On The Sharp Threshold Interval Length of Partially Connected Random Geometric Graphs During K-Means Classification

Robert A. Murphy, Ph.D.

e-mail: robert.a.murphy@wustl.edu

Abstract: In K-means classification, a set of data will form clusters, i.e. classes, if the measured distances between data points (or some common point in each class) are below a certain threshold. With the assumption that the data points are randomly generated throughout some bounded region according to a certain probability distribution, we estimate the mean number of classes to form with high probability.

Keywords and phrases: probability, sharp threshold, random geometric graph, random cluster, K-means, machine learning, classification.

1. Introduction

Suppose infinitely many copies of a bounded structure are used to partition \( \mathbb{R}^2 \) and let \( B \subset \mathbb{R}^2 \) be a bounded subset containing finitely many copies of the bounded structure. Further suppose that structures in the partition are neighbors, if their respective boundaries have non-empty intersection and infinitely many of the bounded structures in the partition are individually occupied by exactly one point at the center of the structure, independently of all other structures. In [24] and [25], it is shown that, if the probability of neighboring structures each containing related points is greater than some critical value, then with probability 1, a path can be traced from any starting occupied bounded structure to any ending occupied bounded structure, with the path in between the start and the end consisting entirely of neighboring occupied structures. From this statement, the contrapositive statement is obtained such that, if the probability of neighboring structures each containing related points is less than or equal to the same critical value, then with probability 1, no such path exists for any two bounded structures. Hence, all points are either related to no other points or only related to finitely many points in neighboring bounded structures. It is this contrapositive condition that is of interest and great use during K-means classification, as classes are formed by groupings of inter-related data points.

Rarely does inter-related, real-world data conform to a predefined, rigid partition, as described above. As such, after removing the rigid partition of \( \mathbb{R}^2 \), suppose that the data are modeled by a node process which randomly generates points within \( B \) according to some predetermined probability distribution. Points in \( B \) are inter-related, if they are within a certain distance of one another or some common point, such as the average of a set of previously-grouped, inter-related points. In [37], it is shown that, with probability 1, there is a path of inter-related points between any two points in \( B \), if the number of points relative to the area of \( B \) is beyond a certain critical number or, if the maximum distance between inter-related points is larger than a certain critical number.

In [39], it is proven in cor. (7.4.39) that an ordered set of data, which is assumed to be spatially uniformly distributed, will form clusters, i.e. classes, if the measured distances between data points
(or some common data point in each class) are below a certain threshold, which is computed as a function of the number of data points sampled from the total population of data. We get around the computational expense by making the same assumptions introduced in Murphy \cite{39,40} of having applying an order statistic to a normally distributed sample and making the argument that this process results in a uniformly distributed sample.

Finally, with the defined partition \cite[Section (2.2)]{40}, it is shown that under certain conditions, no approximation of probabilities in the continuum is required to prove the existence of a path of any order, as in \cite{37}. Instead, the probability of a long range path in the continuum is equivalent to the probability of a long range path, over the same set of points, in the presence of the defined posterior partition when the maximum radial distance between connected points falls within a certain bounded interval. On this bounded interval, the probability of the existence of a long range path rises sharply when points connect at distances within the bounded interval. Lastly, the probability measure in question is found to be a unique random cluster measure which realizes a set of conditional probability measures. As such, the node process samples from the collection of conditional probability measures to form classes, when points connect at a distance less than or equal to the critical length.

2. Related Works

In \cite{10}, Cai, et.al. investigate the problem of partial connectivity of randomly distributed points in a bounded region by making the assumption that, relative to the size of the bounded region, the number of points to be generated is relatively small. As such, a Poisson-distributed node process generates an independent set of points in the designated region. Copies of a hexagon of some fixed, immutable size, which is not dependent upon distances between generated points, are used to partition the bounded region. Points in the region are deemed to connect to form an open edge, if after the region is partitioned, the points lie within the same hexagon or neighboring hexagons, where hexagons are neighbors, if their respective boundaries have non-empty intersection. Otherwise, the edge between two points is closed. Likewise, they define the logical points at the centers of two neighboring hexagons to connect to form an open edge, if each neighboring hexagon independently contains at least one of the generated points. In \cite{10}, Cai, et.al. compute probabilities as a function of the density of hexagons which are occupied by at least one point. They showed that if the number of hexagons in the fixed partition is unbounded and the number of points generated within the continuum of the bounded region is below (or at) a critical threshold, then the probability of a majority of the occupied hexagons (and points contained therein) in the bounded region being connected in a contiguous path will tend to zero. On the other hand, if the density of occupied hexagons is within a short interval around the critical threshold, then a connecting path of hexagons or points, from any start to any end, occurs with probability that rises sharply from some small positive value to a value close to 1 for densities that fall within the range of the short interval.

\cite[Thm. (1.1)]{23} gives an estimate of the length of the short interval. If the area of the bounded region is assumed to be one, without loss of generality, then the estimate of the length of the short interval can never be any better than \( \Theta \left( \log^{1/4}(n) \sqrt{\frac{\log(n)}{n}} \right) \), where \( n \) is the number of points generated within the bounded region by the node process.

This work uses the distance notion of connectivity without the presence of a partition, the same as in \cite{10}. However, it markedly differs in that the prototypical hexagon used in the defined
posterior partition of the bounded region is allowed to change in size, if the node process is stopped and started again after a partition has already been defined. Since the prototypical hexagon is allowed to change in size so that the logical centers are closer to (or further away from) each other and point density is inversely proportional to the maximum connection length, therein lies an added advantage when calculating the probability of a connected path of hexagons or points. Moreover, if the prototypical hexagon always shrinks as the node process generates more points, then it should be expected that the critical threshold and the length of the short interval around the critical threshold are intertwined. It is shown that this is, in fact, the case.

3. Random Geometric Graphs

3.1. Definitions

Definition 1 A node process is a mapping \( X : \mathbb{R}^2 \to \mathbb{R}^2 \) such that for any subset \( B \subset \mathbb{R}^2 \), there is an \( n \in \mathbb{N} \) and a subset \( X_n = \{x_k\}_{1 \leq k \leq n} \subset B \) such that \( X(B) = X_n \).

Definition 2 Suppose \( B \subset \mathbb{R}^2 \) and \( X \) is a node process that randomly generates independent points \( X_n = \{x_k\}_{1 \leq k \leq n} \subset B \) according to some probability distribution. Let \( d : \mathbb{R}^2 \to \mathbb{R}^2 \) be a distance measure. Points \( x, y \in X_n \) are said to be \( r \)-connected and form an \( r \)-open edge, if \( d(x, y) \leq r \), for some fixed \( r > 0 \). Points \( x, y \in X_n \) are \( r \)-disconnected and form an \( r \)-closed edge otherwise.

Definition 3 Let \( E \) be the set of edges between points in \( X_n \). \( G(X_n; r) \) is defined to be the \( r \)-graph of the set of all \( r \)-open and \( r \)-closed edges from \( E \) between points in \( X_n \).

Definition 4 Given points \( x, y \in X_n \), denote the edge between \( x \) and \( y \) as \( < x, y >_r \). A subset of points \( C \subseteq X_n \) forms an \( r \)-connected cluster if and only if given any \( x, y \in C \), there exists \( r \)-open edges \( < x, a_1 >_r, < a_1, a_2 >_r, ..., < a_{k-1}, y >_r \in E \) connecting \( x \) to \( y \), for points \( \{a_1, a_2, ..., a_{k-1}\} \subseteq C \).

Definition 5 Let \( A \) be a set of graphs of \( E \) and \( G(X_n; r) \in A \). \( A \) is said to be an increasing property if and only if for \( r' \neq r \) such that \( G(X_n; r) \subseteq G(X_n; r') \), it is true that \( G(X_n; r') \in A \).

Definition 6 Let \( P \) be a probability measure on \((\Omega, \mathcal{F})\). If \( A \) is a monotone (increasing) property and \( \epsilon \in (0, \frac{1}{2}) \), define

\[
r(n, \epsilon) = \inf\{r > 0 : P(G(X_n; r) \in A) \geq \epsilon\}
\]

and

\[
\Delta(n, \epsilon) = r(n, 1 - \epsilon) - r(n, \epsilon).
\]

If \( \Delta(n, \epsilon) = o(1) \), then \( A \) has a sharp threshold.

3.2. An Important Result

Theorem 7 [23, Thm. (1.1)] For increasing properties \( A \) consisting of graphs of points \( X_n \subset \mathbb{R}^2 \),

\[
\Delta(n, \epsilon) = \Theta(r_c \log^{1/4}(n))
\]

where

\[
r_c = O\left(\sqrt{\frac{\log n}{n}}\right).
\]
These writings will be concerned, at least in part, with estimating the length of this critical interval for a particular property $\mathcal{A}$, using this framework.

### 3.3. Procedure

Let $\mathcal{B} \subset \mathbb{R}^2$ be a bounded region containing the origin $\hat{0} = (0, 0)$ and let $X$ be a node process such that $X(\mathcal{B}) = \mathcal{X}_n$ is a set of $n$ points which are uniformly distributed spatially throughout $\mathcal{B}$, where $n$ is a Poisson random variable which takes a particular value (denoted as $n$) with density parameter $\lambda = \lambda(n)$ indicating the expected number of points generated per unit area of $\mathcal{B}$. For some fixed $r > 0$, points in $\mathcal{X}_n$ will connect if their mutual distance is within $r$. For fixed $\rho \in \left(\frac{1}{2}, 1\right)$, define $\mathcal{A}^r_{n,\rho}$ to be the set of all subsets of $\mathcal{X}_n$ containing at least $100\rho\%$ of all generated points which form a connected subset containing $\hat{0}$.

Let $\epsilon > 0$ be given and let $r(n, \rho, \epsilon)$ be the least connectivity radius $r > 0$ such that $P(\mathcal{A}^r_{n,\rho}) \geq \epsilon$. It will be shown that $P(\mathcal{A}^r_{n,\rho})$ is an increasing function of the connection radius $r$. The aim is to estimate the length of the interval of connectivity radii such that the occurrence of $\mathcal{A}^r_{n,\rho}$ increases in probability from a value of $\epsilon$ to a value of $1 - \epsilon$ on the interval of radii. On this interval will be associated a particular radius such that the probability of the occurrence of $\mathcal{A}^r_{n,\rho}$ is $1/2$. On the left half of the interval, it is more likely that classes will form, with each containing less than half of all points so that no one class contains the majority of data points. No one class containing the majority of data points is important since, in this event, no one class will contain all data points with probability 1, which is guaranteed by [37] in the continuum and [25] in the partitioned continuum.

