at least one of the vertices of $\mathcal{P}(\lambda_n, r_n)$ belongs to a bounded component of order $k \ge 2$ is bounded above by $n \times o\big(q_0(\lambda_n, r_n)\big)$ which is $o(1)$. It follows that

$$Q(\lambda_n, r_n) = Q_0(\lambda_n, r_n) + o(1) = 1 - e^{-e^{-c}} + \zeta_1 + \zeta_2 + o(1) \qquad (n \to \infty).$$

On the other hand, by conditioning on the number of vertices in $\mathcal{P}(\lambda_n, r_n)$, we obtain

$$Q(\lambda_n, r_n) = \sum_{m=1}^{\infty} P(m, r_n) e^{-\lambda_n} \frac{\lambda_n^m}{m!} = \sum_{n-2\beta\sqrt{n} \le m \le n} P(m, r_n) e^{-\lambda_n} \frac{\lambda_n^m}{m!} + \zeta_3 \tag{27}$$

where $0 \le \zeta_3 \le \epsilon$. Now, for each $m \ge 2$, we have the recursive specification

$$P(m, r_n) \le \mathbf{P}\{\text{vertex } m \text{ is isolated in } \mathcal{G}(m, r_n)\} + P(m-1, r_n)$$

so that an easy induction yields

$$P(n, r_n) \le \sum_{j=m+1}^{n} \mathbf{P}\{\text{vertex } j \text{ is isolated in } \mathcal{G}(j, r_n)\} + P(m, r_n).$$

Again from our considerations for isolated vertices, in the range $n - 2\beta\sqrt{n} \le m < j \le n$, we have

$$\mathbf{P}\{\text{vertex } j \text{ is isolated in } \mathcal{G}(j, r_n)\} = e^{-(j-1)r_n^2}\big(1 + o(1)\big)$$
$$= e^{-[n+\mathcal{O}(\sqrt{n})]r_n^2}\big(1 + o(1)\big) = \tfrac{e^{-c}}{n}\big(1 + o(1)\big)$$

so that

$$P(n,r_n) \le \tfrac{(n-m)e^{-c}}{n}\big(1+o(1)\big) + P(m,r_n)$$
$$\le \tfrac{2\beta\sqrt{n}e^{-c}}{n}\big(1+o(1)\big) + P(m,r_n) = P(m,r_n) + \mathcal{O}\big(\tfrac{1}{\sqrt{n}}\big).$$

Returning to (27) we obtain

$$Q(\lambda_n,r_n) \ge P(n,r_n) \sum_{n-2\beta\sqrt{n}\le m\le n} e^{-\lambda_n}\frac{\lambda_n^m}{m!} + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) + \zeta_3 = P(n,r_n) + \zeta_3 + \zeta_4 + o(1)$$

where $|\zeta_4| \le \epsilon$. Thus,

$$P(n,r_n) \le Q(\lambda_n,r_n) - \zeta_3 - \zeta_4 + o(1) = Q_0(\lambda_n,r_n)\big(1+o(1)\big) - \zeta_3 - \zeta_4 + o(1)$$
$$= 1 - e^{-e^{-c}} + \zeta_1 + \zeta_2 - \zeta_3 - \zeta_4 + o(1) \le 1 - e^{-e^{-c}} + 4\epsilon + o(1).$$

On the other hand, $\mathcal{G}(n,r_n)$ is certainly not connected if there exist any isolated vertices, whence

$$P(n,r_n) \ge P_0(n,r_n) = 1 - e^{-e^{-c}} + o(1).$$

It follows that

$$1 - e^{-e^{-c}} + o(1) \le P(n,r_n) \le 1 - e^{-e^{-c}} + 4\epsilon + o(1).$$

As $\epsilon > 0$ may be taken arbitrarily small, we obtain $P(n,r_n) \to 1 - e^{-e^{-c}}$ as $n \to \infty$, or, equivalently,

$$\mathbf{P}\{\mathcal{G}(n,r_n) \text{ is connected}\} \to e^{-e^{-c}} \qquad (n \to \infty).$$

This concludes the proof of Theorem **??**.

References

[1] B. Bollobás, *Random Graphs*. New York: Academic Press, 1985.

[2] L. Booth, J. Bruck, M. Franceschetti, and R. Meester, "Covering algorithms, continuum percolation and the geometry of wireless networks." http://www.paradise.caltech.edu/ETR.html

[3] P. Erdös and A. Renyi, "On the evolution of random graphs," *Magyar Tud. Akad. Mat. Kut. Int. Kozl.*, vol. 5, pp. 17–61, 1960.

[4] W. Feller, *An Introduction to Probability Theory and Its Applications, Volume I*. New York: Wiley, 1965.

[5] P. Gupta and P. R. Kumar, "Critical power for asymptotic connectivity in wireless networks," in *Stochastic Analysis, Control, Optimization and Applications: A Volume in Honor of W.H. Fleming* (eds. W. M. McEneany, G. Yin, and Q. Zhang). Boston, MA: Birkhauser, pages 547–566, 1998.

[6] O. Häggström and R. Meester, "Nearest neighbor and hard sphere models in continuum percolation," *Random Structures and Algorithms*, vol. 9, pp. 295–315, 1996.

[7] P. Hall, *Introduction to the Theory of Coverage Processes*. New York: John Wiley and Sons, 1988.

[8] S. S. Kunniyur and S. S. Venkatesh, "Threshold functions, node isolation, and emergent lacunae in sensor networks," preprint, 2004.

[9] R. Meester and R. Roy, *Continuum Percolation*. Cambridge, UK: Cambridge University Press, 1996.

[10] T. K. Philips, S. S. Panwar, and A. N. Tantawi, "Connectivity properties of a packet radio network model," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1044–1047, 1989.

[11] F. Xue and P. R. Kumar, "The number of neighbours needed for connectivity of wireless networks. http://decision.csl.uiuc.edu/ prkumar, April 2002.

DEPARTMENT OF ELECTRICAL AND SYSTEMS ENGINEERING, UNIVERSITY OF PENNSYLVANIA, PHILADELPHIA, PA 19104

E-mail address: venkatesh@ee.upenn.edu