As an integral step in estimating the length of the interval of radii, continuity in $r$ and $\rho$ of $P(\mathcal{A}^r_{n,\rho})$ will be shown. As such, by continuity in $\rho$, for small $\delta > 0$, the probability of $\mathcal{A}^r_{n,\rho}$ is close to the probabilities of $\mathcal{A}^r_{n,\rho+\delta}$ and $\mathcal{A}^r_{n,\rho-\delta}$. Furthermore, by continuity and the increasing nature of $P(\mathcal{A}^r_{n,\rho})$ in $r$, there exists $r_0 = r_0(n, \rho, \epsilon)$ such that $P(\mathcal{A}^{r_0}_{n,\rho}) = 1/2$. This particular radius of connectivity demarcates the point, beyond which, the generated set of points will transition from a set of disjoint classes to one giant, inter-related class of points, almost surely. Furthermore, for $\epsilon > 0$, this radius of connectivity is the center of the estimated interval of radii, upon which, $P(\mathcal{A}^r_{n,\rho})$ increases from $\epsilon$ to $1 - \epsilon$.

### 3.4. Definitions

**Definition 8** Given a fixed point, $y \in \mathcal{X}_n$, an $r$-connected component containing $y$ is the subset of points $< C_y >_r \subseteq \mathcal{X}_n$ containing $y$ and every $x \in \mathcal{X}_n \setminus \{y\}$ having a chain of $r$-open edges connecting $x$ to $y$.

**Definition 9** Given an $r$-open edge, $e = < x, y >_r \in G(\mathcal{X}_n; r)$, an $r$-connected component containing $e$ is the subset of points $< C_e >_r \subseteq \mathcal{X}_n$ containing $x$ and $y$ together with every $z \in \mathcal{X}_n \setminus \{x, y\}$ having a chain of $r$-open edges connecting $z$ to both $x$ and $y$.

**Definition 10** Let $\mathcal{E}$ be any $\sigma$-algebra of subsets of $E$ such that $\emptyset, E \in \mathcal{E}$, any $A \in \mathcal{E}$ implies $A^c \in \mathcal{E}$ and all countable unions of subsets of $\mathcal{E}$ is again in $\mathcal{E}$. Suppose $\{\eta_k\}_{k \geq 1}$ is a sequence of random variables on $E$ with values in $\mathbb{R}$. It will be said that $\eta_k$ converges weakly to a random
variable \( \eta : E \to \mathbb{R} \) (written \( \eta_k \Rightarrow \eta \)), if
\[
\lim_{k \to \infty} F_k(x) = \lim_{k \to \infty} P(\eta_k \leq x) = P(\eta \leq x) = F(x)
\]
for all \( x \in \mathbb{R} \).

### 3.5. The Event

#### 3.5.1. Bounded Number of Nodes

Let \( < C >_r \subseteq X_n \) be an \( r \)-connected component containing \( \hat{0} \) such that \( | < C >_r | = N \) and define \( \rho_n(C) = \frac{N}{n} \). For \( \rho \in (\frac{1}{2}, 1) \), define the graph property of all connected components containing at least \( 100\rho \% \) of all available points by

\[
\mathcal{A}^r_{[n, \rho]} = \{ < C >_r \subseteq X_n : \rho_n(C) \geq \rho \}.
\]

As in [23], for \( \epsilon \in (0, \frac{1}{2}) \), define

\[
r(n, \rho, \epsilon) = \inf\{ r > 0 : P(\mathcal{A}^r_{[n, \rho]} ) \geq \epsilon \}
\]

to be the critical radius at which \( \mathcal{A}^r_{[n, \rho]} \) occurs with probability at least \( \epsilon \) and define

\[
\Delta(n, \rho, \epsilon) = r(n, \rho, 1 - \epsilon) - r(n, \rho, \epsilon)
\]

to be the length of the continuum of radii upon which \( \mathcal{A}^r_{[n, \rho]} \) increases in probability of occurrence from \( \epsilon > 0 \) to \( 1 - \epsilon > 0 \).

#### 3.5.2. Unbounded Number of Nodes

In the case of \( n \) being unbounded, define the corresponding graph property to be

\[
\mathcal{A}^r = \{ < C >_r \subseteq X_\infty : | < C >_r | = \infty \}.
\]

Define

\[
r(\epsilon) = \inf\{ r > 0 : P(\mathcal{A}^r) \geq \epsilon \}
\]

to be the critical radius at which \( \mathcal{A}^r \) occurs with probability at least \( \epsilon \) and define

\[
\Delta(\epsilon) = r(1 - \epsilon) - r(\epsilon)
\]

to be the length of the continuum of radii upon which \( \mathcal{A}^r \) increases in probability of occurrence from \( \epsilon > 0 \) to \( 1 - \epsilon > 0 \).
### 3.6. Continuity Results

In order to prove the existence of \( r_0 > 0 \) such that \( P(\mathcal{A}_{[n,\rho]}^{r_0}) = 1/2 \), it will be shown that \( P(\mathcal{A}_{[n,\rho]}^r) \) is a continuous function of \( r \). By properties of probabilities measures, \( P(\mathcal{A}_{[n,\rho]}^r) \in [0,1] \) and by prop. (62), it is true that \( P(\mathcal{A}_{[n,\rho]}^r) \) is non-decreasing as a function of increasing \( r \) > 0. By thm. (7), it is true that \( P(\mathcal{A}_{[n,\rho]}^r) \) increases from \( \varepsilon > 0 \) to \( 1 - \varepsilon > 0 \) for fixed \( \varepsilon \in (0,\frac{1}{2}) \). Then, by continuity, there exists \( r_0 > 0 \) such that \( P(\mathcal{A}_{[n,\rho]}^{r_0}) = 1/2 \). If \( I \) is any continuum of radii and \( P(\mathcal{A}_{[n,\rho]}^r) \) is defined to be the set \( \{P(\mathcal{A}_{[n,\rho]}^r) : r \in I\} \), then it is easily seen that \( r_0 \) is in the interior of any compact interval of radii \( I_\varepsilon \) such that \( P(\mathcal{A}_{[n,\rho]}^{r_0}) = [\varepsilon,1-\varepsilon] \). Seeking a contradiction, suppose \( r_0 \) is in the boundary of \( I_\varepsilon \). Since \( I_\varepsilon \) is compact and \( P(\mathcal{A}_{[n,\rho]}^r) \) is continuous in \( r \), then \( P(\mathcal{A}_{[n,\rho]}^{r_0}) = \varepsilon \) or \( P(\mathcal{A}_{[n,\rho]}^{r_0}) = 1 - \varepsilon \). Therefore, \( P(\mathcal{A}_{[n,\rho]}^{r_0}) = 1/2 \) implies \( \varepsilon = 1/2 \). This is a contradiction since \( \varepsilon \in (0,\frac{1}{2}) \). Thus, \( r_0 \) is in the interior of \( I_\varepsilon \). Q.E.D.

Now, if it can be shown that \( r_0 \) is independent of \( \varepsilon \), then \( r_0 \in I_\varepsilon \) for all \( \varepsilon \in (0,\frac{1}{2}) \). Note that \( r_0 \in I = \bigcap_k I_{\varepsilon_k} \) for any sequence \( \varepsilon_k \to 1/2 \). Clearly \( I \) is compact so that \( r_0 \) is in the interior of \( I \). Therefore, either \( I \) is an interval or \( I = \{r_0\} \). Suppose \( I \) is an interval of radii. Since \( r_0 \) is in the interior of \( I \), then there exists \( r_0' < r_0 \in I \). Now, since \( \varepsilon_k \to 1/2 \), then \( P(\mathcal{A}_{[n,\rho]}^0) = 1/2 \) and \( r_0' < r_0 = \inf\{r > 0 : P(\mathcal{A}_{[n,\rho]}^r) = \frac{1}{2}\} \). This is a contradiction. Therefore, \( I = \{r_0\} \) so that \( r_0 \) is unique. Q.E.D.

Continuity of \( P(\mathcal{A}^r) \) in \( r \) is proven in [37] and can be used for proving continuity of \( P(\mathcal{A}_{[n,\rho]}^r) \) in \( r \) as follows. Let \( \partial B \) denote the boundary of \( B \) and define \( \mathcal{A}_B^* = \{0 \leftrightarrow \partial B\} \) to be the property that there is an \( r \)-connected cluster containing \( 0 \) and a point in \( \partial B \). By arguments in [37], continuity of \( P(\mathcal{A}^r) \) in \( r \) is equivalent to continuity of \( P(\mathcal{A}_B^*) \) in \( r \) for all bounded regions \( B \) containing \( 0 \). Clearly, \( P(\mathcal{A}_B^*) = P(\mathcal{A}_B^* - \mathcal{A}_{[n,\rho]}^r) + P(\mathcal{A}_B^* \cap \mathcal{A}_{[n,\rho]}^r) \) so that continuity of \( P(\mathcal{A}_B^*) \) in \( r \) implies continuity of \( P(\mathcal{A}_B^* \cap \mathcal{A}_{[n,\rho]}^r) \) in \( r \). Now, there exists \( r_0 > 0 \) such that \( P(\mathcal{A}_B^*) = 1 \) for all \( r \geq r_0 \). Then, it follows that \( P(\mathcal{A}_{[n,\rho]}^r) = P(\mathcal{A}_B^* \cap \mathcal{A}_{[n,\rho]}^r) \) is continuous when \( r \geq r_0 \). In particular, \( P(\mathcal{A}_{[n,\rho]}^r) \) is continuous at \( r_0 \). So, there is \( \delta > 0 \) such that \( P(\mathcal{A}_{[n,\rho]}^r) \) is continuous upon \( [r_0 - \delta, r_0 + \delta] \). Continuing this argument, continuity of \( P(\mathcal{A}_{[n,\rho]}^r) \) extends until \( r_0' - \delta = 0 \) so that \( P(\mathcal{A}_{[n,\rho]}^r) \) is continuous for all \( r \geq 0 \). Q.E.D.

**Theorem 11** [37, Thm. (3.8)] Suppose \( \{r_k\}_{k \geq 1} \) is a sequence of radii such that \( 0 < r_k \leq R \) for some \( R > 0 \) and \( \{\eta_k\}_{k \geq 1} \) is a sequence of random variables which take values \( r_k \) with probability \( 1 \). If \( 0 < r \leq R \) and \( \eta \) is a random variable taking the value \( r \) with probability \( 1 \) such that \( \eta_k \Rightarrow \eta \) as \( k \to \infty \). Then, \( P(\mathcal{A}^\eta) \to P(\mathcal{A}^r) \) as \( k \to \infty \).

**Proof** This is just a restatement of [37, Thm. (3.8)] for the special case of random variables \( \eta_k \) and \( \eta \) such that \( P(\eta_k = r_k) = 1 = P(\eta = r) \) for all \( k \geq 1 \).

**Corollary 12** (to Theorem 11) \( P(\mathcal{A}_{[n,\rho]}^r) \) is a continuous function of \( r \).

**Proof** Continuity of \( P(\mathcal{A}^r) \) in \( r \) follows from thm. (11). Therefore, the result follows by the discussion preceding the statement of thm. (11).

**Theorem 13** \( r = r(n,\rho,\epsilon) \) is a continuous function of \( \epsilon \) if and only if \( P(\mathcal{A}_{[n,\rho]}^r) \) is a continuous function of \( r \).

**Proof** Suppose \( r(n,\rho,\epsilon) \) is a continuous function of \( \epsilon \) and let \( \{\epsilon_k \in (0,\frac{1}{2})\}_{k \geq 1} \) be a sequence of positive real numbers such that \( \epsilon_k \to \epsilon_0 \) as \( k \to \infty \). Let \( \{X(\epsilon)\}_{\epsilon \in G(A_X, r)} \) be a finite sequence
of uniformly distributed random variables with values in $[0, 1]$ and define a sequence of random variables $\{\eta_k\}_{k \geq 1}$ by $\eta_k(e) = r(n, \rho, \epsilon_k) \equiv r_k$ when $X(e) < 1$ and 0 otherwise. Clearly, $\eta_k = r_k$ with probability 1 for all $k \geq 1$. Likewise, define a random variable $\eta_0$ by $\eta_0(e) = r(n, \rho, \epsilon_0) \equiv r_0$ when $X(e) < 1$ and 0 otherwise so that $\eta_0 = r_0$ with probability 1. Since $r(n, \rho, \epsilon)$ is continuous in $\epsilon$, then $r_k \to r_0$ as $k \to \infty$ so that $\eta_k \Rightarrow \eta_0$ as $k \to \infty$. Now, define $R = 2*\max\{d(x, y) : x, y \in X_n\}$.

By lemma (68), $0 < \eta_k \leq R$ for all $k \geq 0$. Therefore, $P(A_{n, \rho}^\epsilon) \to P(A_{n, \rho}^0)$ as $k \to \infty$ by cor. (12) since $r_k \to r_0$ as $k \to \infty$. Thus, $P(A_{n, \rho}^\epsilon)$ is a continuous function of $r$. Conversely, suppose $P(A_{n, \rho}^\epsilon)$ is a continuous function of $r$ and let $\{\epsilon_k \in (0, \frac{1}{2})\}_{k \geq 1}$ be any convergent sequence such that $\epsilon_k \to \epsilon_0$. Define $r_k = r(n, \rho, \epsilon_k)$ and $r_0 = r(n, \rho, \epsilon_0)$. Given $\xi > 0$, it is true that $\Xi \equiv \{k \geq 1 : |P(A_{n, \rho}^\epsilon) - P(A_{n, \rho}^0)| \geq \xi\}$ is a set of measure zero by the continuity assumption. Therefore, $r_k \to r_0$ as $k \to \infty$ by prop. (69). Thus, suppose that $\Xi \neq \emptyset$. Then, $\Xi$ is at most countable so that $\Xi = \emptyset$ a.s. Hence, $r_k \to r_0$ as $k \to \infty$ by prop. (69) so that $r(n, \rho, \epsilon)$ is a continuous function of $\epsilon$. \]

3.7. Continuum Giant Component

**Theorem 14** There exists $r_0 = r_0(n, \rho, \epsilon) < \infty$, independent of $\epsilon$, such that

$$P(A_{n, \rho}^0) = \frac{1}{2}.$$  

**Proof** Let $\epsilon \in (0, \frac{1}{2})$ be given. Since $A_{n, \rho}^\epsilon$ is an increasing property in $r$ by prop. (59), thm. (7) applies. Thus, there exists an interval $I_\epsilon$ of length $\Delta(n, \rho, \epsilon)$ such that $P(A_{n, \rho}^\epsilon) \in [\epsilon, 1-\epsilon]$ for $r \in I_\epsilon$. Since $P(A_{n, \rho}^\epsilon)$ is a continuous function of $r$ by cor. (12) and non-decreasing in $r$ by prop. (62) and $\frac{1}{2} \in [\epsilon, 1-\epsilon]$, then there exists $r_0 \in I_\epsilon$ such that $P(A_{n, \rho}^0) = 1/2$. If $R = 2*\max\{d(x, y) : x, y \in X_n\}$, then by lemma (68), it is true that $0 < r_0(n, \rho, \epsilon) \leq R < \infty$. It remains to be shown that $r_0 = r_0(n, \rho, \epsilon)$, independent of $\epsilon$.

Recall that $\rho \in (\frac{1}{4}, 1)$ and note that the maximum distance between any two connected points in $B$ is inversely proportional to $n$. Then, the particular $r_0$, which meets the requirements of thm. (14), is the exact radius, such that, it is equally probable that more than half of all points are connected contiguously, in which case, only one such cluster exists, with all other clusters being disjoint and sparsely connected throughout $B$, or all connected clusters disjointly contain half (or less than half) of all available points, in which case, more than one such cluster can exist. As such, $r_0$ demarcates the radial connection length at which the property $A_{n, \rho}$ undergoes a phase transition such that the graph $G(X_n; r)$ is likely to be sparsely connected and form disjoint connected classes of points $[37, Theorems (3.3), (3.6)]$, almost surely, when $r \in [0, r_0]$, while $G(X_n; r)$ is more likely to be fully connected and form one connected class of points $[37, Theorems (3.3), (3.6)]$, almost surely, when $r \in (r_0, 1]$.

**Lemma 15** $r_0 = r_0(n, \rho, \epsilon)$ is independent of $\epsilon$.

**Proof** Let $\epsilon_1, \epsilon_2 \in (0, \frac{1}{2})$ and suppose $r_{0,1} = r_0(n, \rho, \epsilon_1), r_{0,2} = r_0(n, \rho, \epsilon_2)$ such that

$$P(A_{n, \rho}^{r_{0,1}}) = \frac{1}{2} = P(A_{n, \rho}^{r_{0,2}}).$$  

(7)

It has to be shown that $r_{0,1} = r_{0,2}$. Let $\{\epsilon_k\}_{k \geq 1}$ be a sequence such that $\epsilon_k = \epsilon_1$ for all $k \geq 1$ and define $r_{0,k} = r_0(n, \rho, \epsilon_k)$. Then, for arbitrary $\xi > 0$, it is true that

$$\Xi \equiv \{k \geq 1 : |P(A_{n, \rho}^{r_{0,k}}) - P(A_{n, \rho}^{r_{0,2}})| \geq \xi\} = \emptyset$$  

(8)
since \( r_{0,k} = r_{0,1} \) for all \( k \geq 1 \). Hence, by prop. \((69)\), \( r_{0,k} \to r_{0,2} \) as \( k \to \infty \). But, \( r_{0,k} = r_{0,1} \) for all \( k \geq 1 \) so that \( r_{0,1} = r_{0,2} \). Thus, \( r_0 = r_0(n, \rho) \), independent of \( \epsilon \). 

**Remark 16** As a result of thm. \((15)\), \( r(\epsilon) \) is independent of \( \epsilon > 0 \), since \( r(n, \rho, \epsilon) \to r(\epsilon) \) as \( E[n] \to \infty \). As such, \( \Delta(\epsilon) = o(1) \) so that \( A^r \) has a sharp threshold, by definition \((6)\).

**Corollary 17** The critical radius, associated with the property \( A^r \), is unique.

**Proof** \( r(\epsilon) \) is the limit of \( r(n, \rho, \epsilon) \) as \( E[n] \to \infty \). As such, \( r_0 \) is the constant limit of \( r_0(n, \rho) \) as \( E[n] \to \infty \).

**Corollary 18** Given \( r > 0 \), there exists a density of points \( \lambda_0 = \lambda(n_0) \) such that

\[
P(A^r_{[n_0, \rho]} = \frac{1}{2}.
\]

**Proof** By lemma \((15)\), let \( n_0 = n_0(r, \rho) \) be the minimum of all positive (real) solutions to \( r = r_0(n, \rho) \) for some fixed \( r > 0 \). The result follows.

Since \( n \) is inversely proportional to connection distance \( r \) (requiring that \( n \in [1, \infty) \)), then the particular \( n_0 \), which meets the requirements of cor. \((18)\), is the exact number of points, such that, it is equally probable that more than half of all points are connected continuously. In this case, only one such cluster exists, with all other clusters being disjoint and sparsely connected throughout \( B \). Otherwise, all connected clusters disjointly contain half (or less than half) of all available points, in which case, more than one such cluster can exist. As such, \( n_0 \) demarcates the number of points at which the property \( A^r_{[n_0, \rho]} \) undergoes a phase transition such that the graph \( G(X_n; r) \) is likely to be sparsely connected and form disjoint connected classes of points \([37, \text{Theorems (3.3), (3.6)}]\), almost surely, when \( n \in [1, n_0] \). Alternatively, \( G(X_n; r) \) is more likely to be fully connected and form one connected class of points \([37, \text{Theorems (3.3), (3.6)}]\), almost surely, when \( n \in (n_0, \infty) \).

### 3.8. Continuum Sharp Threshold Interval Length

Given the particular radius guaranteed by thm. \((14)\), then thm. \((7)\) can be used to find an estimate of the length of the sharp threshold interval such that \( P(A^r_{[n_0, \rho]} \) increases sharply from some \( \epsilon \in (0, \frac{1}{2}) \) to \( 1 - \epsilon \). By lemma \((15)\), it is true that \( r_0 \) is independent of any particular \( \epsilon \). Thus, the interval and its length must be fixed, given \( n \) and \( \rho \in (\frac{1}{2}, 1) \).

**Theorem 19** \( \Delta(n, \rho) = \Theta(r_0 \log^{\frac{1}{2}} n) \).

**Proof** For \( \delta \in (0, \frac{1}{2}) \), let \( \epsilon_\delta = \frac{1}{2} - \delta \). By thm. \((7)\) and thms. \((14)\) and \((15)\),

\[
\Delta(n, \rho) = \lim_{\delta \to 0^+} \Delta(n, \rho, \epsilon_\delta) = \lim_{\delta \to 0^+} \Theta(r(n, \rho, \epsilon_\delta) \log^{\frac{1}{2}} n) = \Theta(r_0 \log^{\frac{1}{2}} n).
\]

Theorem \((19)\) gives an expected result, given thm. \((7)\) above. However, in \([10]\), a much more practical estimate of this length is obtained after the bounded region is partitioned by hexagons of a known size. If \( M \) is the number of these hexagons in the bounded region, then it is shown that a good estimate of the sharp interval length is a polynomial in \( 1/M \).
Theorem 20 There is a constant $c > 0$, independent of $M$, such that for all $\epsilon_1 > 0$ and every fixed small $\delta > 0$

$$P(A^r_{[n,\rho+\delta]}) \leq \left(\frac{1}{2} + \epsilon_1\right)M^{-c(r_0-r)}$$

(9)

for all $r \leq r_0$ and

$$P(A^r_{[n,\rho-\delta]}) \geq 1 - \left(\frac{1}{2} + \epsilon_1\right)M^{-c(r-r_0)}$$

(10)

for all $r \geq r_0$.

Theorem 21 $P(A^r_{[n,\rho]})$ is a continuous function of $\rho$.

Remark 22 The proof of thm. (21) requires thm. (20) which will be proven later. For now, the result of thm. (21) is assumed. By thm. (21), for small $\delta > 0$,

$$P(A^r_{[n,\rho-\delta]}) \approx P(A^r_{[n,\rho]}) \approx P(A^r_{[n,\rho+\delta]}).$$

Theorem (20) asserts that if $r_1 < r_0 < r_2$ and for some $\epsilon \in (0, \frac{1}{2})$, it is true that $P(A^{r_1}_{[n,\rho]}) = \epsilon$ and $P(A^{r_2}_{[n,\rho]}) = 1 - \epsilon$, then $r_2 - r_1$ is an estimate of the sharp threshold interval length for the property, $A^r_{[n,\rho]}$. Later, a similar result will be stated and proven which can be used in the proof of thm. (21).

4. Hexagonal Partition Model

It was seen in section (3.7) that $r_0 > 0$ exists such that the probability is 1/2 for the occurrence of the property that at least half of all points connect in the bounded region, $B$. By thm. (7),

$$r_c = r_c(n) = O\left(\sqrt{\frac{\log n}{n}}\right) \leq r_0(n) = r_0$$

(11)

where $r_c$ defines the critical radius at which the continuum property occurs with arbitrarily small, positive probability.

For fixed $r \in (r_c, r_0)$, let $h^r$ be the largest hexagon that can be inscribed into a circle of radius $r/4$. Let $H_r$ be a countably infinite collection of copies of $h^r$ such that

$$\mathbb{R}^2 = \bigcup_{h^r_{i,j} \in H_r} h^r_{i,j}$$

(12)

and for $h^r_{i,j}, h^r_{i',j'} \in H_r$, it is true that $h^r_{i,j} \neq h^r_{i',j'}$ whenever $|i - i'| + |j - j'| \neq 0$. Connectivity between $x, y \in X_n$ is then defined as $x$ and $y$ both lying in the same hexagon or neighboring hexagons.

With the bounded region $B$ partitioned into hexagons contained within $B \cap H_r$, the analysis proceeds whereby the original problem of estimating the sharp threshold interval length in the continuum is now replaced by the problem of estimating the length in the hexagonal partition framework. As such, definitions of connectivity and the increasing property are defined in the new framework. Then, the continuity and existence results are shown to still hold in the new framework. Later, an analogue to thm. (20) is stated and proven.
4.1. Definitions

Definition 23 A hexagonal partition of \( \mathcal{B} \) is a finite collection of hexagons from \( H_r \) such that \( \mathcal{B} \) is a union of all hexagons in the finite collection.

Definition 24 The Hamming distance between elements, \( h_{i,j}^r, h_{i',j'}^r \in H_r \) is defined to be the quantity
\[
h(h_{i,j}^r, h_{i',j'}^r) = |i - i'| + |j - j'|.
\]

Definition 25 Points \( x, y \in \mathcal{X}_n \) are \( H_r \)-connected and \( x, y >_{H_r} \) is an \( H_r \)-open edge, if there exists \( h_{i,x,j,y}^r, h_{i',x,j',y}^r \in H_r \) such that \( x \in h_{i,x,j,y}^r \) and \( y \in h_{i',x,j',y}^r \) where \( h(h_{i,x,j,y}^r, h_{i',x,j',y}^r) \leq 2 \) with \( |i_x - i_y| \leq 1 \) and \( |j_x - j_y| \leq 1 \). Points in \( \mathcal{X}_n \) are \( H_r \)-disconnected and form an \( H_r \)-closed edge otherwise.

Definition 26 Given a \( y \in \mathcal{X}_n \), an \( H_r \)-connected component containing \( y \) is the subset of points \( C_y >_{H_r} \subseteq \mathcal{X}_n \) containing \( y \) and every \( x \in \mathcal{X}_n \setminus \{y\} \) having an \( H_r \)-open set of edges connecting \( x \) to \( y \).

Definition 27 Given an \( H_r \)-connected edge, \( e = (x, y) >_{H_r} \), an \( H_r \)-connected component containing \( e \) is the subset of points \( C_e >_{H_r} \subseteq \mathcal{X}_n \) containing \( x \) and \( y \) and every \( z \in \mathcal{X}_n \setminus \{x, y\} \) having an \( H_r \)-open set of edges connecting \( z \) to both \( x \) and \( y \).

4.2. The Increasing Property

4.2.1. Bounded Number of Nodes

Let \( C >_{H_r} \subseteq \mathcal{X}_n \) be an \( H_r \)-connected component such that \( C >_{H_r} \) and let \( \rho_n(C) = \frac{N}{n} \) be defined as in section (3.5.1). For \( \rho \in \left( \frac{1}{2}, 1 \right) \), define the graph property of all connected components containing at least \( 100\rho\% \) of all available points by
\[
\mathcal{A}^{H_r}_{[n, \rho]} = \{ C >_{H_r} \subseteq \mathcal{X}_n : \rho_n(C) \geq \rho \}.
\] (13)

As in [23], for \( \epsilon \in (0, \frac{1}{2}) \), define
\[
r^\ast(n, \rho, \epsilon) = \inf\{ r > 0 : P(\mathcal{A}^{H_r}_{[n, \rho]}) \geq \epsilon \}
\] (14)

As the critical radius at which \( \mathcal{A}^{H_r}_{[n, \rho]} \) occurs with probability at least \( \epsilon \) and define
\[
\Delta^\ast(n, \rho, \epsilon) = r^\ast(n, \rho, 1 - \epsilon) - r^\ast(n, \rho, \epsilon)
\] (15)

As the length of the continuum of radii upon which \( \mathcal{A}^{H_r}_{[n, \rho]} \) increases in probability of occurrence from \( \epsilon > 0 \) to \( 1 - \epsilon > 0 \).

4.2.2. Unbounded Number of Nodes

In the event that \( n \) is unbounded, define the corresponding graph property to be
\[
\mathcal{A}^{H_r} = \{ C >_{H_r} \subseteq \mathcal{X}_\infty : |C >_{H_r}| = \infty \}.
\] (16)
Define
\[ r^*(\epsilon) = \inf\{ r > 0 : P(A_{H_r}) \geq \epsilon \} \] (17)
to be the critical radius at which \( A_{H_r} \) occurs with probability at least \( \epsilon \) and define
\[ \Delta^*(\epsilon) = r^*(1 - \epsilon) - r^*(\epsilon) \] (18)
to be the length of the continuum of radii upon which \( A_{H_r} \) increases in probability of occurrence from \( \epsilon > 0 \) to \( 1 - \epsilon > 0 \).

4.3. Continuity Results and Some Continuum Relationships

The continuity results of section (3.6) hold for the properties defined after the bounded region \( B \) is partitioned by copies of the hexagon \( h^r \), since connectivity is now characterized by points lying within distance \( r/2 \) (within neighboring hexagons). As such, the hexagonal partition connectivity model is only another way of viewing the continuum connectivity model. Then, by thm. (14), there exists \( r^*_0 = r^*_0(n, \rho) \) which satisfies the criteria of the theorem for the property \( A_{H_r}^{[n, \rho]} \).

**Definition 28** \( G(\mathcal{X}_n; H_r) \) is defined to be the \( H_r \)-graph of all \( H_r \)-open and \( H_r \)-closed edges between points in \( \mathcal{X}_n \subset B \).

**Lemma 29** \( G(\mathcal{X}_n; H_r) \subseteq G(\mathcal{X}_n; r) \).

**Proof** Suppose \( < x, y >_{H_r} \in G(\mathcal{X}_n; H_r) \) is any \( H_r \)-connected edge. Without loss of generality, choose a coordinate system on \( \mathbb{R}^2 \) so that \( x \neq y \) lies on a coordinate axis with \( 0 = (0,0) \) defined such that \( d(x,0) = \frac{d(x,y)}{2} = d(0,y) \). Since \( x,y \in \mathcal{X}_n \subset B \) and \( H_r \) is a partition of \( B \), then there exists \( h_{r_{x,jx}}, h_{r_{y,jy}} \in H_r \) such that \( x \in h_{r_{x,jx}} \), \( y \in h_{r_{y,jy}} \) and \( h(h_{r_{x,jx}}, h_{r_{y,jy}}) \leq \max\{|i_x - i_y|, |j_x - j_y|\} \leq 1 \). Each of \( h_{r_{x,jx}} \) and \( h_{r_{y,jy}} \) are copies of \( h^r \) and can be inscribed into copies of a circle of radius \( \frac{r}{4} \). Therefore, \( d(x, y) = d(x, \partial h_{r_{x,jx}}) + d(\partial h_{r_{y,jy}}, y) \leq \frac{r}{2} + \frac{r}{2} = r \) so that \( x,y \in \mathcal{X}_n \) are \( r \)-connected. Thus, \( < x, y >_{H_r} \in G(\mathcal{X}_n; r) \), which shows that \( G(\mathcal{X}_n; H_r) \subseteq G(\mathcal{X}_n; r) \).

**Lemma 30** \( P(A_{r_{[0, \rho]}}^{H_r}) \leq P(A_{r_{[0, \rho]}}^r) \).

**Proof** By lemma (29), it is true that \( A_{r_{[0, \rho]}}^{H_r} \subseteq A_{r_{[0, \rho]}}^r \).

**Lemma 31** \( r_0 \leq r^*_0 \).

**Proof** Seeking a contradiction, suppose \( r_0 > r^*_0 \). Then,
\[
\frac{1}{2} = P(A_{r_0}^{[n, \rho]}) \geq P(A_{r_0}^{[n, \rho]}) \geq P(A_{r_0}^{H_{r_0}}) = \frac{1}{2}
\]
where equality (19) follows by thm. (14), ineq. (20) follows by properties of probability measures and by hypothesis, ineq. (21) follows by lemma (30) and equality (22) follows by thm. (14). It follows that \( P(A_{r_0}^{H_{r_0}}) = 1/2 \). Therefore, \( r^*_0 \in \{ r > 0 : P(A_{r_0}^{[n, \rho]}) = \frac{1}{2} \} \) and \( r^*_0 < r_0 = \inf\{ r > 0 : P(A_{r_0}^{[n, \rho]}) = \frac{1}{2} \} \). This is a contradiction. Thus, \( r_0 \leq r^*_0 \).
4.4. Hexagonal Sharp Threshold Interval Length

Given the particular radius guaranteed by thm. (14), then thm. (7) can be used to find an estimate of the length of the sharp threshold interval such that \( P(A_{[n,\rho]}^H) \) increases sharply from some \( \epsilon \in (0, \frac{1}{2}) \) to \( 1 - \epsilon \). By lemma (15), it is true that \( r_0^* \) is independent of any particular \( \epsilon \). Thus, the interval and its length must be fixed given \( n \) and \( \rho \in (\frac{1}{2}, 1) \).

**Theorem 32** \( \Delta^*(n, \rho) = \Theta(r_0^* \log^{\frac{1}{4}} n) \).

**Proof** For \( \delta \in (0, \frac{1}{2}) \), let \( \epsilon_\delta = \frac{1}{2} - \delta \). By thm. (7) and thms. (14) and (15),

\[
\Delta^*(n, \rho) = \lim_{\delta \to 0^+} \Delta^*(n, \rho, \epsilon_\delta) = \lim_{\delta \to 0^+} \Theta(r_0^*(n, \rho, \epsilon_\delta) \log^{\frac{1}{4}} n) = \Theta(r_0^* \log^{\frac{1}{4}} n).
\]

As in thm. (19) above, thm. (32) gives an expected result, given thm. (7) above. Likewise, a similar result to [10, Thm. (3.3.1)] can be stated and later proven, as in the case of thm. (20).

**Theorem 33** There is a constant \( c > 0 \), independent of \( M \), such that for all \( \epsilon_1 > 0 \) and every fixed small \( \delta > 0 \)

\[
P(A_{[n,\rho+\delta]}^H) \leq \left( \frac{1}{2} + \epsilon_1 \right) M^{-c(r_0^*-r)}
\]

for all \( r \leq r_0^* \) and

\[
P(A_{[n,\rho-\delta]}^H) \geq 1 - \left( \frac{1}{2} + \epsilon_1 \right) M^{-c(r-r_0^*)}
\]

(23)

for all \( r \geq r_0^* \).

Let \( M^2 \) be the number of hexagons partitioning the region \( B \) and let \( H_B(r) = H_r \cap B \). Given \( < C >_{H_r} \subseteq \chi_n \), define \( H_C = \{ h_B \in H_B(r) : h_B \cap < C >_{H_r} \neq \emptyset \} \) to be the connected cluster of hexagons such that each hexagon contains at least one point from the connected cluster of points, \( < C >_{H_r} \).

**Lemma 34** \( E[\rho_n(C)] = \frac{E[|H_C|]}{M^2} \).

**Proof** Let \( < C >_{H_r} \subseteq \chi_n \) be an \( H_r \)-connected cluster and let \( K_{H_C} \) be a random variable taking as values the number of points in the region \( R_{H_C} \) defined by the hexagons in \( H_C \). Since the \( n \) points are uniformly distributed spatially and \( B \) is partitioned into \( M^2 \) copies of the prototypical hexagon \( h_r \), then

\[
E[K_{H_C}] = \frac{n}{\text{area}(B)} \frac{E[\text{area}(R_{H_C})]}{\text{area}(B)} = \frac{n}{M^2} \frac{E[|H_C|] \times \text{area}(h_r)}{\text{area}(h_r)} = \frac{n}{M^2} \frac{E[|H_C|]}{M^2}.
\]
But, \( E[ K_{Hr} ] = E[ | < C >_{Hr} | ] \). Therefore,

\[
E[ | < C >_{Hr} | ] = n \frac{E[ | H_C | ]}{M^2}
\]

implies

\[
E[ \rho_n(C) ] = \frac{E[ | H_C | ]}{M^2}.
\]

Define \( D^r_{[n,\rho]} = \{H_C \subseteq H_B(r) : E[ \rho_n(C) ] \geq \rho\} \). With \( D^r_{[n,\rho]} \) defined as such, the original problem of estimating the length of the sharp threshold for the property \( A^r_{[n,\rho]} \) in the continuum is now recast as a site percolation problem on a hexagonal lattice. As will be defined later, a site in the lattice will be deemed open if the corresponding hexagon is occupied by at least one of the points from \( X_n \) and it will be deemed closed otherwise. Likewise, two sites are connected and belong to the same connected cluster if both sites are open and their hamming distance is less than or equal to one. Later, a torus on the lattice will be formed by defining a countable collection of permutations of the hexagons in the partition so that the length of the sharp threshold for the property \( D^r_{[n,\rho]} \) can be approximated by the length for another property \( \hat{D}^r_{[n,\rho]} \) on the torus. In this way, boundary connection issues for sites in the partition of \( B \) are mitigated and the length of the sharp threshold interval for the property \( \hat{D}^r_{[n,\rho]} \) approximates the length for \( D^r_{[n,\rho]} \), which finally approximates the length for \( A^r_{[n,\rho]} \), the original property in the continuum.

**Theorem 35** There is a constant \( c > 0 \), independent of \( M \), such that

\[
P(D^r_{[n,\rho]}) \leq \frac{1}{2} M^{-c(r^*_0 - r)}
\]

for all \( r \leq r^*_0 \). Similarly, for some fixed small \( \delta > 0 \) and for all \( \epsilon_1 > 0 \), there is an \( M_0(\delta, \epsilon_1) \) such that for all \( M > M_0(\delta, \epsilon_1) \)

\[
P(D^r_{[n,\rho-\delta]}) \geq 1 - \left( \frac{1}{2} + \epsilon_1 \right) M^{-c(r - r^*_0)}
\]

for all \( r \geq r^*_0 \).

An important part of the proof of thm. (35) relies upon the sharp threshold inequality results of [9] and [18]. In order to apply these results, connectivity in the hexagon lattice structure should be extended to the case of a torus, whereby any boundary connectivity issues are mitigated. As such, make \( H_B(r) \) into a torus by identifying \( h_{i,j} \in H_B(r) \) with an element \( h_{i',j'} \) in a copy of \( H_B(r) \), if \( i' = i \mod M \) and \( j' = j \mod M \). For every \( k, l \in \mathbb{Z} \), the mapping \( \tau_{k,l} : h_{i,j} \rightarrow h_{i+k,j+l} \) defines a shift translation. In this way, a subgroup of automorphisms \( \tau = \{\tau_{k,l} : k, l \in \mathbb{Z}\} \) with the transitivity property is formed. Thus, any hexagon \( h_{i,j} \) can be shifted to any other hexagon \( h_{i',j'} \) with the translation, \( \tau_{i'-i,j'-j} \). Now, hexagons in the 1st row (column) are allowed to be joined in a connected cluster with hexagons in the Mth row (column), provided that all hexagons in question are occupied.

**Proposition 36** Define \( \tau(H_B(r)) \) to be the torus created by translations of hexagons in \( H_B(r) \) under the action of permutations in \( \tau \) and define \( \hat{D}^r_{[n,\rho]} = \{H_C \subseteq \tau(H_B(r)) : E[ \rho_n(C) ] \geq \rho\} \). Then, \( D^r_{[n,\rho]} \subset \hat{D}^r_{[n,\rho]} \) and \( D^r_{[n,\rho]} \neq \hat{D}^r_{[n,\rho]} \).
Proof Since $\mathcal{D}_{[n, \rho]}^r$ contains all of the connected hexagons from $\mathcal{D}_{[n, \rho]}^r$ and any connections between the 1st and Mth rows (columns) while $\mathcal{D}_{[n, \rho]}^r$ contains no connection between the 1st and Mth columns (rows), then the result follows.

Definition 37 To each hexagon in the partition of $\mathcal{B}$, associate a site $i \in \{1, 2, \ldots, M^2\}$ as the center of the hexagon. For sites $i \in \{1, 2, \ldots, M^2\}$, define $s_i \in \{0, 1\}$ to be the state on site $i$. A site $i$ is said to be open if $s_i = 1$ and closed otherwise. There exists an edge $e_{i, j}$ between sites $i, j \in \{1, 2, \ldots, M^2\}$ if and only if there exists a hexagon $h^r_{i,j} \ni i, j$ or there exists neighboring hexagons $h^r_i \ni i$ and $h^r_j \ni j$ in the partition of $\mathcal{B}$. Define $e_{i,j}$ to be open if and only if $s_i = 1 = s_j$ and closed otherwise.

Definition 38 The conditional influence of $i$ on the property $\mathcal{D}_{[n, \rho]}^r$ is defined to be

$$I(i) = P(\mathcal{D}_{[n, \rho]}^r | s_i = 1) - P(\mathcal{D}_{[n, \rho]}^r | s_i = 0)$$

and it is a measure of the change in the probability of $\mathcal{D}_{[n, \rho]}^r$ due to a state change from $s_i = 0$ to $s_i = 1$ at site, $i$.

For completeness, [10, Lemma (4.1.1)] is stated without proof, which gives an upper bound on the change in $P(\mathcal{D}_{[n, \rho]}^r)$ as a function of the point density $\lambda$. Utilizing the chain rule for derivatives, a lower bound on the change in $P(\mathcal{D}_{[n, \rho]}^r)$ as a function of $r$ is found and the resulting inequality relationship is used to estimate upper and lower bounds on $P(\mathcal{D}_{[n, \rho]}^r)$, which will approximate the inequality results of thm. (35).

Lemma 39 [10, Lemma (4.1.1)] There is a constant $z > 0$, independent of $M$ and $\lambda$, such that

$$\frac{d}{d\lambda} P(\mathcal{D}_{[n, \rho]}^r) \leq z^*(\lambda) \min\{P(\mathcal{D}_{[n, \rho]}^r), 1 - P(\mathcal{D}_{[n, \rho]}^r)\} \log M$$

where $A_{h^r}$ is the area of the prototypical hexagon $h^r$ and $z^*(\lambda) = -zA_{h^r} e^{-A_{h^r} \lambda}$.

Lemma 40 There is a constant $c > 0$, independent of $M$ and $\lambda$, such that

$$\frac{d}{dr} P(\mathcal{D}_{[n, \rho]}^r) \geq c^*(\lambda) \min\{P(\mathcal{D}_{[n, \rho]}^r), 1 - P(\mathcal{D}_{[n, \rho]}^r)\} \log M$$

where $A_{h^r}$ is the area of the prototypical hexagon $h^r$ and $c^*(\lambda) = c(\lambda)A_{h^r} e^{-A_{h^r} \lambda}$, with $c(\lambda) = -cg(\lambda)$ for some function $g(\lambda)$.

Proof As in cor. (18), let $n^*$ be the inverse of $r^*$ and seeking a contradiction, suppose $dr/d\lambda = 0$. Let $\epsilon \in (0, \frac{1}{2})$. By lemma (39), $dP/d\lambda$ exists. Now, the existence of $dP/dr$ will be shown by proving a Lipschitz condition on the probability distribution $P(\mathcal{D}_{[n, \rho]}^r)$ as a function of $r$. Assume area($\mathcal{B}$) = 1. Without loss of generality, it can be assumed that $r \in [0, 1]$. Without further loss of generality, let $r_1^*, r_2^* \in [0, 1]$ such that $r_0^*$ is the midpoint of $[r_1^*, r_2^*]$, i.e. $r_0^* = (r_2^* - r_1^*)/2$. Then, by thm. (32),

$$|P(\mathcal{D}_{[n, \rho]}^{r_2^*}) - P(\mathcal{D}_{[n, \rho]}^{r_1^*})| \leq 1 = (\Delta^*(n, \rho))^{-1}|r_2^* - r_1^*|.$$ 

Therefore, $P(\mathcal{D}_{[n, \rho]}^r)$ is Lipschitz continuous with respect to $r$. Hence, $dP/dr$ exists. Now, since $dP/d\lambda$, $dP/dr$ and $dr/d\lambda$ all exist, then the Chain Rule for derivatives yields,

$$\frac{d}{d\lambda} P(\mathcal{D}_{[n, \rho]}^r) = \frac{d}{dr} P(\mathcal{D}_{[n, \rho]}^r) \times \frac{dr}{d\lambda}.$$
Note that the existence of \( dP/dr \) requires that \(|dP/dr| < \infty\). Therefore, since \( dr/d\lambda = 0 \), then
\[
\frac{d}{d\lambda} P(\hat{D}_{[n,\rho]}^r) = \frac{d}{dr} P(\hat{D}_{[n,\rho]}^r) \times 0 = 0.
\]

As a result, \( P(\hat{D}_{[n,\rho]}^r) \) is constant as a function of \( \lambda \). So, suppose that \( 0 < n < n^* \). Then, \( P(\hat{D}_{[n,\rho]}^r) = 0 \), which implies that \( P(\hat{D}_{[n,\rho]}^r) \equiv 0 \). This is a contradiction, since \( P(\hat{D}_{[n,\rho]}^r) \) is a probability distribution. Hence, \( dr/d\lambda \neq 0 \). Now, by \([25, \text{Thm.}~(2.28)]\), there is a constant \( c > 0 \), independent of \( M \) and \( \lambda \), such that
\[
I(i) \geq c \min\{P(\hat{D}_{[n,\rho]}^r), 1 - P(\hat{D}_{[n,\rho]}^r)\} \frac{\log M}{M^2}.
\]

Under the action of \( \tau \), each hexagon in the bounded region \( B \) is translated to another hexagon in a copy of \( B \). Therefore, \( \hat{D}_{[n,\rho]}^r \) and \( P(\hat{D}_{[n,\rho]}^r) \) are invariant under the action of \( \tau \). Hence, \( I(i) = I(j) \) whenever, \( \tau(i) = j \), where \( \tau(i) \) is defined to be the translation of the hexagon \( h_i^r \) to the hexagon \( h_j^r \) in the copy of the partition of \( B \). From \([10]\), in the proof of thm. \((39)\), the following identity holds
\[
\frac{d}{d\lambda} P(\hat{D}_{[n,\rho]}^r) = \frac{d}{d\lambda} \left( e^{-A_{h_r^r}^r \lambda} \sum_{i=1}^{M^2} I(i) \right)
= -A_{h_r^r} e^{-A_{h_r^r}^r \lambda} \sum_{i=1}^{M^2} I(i).
\]

For \( \gamma > 0, r > 0 \) and \( k > 0 \), any \( H_r \)-connected component in \( \chi_n \) containing at least \( \gamma(n + k)/2 \) points will inherently contain an \( H_r \)-connected component of size at least \( \gamma n/2 \). Hence, \( A_{\gamma(n+k),\rho}^{H_r} \subseteq A_{\gamma n,\rho}^{H_r} \). It follows that \( P(A_{\gamma(n+k),\rho}^{H_r}) \leq P(A_{\gamma n,\rho}^{H_r}) \). Therefore, \( r^*(\gamma n, \rho, \epsilon) \in \{ r > 0 : P(A_{\gamma(n+k),\rho}^{H_r}) \geq \epsilon \} \), which implies \( r^*(\gamma(n + k), \rho, \epsilon) \leq r^*(\gamma n, \rho, \epsilon) \) for \( k > 0 \). Hence,
\[
r^*(\gamma(n + k), \rho, \epsilon) - r^*(\gamma n, \rho, \epsilon) \leq 0.
\]

Since point density \( \lambda \) is proportional to point count \( n \) for any bounded region \( B \), then using ineq. \((25)\) yields
\[
\frac{dr}{d\lambda} = \lim_{k \to 0} \frac{r^*(\gamma(n + k), \rho, \epsilon) - r^*(\gamma n, \rho, \epsilon)}{\gamma k} \leq 0,
\]
for some \( \gamma > 0 \). Since \( dr/d\lambda \neq 0 \), it follows that
\[
\frac{dr}{d\lambda} < 0.
\]

Since \( dr/d\lambda \) exists, then \(|dr/d\lambda| < \infty \). Thus, by substituting
\[
I(i) \geq c \min\{P(\hat{D}_{[n,\rho]}^r), 1 - P(\hat{D}_{[n,\rho]}^r)\} \frac{\log M}{M^2}.
\]
into (24), it follows that
\[
\frac{d}{d\lambda} P(\hat{D}_{[n, \rho]}^r) = -A_h r e^{-A_h r \lambda} \sum_{i=1}^{M^2} I(i)
\]
\[
\leq -c A_h r e^{-A_h r \lambda} \sum_{i=1}^{M^2} \min\{P(\hat{D}_{[n, \rho]}^r), 1 - P(\hat{D}_{[n, \rho]}^r)\} \log M
\]
\[
= -c A_h r e^{-A_h r \lambda} \min\{P(\hat{D}_{[n, \rho]}^r), 1 - P(\hat{D}_{[n, \rho]}^r)\} \log M.
\]
Therefore,
\[
\frac{d}{d\lambda} P(\hat{D}_{[n, \rho]}^r) = \frac{d}{dr} P(\hat{D}_{[n, \rho]}^r) \times \frac{dr}{d\lambda}
\]
\[
\leq -c A_h r e^{-A_h r \lambda} \min\{P(\hat{D}_{[n, \rho]}^r), 1 - P(\hat{D}_{[n, \rho]}^r)\} \log M
\] (26)
so that
\[
\frac{d}{dr} P(\hat{D}_{[n, \rho]}^r) \geq -c A_h r e^{-A_h r \lambda} \left(\frac{dr}{d\lambda}\right)^{-1} \min\{P(\hat{D}_{[n, \rho]}^r), 1 - P(\hat{D}_{[n, \rho]}^r)\} \log M.
\] (27)
Defining \(g(\lambda) = (dr/d\lambda)^{-1}\), the result follows.

**Remark 41** Let \(\epsilon > 0\) be given. At the risk of ambiguity, denote \(n = E[n]\) and define \(\lambda^*(n, \rho, \epsilon) = \inf\{n > 0 \mid P(\hat{D}_{[n, \rho]}^r) \geq \epsilon\}\). Inequality (26) implies that \(P(\hat{D}_{[n, \rho]}^r)\) is increasing as a function of decreasing node density \(\lambda = \lambda^*(n, \rho, \epsilon)\) such that the event \(\{P(\hat{D}_{[n, \rho]}^r) \geq \epsilon\}\) first occurs. Likewise, since the maximum distance between connected points is inversely proportional to node density, then ineq. (27) implies that \(P(\hat{D}_{[n, \rho]}^r)\) is decreasing as a function of increasing maximum distance \(r = r^*(n, \rho, \epsilon)\) between connected points such that the event \(\{P(\hat{D}_{[n, \rho]}^r) \geq \epsilon\}\) first occurs.

**Lemma 42** Let \(c > 0\) be as in thm. (40). Then, there exists \(r^*_0\), independent of \(M\), such that
\[
P(\hat{D}_{[n, \rho]}^r) \leq \frac{1}{2} M^{-c(r^*_0 - r)}
\]
for all \(r \leq r^*_0\) and
\[
P(\hat{D}_{[n, \rho]}^r) \geq 1 - \frac{1}{2} M^{-c(r - r^*_0)}
\]
for all \(r \geq r^*_0\).

**Proof** Arguing as in the proof to thm. (14), there exists \(r^*_0\) such that \(P(\hat{D}_{[n, \rho]}^{r^*_0}) = 1/2\). Arguing similarly to cor. (12), \(P(\hat{D}_{[n, \rho]}^r)\) is continuous in \(r\). Therefore, \(P(\hat{D}_{[n, \rho]}^r) \leq 1 - P(\hat{D}_{[n, \rho]}^r)\) for \(r \leq r^*_0\) and \(P(\hat{D}_{[n, \rho]}^r) \geq 1 - P(\hat{D}_{[n, \rho]}^r)\) for \(r \geq r^*_0\). Thus, the result of lemma (40) takes the form
\[
\frac{d}{dr} P(\hat{D}_{[n, \rho]}^r) \geq c^*(\lambda) P(\hat{D}_{[n, \rho]}^r) \log M
\]
for \(r \leq r^*_0\) and
\[
\frac{d}{dr} P(\hat{D}_{[n, \rho]}^r) \geq c^*(\lambda)(1 - P(\hat{D}_{[n, \rho]}^r)) \log M
\]
for \( r \geq r^*_0 \). The last two inequalities can be written
\[
\frac{d}{dr} \log P(\hat{D}_r^{[n,\rho]} \geq c^*(\lambda) \log M
\]
for \( r \leq r^*_0 \) and
\[
\frac{d}{dr} \log (1 - P(\hat{D}_r^{[n,\rho]})) \leq -c^*(\lambda) \log M
\]
for \( r \geq r^*_0 \), respectively. Consider \( r \leq r^*_0 \). Both sides of
\[
\frac{d}{dr} \log P(\hat{D}_r^{[n,\rho]} \geq c^*(\lambda) \log M
\]
are integrated in the direction of increasing point density since \( P(\hat{D}_r^{[n,\rho]} \geq c^*(\lambda) \log M
\) decreases as a function of point density \( \lambda \) by the proof to lemma (40). It was also shown that \( dr/d\lambda < 0 \), i.e. \( r \) is decreasing as a function of point density. Therefore, the integration limits for the interval \( [r, r^*_0] \) are from \( r^*_0 \) to \( r \). Noting that the inequality is reversed for backward integration, the following is obtained for \( c > 0 \) and some \( K_1(\lambda) \geq 0 \),
\[
\log P(\hat{D}_r^{[n,\rho]} \geq c^*(\lambda) \log M
\]
which can be rewritten as
\[
\log P(\hat{D}_r^{[n,\rho]} \leq K_1(\lambda) \log M^{-c(r^*_0 - r)}
\]
This implies
\[
P(\hat{D}_r^{[n,\rho]} \leq K_2(\lambda) M^{-c(r^*_0 - r)}
\]
for some \( K_2(\lambda) \geq 0 \). Therefore, using the initial condition \( P(\hat{D}_r^{[n,\rho]} = 1/2 \) yields \( K_2(\lambda) = 1/2 \). Thus,
\[
P(\hat{D}_r^{[n,\rho]} \leq \frac{1}{2} M^{-c(r^*_0 - r)}.
\]
Now, consider \( r \geq r^*_0 \). Similarly, both sides of
\[
\frac{d}{dr} \log (1 - P(\hat{D}_r^{[n,\rho]})) \leq -c^*(\lambda) \log M
\]
are integrated in the direction of increasing connection radii on \( [r^*_0, r] \) since \( P(\hat{D}_r^{[n,\rho]} \) increases as a function of connection radii \( r \) by the proof to lemma (40). Therefore, the integration limits are from \( r^*_0 \) to \( r \). The following is obtained for \( c > 0 \) and some \( K_3(\lambda) \geq 0 \),
\[
\log (1 - P(\hat{D}_r^{[n,\rho]})) \leq -K_3(\lambda) \log M^{-c(r^*_0 - r)}
\]
which can be rewritten as
\[
\log (1 - P(\hat{D}_r^{[n,\rho]})) \leq -K_3(\lambda) \log M^{-c(r^*_0 - r)} = K_3(\lambda) \log M^{-c(r^*_0 - r)}.
\]
This implies

$$1 - P(\hat{D}_{[n,\rho]}^r) \leq K_4(\lambda)M^{-c(r-r_0^*)}$$

for some $K_4(\lambda) \geq 0$. Therefore, using the initial condition $P(\hat{D}_{[n,\rho]}^r) = 1/2$ yields $K_4(\lambda) = 1/2$. Hence,

$$P(\hat{D}_{[n,\rho]}^r) \geq 1 - \frac{1}{2}M^{-c(r-r_0^*)}.$$  

By prop. (36), there are cases when $\mathcal{D}_{[n,\rho]}^r \subset \hat{\mathcal{D}}_{[n,\rho]}^r$, but $\mathcal{D}_{[n,\rho]}^r \neq \hat{\mathcal{D}}_{[n,\rho]}^r$ so that the occurrence of $\hat{\mathcal{D}}_{[n,\rho]}^r$ does not imply the occurrence of $\mathcal{D}_{[n,\rho]}^r$. To exclude these possibilities, the arguments of [10] are followed whereby a slightly larger property $\mathcal{D}_{[n,\rho-\delta]}^r$ is considered for some small $\delta > 0$ such that the occurrence of $\hat{\mathcal{D}}_{[n,\rho]}^r$ implies the occurrence of $\mathcal{D}_{[n,\rho-\delta]}^r$.

As in [10], let $\phi(M)$ be any $M$-dependent integer such that $\phi(M) \to \infty$ as $M \to \infty$ and

$$\phi(M) = o(c(r-r_0^*) \log M).$$

Choose a coordinate system so that $\mathcal{B}$ has its lower left corner at the origin. Define the top, bottom, left and right boundary strips of $\mathcal{B}$ as $H_i, i = 1, 2, 3, 4$ with sizes $\phi(M) \times M$, $\phi(M) \times M$, $M \times \phi(M)$ and $M \times \phi(M)$ by

$$H_1 = \{H_{i,j} : i = M - \phi(M) + 1, ..., M, j = 1, ..., M\}$$

$$H_2 = \{H_{i,j} : i = 1, ..., \phi(M), j = 1, ..., M\}$$

$$H_3 = \{H_{i,j} : i = 1, ..., M, j = 1, ..., \phi(M)\}$$

$$H_4 = \{H_{i,j} : i = 1, ..., M, j = M - \phi(M) + 1, ..., M\}.$$ 

Let $E_i$ be the event that there is a connected path of occupied hexagons crossing $H_i$, long way.

**Lemma 43** For $i = 1, 2, 3, 4$, there are constants $c_i > 0$ such that for large $M$ and $r \geq r_0^*$

$$P(E_i) \geq 1 - e^{-c_i \phi(M)}.$$ 

**Proof** As in [10], by the duality property, the occurrence of $E_i, i = 1, 2, 3, 4$ is equivalent to the non-occurrence of the event that there is a connected path of unoccupied hexagons crossing $H_i, i = 1, 2, 3, 4$, short way. The rest of the proof follows [10] with the edge probability as a function of point density $p(\lambda_0)$ replaced by $r_0^*$ and the critical probability for the occurrence of an infinite cluster of occupied hexagons $p_0$ replaced by $r^*(n, \rho, \epsilon)$.

**Proof** (Theorem 35) By prop. (36), $\mathcal{D}_{[n,\rho]}^r \subset \hat{\mathcal{D}}_{[n,\rho]}^r$ so that $P(\hat{\mathcal{D}}_{[n,\rho]}^r) \leq P(\mathcal{D}_{[n,\rho]}^r)$. To estimate $P(\mathcal{D}_{[n,\rho-\delta]}^r)$ for $r > r_0$ and any given $\delta > 0$, let $E = E_1 \cap E_2 \cap E_3 \cap E_4$ and consider $F = \hat{\mathcal{D}}_{[n,\rho]}^r \cap E$.

Since $P(F) = P(F \cap \mathcal{D}_{[n,\rho-\delta]}^r) + P(F - \mathcal{D}_{[n,\rho-\delta]}^r)$, then

$$P(\mathcal{D}_{[n,\rho-\delta]}^r) \geq P(F) - P(F - \mathcal{D}_{[n,\rho-\delta]}^r).$$
Noting that \( P(E_1) = P(E_2) \) and \( P(E_3) = P(E_4) \), then the FKG inequality of [24] yields
\[
P(F) \geq P(D^r_{(n, \rho)}) P^2(E_1) P^2(E_3).
\]
By lemma (43), there exists \( b > 0 \) such that for all sufficiently large \( M \),
\[
P(F) \geq 1 - \frac{1}{2} M^{-c(r-r_0^*)} - O(e^{-b\phi(M)}).
\]
Using \( \phi(M) = o(c(r-r_0^*) \log M) \), this implies that for any given \( \epsilon_1 > 0 \) and all sufficiently large \( M \) depending upon \( \epsilon_1 \),
\[
P(F) \geq 1 - \left( \frac{1}{2} + \epsilon_1 \right) M^{-c(r-r_0^*)}.
\]
It is now claimed that \( F - D^r_{(n, \rho - \delta)} = \emptyset \), requiring that \( P(F - D^r_{(n, \rho - \delta)}) = 0 \) for all large \( M \). Following [10], the occurrence of \( F \) implies that there is a connected path of hexagons which encloses the sub-lattice given by \( H_B(r) - \bigcup_{i=1}^{4} H_i \). Because the points in \( X_n \) are uniformly distributed, then there is a connected cluster of hexagons within the original lattice totaling at least \( \rho M^2 - (2M\phi(M) + 2\phi(M)(M - 2\phi(M))) \) hexagons, where \( \rho M^2 \) is a lower bound on the number of occupied hexagons in the largest connected cluster and \( 2M\phi(M) + 2\phi(M)(M - 2\phi(M)) \) is the total number of hexagons in the strips, \( H_i, i = 1, 2, 3, 4 \). Let \( \delta_1 = (2M\phi(M) + 2\phi(M)(M - 2\phi(M)))/M^2 \). It follows that \( F \subset D^r_{(n, \rho - \delta_1)} \), since \( F \) occurs in those hexagons of \( B \) that are not near the boundary of \( B \) by a simple translation \( \tau \) of hexagons \( h \in \bigcup_{i=1}^{4} H_i \) to hexagons \( h \in H_B(r) - \bigcup_{i=1}^{4} H_i \). Thus, if \( M \) is large enough so that \( \delta_1 < \delta \), then \( F \subset D^r_{(n, \rho - \delta_1)} \subset D^r_{(n, \rho - \delta)} \).

**Proof (Theorem 33)** Consider \( r \leq r_0^* \). Since
\[
P(A^r_{(n, \rho + \delta)}) = P(A^r_{(n, \rho + \delta)}, D^r_{(n, \rho)}) + P(A^r_{(n, \rho + \delta)} - D^r_{(n, \rho)})
\]
then
\[
P(A^r_{(n, \rho + \delta)}) \leq P(D^r_{(n, \rho)}) + P(A^r_{(n, \rho + \delta)} - D^r_{(n, \rho)}).
\]
It will be shown that \( P(A^r_{(n, \rho + \delta)} - D^r_{(n, \rho)}) = o(M^{-c(r_0^* - r)}) \). Let \( x \) be a configuration of states across hexagons in \( H_B(r) \) and let \( \mathcal{C}(x) = \{C_1, ..., C_K\} \) be the set of clusters in \( x \). For \( i = 1, ..., K \), let \( N_{C_i} \) be the number of points in the cluster, \( C_i \). Then, \( \{N_{C_i} \mid \mathcal{C}(x), n\} \sim B(n, |H_{C_i}|/M^2) \). Suppose \( C_{i_0} \in \mathcal{C}(x) \) is any cluster such that \( \rho_n(C_{i_0}) \geq \rho + \delta \). Since the occurrence of the property \( (D^r_{(n, \rho)})^c \) implies \( |H_{C_{i_0}}| < \rho \), then
\[
A^r_{(n, \rho + \delta)} - D^r_{(n, \rho)} \subset \left\{ \rho_n(C_{i_0}) \geq \rho + \delta, \frac{|H_{C_{i_0}}|}{M^2} < \rho \right\}.
\]
By arguments in [15] and [10], there is an \( \alpha = \alpha(\rho, \delta) > 0 \) such that
\[
P \left( \rho_n(C_{i_0}) \geq \rho + \delta \left\{ \frac{|H_{C_{i_0}}|}{M^2} < \rho \right\}, \mathcal{C}(x), n \right) \leq e^{-\alpha(\rho, \delta)n}.
\]
It follows that
\begin{align}
P(A_{[n,\rho+\delta]}^H - D_{[n,\rho]}) & \leq P \left( \{ \rho_n(C_{i_0}) \geq \rho + \delta \}, \left\{ \frac{|H_{C_{i_0}}|}{M^2} < \rho \right\} \right) \\
& = P \left( \rho_n(C_{i_0}) \geq \rho + \delta \left| \frac{|H_{C_{i_0}}|}{M^2} < \rho \right\} \right) \times P \left( \frac{|H_{C_{i_0}}|}{M^2} < \rho \right) \\
& \leq P \left( \rho_n(C_{i_0}) \geq \rho + \delta \left| \frac{|H_{C_{i_0}}|}{M^2} < \rho \right\} \right) \\
& = E \left[ \rho_n(C_{i_0}) \geq \rho + \delta \left| \left\{ \frac{|H_{C_{i_0}}|}{M^2} < \rho \right\} \right. , C(x), n \right] \\
& \leq E [\rho_n(1 - e^{-\alpha})] \\
& = \exp \{- n(1 - e^{-\alpha})\}
\end{align}
where eq. (28) follows since $P \left( \frac{|H_{C_{i_0}}|}{M^2} < \rho \right) \leq 1$ and eq. (29) follows since $E [e^{-\alpha n}]$ is simply the moment generating function, $\exp \{- n(1 - e^{-\alpha})\}$.

Now, since $n(1 - e^{-\alpha}) > d \log M$ implies $\exp \{- n(1 - e^{-\alpha})\} < M^{-d}$, then for any $d > 0$ and every fixed $\delta > 0$, it follows that $P(A_{[n,\rho+\delta]}^H - D_{[n,\rho]})$ decays to zero at a rate faster than $M^{-d}$ for $n$ large enough. The case of $r \geq r_0^*$ is proven with similar arguments.

**Theorem 44** $P(A_{[n,\rho]}^H)$ is a continuous function of $\rho$.

**Proof** Let $\sigma = 1 - \rho$ in eq. (13). Then, $A_{[n,\rho]}^H$ is an increasing property in $\sigma$ for increasing $\rho \in (\frac{1}{2}, 1)$. Therefore, by [25, Thm. (2.48)], it is true that $A_{[n,\rho]}^H$ has a sharp threshold in $\sigma$, and hence, in $\rho$. Thus, by [25, Ineq. (2.49)], $P(A_{[n,\rho]}^H)$ is differentiable in $\rho$, which implies that $P(A_{[n,\rho]}^H)$ is continuous as a function of $\rho$.

**Remark 45** By thm. (44), for small $\delta > 0$,
\[
P(A_{[n,\rho-\delta]}^H) \approx P(A_{[n,\rho]}^H) \approx P(A_{[n,\rho+\delta]}^H).
\]
In this light, thm. (33) asserts that if $r_1^* < r_0^* < r_2^*$ and for some $\epsilon \in (0, \frac{1}{2})$, it is true that $P(A_{[n,\rho]}^{H_1}) = \epsilon$ and $P(A_{[n,\rho]}^{H_2}) = 1 - \epsilon$, then $r_2^* - r_1^*$ is an estimate of the sharp threshold interval length for the property, $A_{[n,\rho]}^H$.

**Proof** (Theorem 20) Since $P(A_{[n,\rho]}^r)$ and $P(A_{[n,\rho]}^H)$ are continuous functions of $r$, then by thm. (33) and lemma (31), for every $r \in [0, r_0]$ there exists $r' \leq r$ such that
\begin{align}
P(A_{[n,\rho+\delta]}^r) & \leq P(A_{[n,\rho+\delta]}^H) \\
& \leq \left( \frac{1}{2} + \epsilon_1 \right) M^{-c(r_0^*-r)} \\
& \leq \left( \frac{1}{2} + \epsilon_1 \right) M^{-c(r_0-r)}.
\end{align}
Consider $r_0 \in [0, r_0]$. Then, continuity of $P(A_{[n,\rho]}^r)$ in $r$ and the non-decreasing property of $P(A_{[n,\rho]}^r)$ in $r$ implies ineq. (31) for all $r \in [0, r']$. It is claimed that $r' = r_0$. Seeking a contradiction
if \( r' < r_0 \), suppose \( P(\mathcal{A}_r^{\rho}) \leq (\frac{1}{2} + \epsilon_1)M^{-c(r_0 - r)} \) for all \( r \in [0, r'] \) and \( P(\mathcal{A}_r^{\rho}) > (\frac{1}{2} + \epsilon_1)M^{-c(r_0 - r)} \) for all \( r > r' \). By hypothesis, \( r_0 > r' \) so that when \( r = r_0 \), it follows that \( P(\mathcal{A}_r^{\rho}) > 1/2 \). Now, since for any connected cluster \( < C > \), such that \( \rho_n(C) \geq \rho + \delta \) for \( \delta > 0 \), the statement \( \rho_n(C) \geq \rho \) is implied, then \( \mathcal{A}_r^{\rho} \subseteq \mathcal{A}_r^{\rho} \) for all \( r \in [0, r_0] \). Hence, \( r' < r_0 \) leads to

\[
P(\mathcal{A}_r^{\rho}) \geq \limsup_{\delta \to 0^+} P(\mathcal{A}_r^{\rho}) \geq P(\mathcal{A}_r^{\rho}) > \frac{1}{2}.
\]

In particular, ineq. (32) gives \( P(\mathcal{A}_r^{\rho}) > 1/2 \). This is a contradiction since \( P(A_r^{\rho}) = 1/2 \) by thm. (14). It follows that \( r' = r_0 \) and

\[
P(\mathcal{A}_r^{\rho}) \leq \left( \frac{1}{2} + \epsilon_1 \right)M^{-c(r_0 - r)}
\]

for \( r \geq r_0 \).

The implication of the proof to thm. (20) is that \( P(\mathcal{A}_r^{\rho}) = P(A_r^{H_{\rho}}) \) for \( r \in [0, r_0] \). By [25, Thm. (1.16)], the random cluster measure gives rise to a collection of conditional probability measures of connection events in the identified classes during K-means classification. Therefore, the node process \( X \) samples from each element of the collection.

**Theorem 46** \( P(\mathcal{A}_r^{\rho}) = P(A_r^{H_{\rho}}) \) for \( r \in [0, r_0] \).

**Proof** By continuity in \( \rho \) of \( P(A_r^{H_{\rho}}) \) as given by thm. (44), it is true that

\[
\lim_{\delta \to 0^+} P(\mathcal{A}_r^{H_{\rho + \delta}}) = P(\mathcal{A}_r^{H_{\rho}}).
\]

Suppose \( \delta_1 > \delta_2 \) such that \( \rho + \delta_1, \rho + \delta_2 \in (\frac{1}{2}, 1) \) and let \( < C >, r \in \mathcal{A}_r^{\rho + \delta_1} \). Then, \( \rho_n(C) \geq \rho + \delta_1 > \rho + \delta_2 \) so that \( < C >, r \in \mathcal{A}_r^{\rho + \delta_1} \). Hence, \( \mathcal{A}_r^{\rho + \delta_1} \subseteq \mathcal{A}_r^{\rho + \delta_2} \). By properties of probability measures, \( P(A_r^{H_{\rho + \delta}}) \) is monotone non-decreasing as a function of decreasing \( \rho \). By ineq. (30), it follows that for some fixed \( r \in [0, r_0] \), there exists \( r' \in [0, r_0] \) such that \( P(A_r^{\rho + \delta}) \leq P(A_r^{H_{\rho + \delta}}) \) for all \( r'' \in [0, r'] \) so that

\[
\limsup_{\delta \to 0^+} P(\mathcal{A}_r^{\rho + \delta}) \leq \limsup_{\delta \to 0^+} P(A_r^{H_{\rho + \delta}}) = P(A_r^{H_{\rho}}).
\]

From the proof of thm. (20), it was shown that \( r' = r_0 \). Therefore, by continuity of \( P(A_r^{H_{\rho}}) \) in \( r \), ineq. (33) holds for all \( r \in [0, r_0] \), with \( r'' \) replaced by \( r \). The Monotone Convergence Theorem [44] applied to \( E[1 \mathcal{A}_r^{\rho}] \) and \( E[1 \mathcal{A}_r^{H_{\rho}}] \) guarantees that \( P(A_r^{\rho}) \to P(A_r^{H_{\rho}}) \) as \( \delta \to 0^+ \). Therefore, ineq. (33) becomes

\[
P(\mathcal{A}_r^{\rho}) = \limsup_{\delta \to 0^+} P(\mathcal{A}_r^{\rho + \delta}) \leq P(A_r^{H_{\rho}}).
\]

In particular, \( P(A_r^{\rho}) \leq P(A_r^{H_{\rho}}) \) so that with the result of lemma (30), namely \( P(A_r^{H_{\rho}}) \leq P(A_r^{\rho}) \), the theorem follows.
Corollary 47 $P(A^r) = P(A^{Hr})$ for $r \in [0, r_0]$.

Proof By thm. (46), it is true that $P(A^r_{[n, \rho]}) = P(A^{Hr}_{[n, \rho]})$ for all $r \in [0, r_0]$ and all $n \geq 1$. By prop. (63), it follows that $P(A^{Hr}) \leq P(A^{Hr}_{[n, \rho]}) = P(A^r_{[n, \rho]})$. In particular, $P(A^{Hr}) \leq P(A^r_{[n, \rho]})$. Without loss of generality, assume that area($B$) = 1. From [10], differentiability of $P(A^r_{[n, \rho]})$ in point density $\lambda = \lambda(n) = E[n]$ implies continuity of $P(A^r_{[n, \rho]})$ in $\lambda$ so that the following holds

$$\lim_{E[n] \to \infty} P(A^r_{[n, \rho]}) = P(A^r).$$

Therefore, $P(A^{Hr}) \leq P(A^r_{[n, \rho]})$ and eq. (35) implies $P(A^{Hr}) \leq P(A^r)$. Similarly, $P(A^r) \leq P(A^{Hr})$ so that the corollary follows.

Corollary 48 $r_0 = r_0^*$.

Proof By thm. (46), it is true that $1/2 = P(A^r_{[n, \rho]}) = P(A^{Hr}_{[n, \rho]})$. In particular, $1/2 = P(A^{Hr}_{[n, \rho]})$. Since $P(A^r_{[n, \rho]}) = 1/2 = P(A^{Hr}_{[n, \rho]})$, by the discussion preceding thm. (32) and by thm. (14), then the uniqueness of $r_0^*$ and $r_0$ guarantees that $r_0^* = r_0$.

Proof (Theorem 21) Follows directly from thms. (44) and (46).

By thm. (46) and cor. (48), the problem of estimating the probabilities and length of the sharp threshold interval in the continuum can be re-cast as problems of estimation in the presence of a hexagonal partition of the bounded region. As such, tools from percolation [24] and the random cluster model [25] can readily be employed. This fact will be of paramount importance in applications to $K$-means classification where a data set consisting of multi-dimensional points is partitioned into disjoint, connected subsets. As it is advantageous to not have one connected cluster containing at least $100\%$ of all points, since otherwise there may exist a single cluster containing almost all points by lemma (58), the connection radius for points in the continuum must be in the sub-critical range $r \in [0, r_0]$ when classifying data into more than 2 classes. Since $P(A^r_{[n, \rho]}) = P(A^{Hr}_{[n, \rho]})$ for $r \in [0, r_0]$, disjoint clusters of points in the continuum are equivalent to disjoint clusters of occupied hexagons in the hexagonal partition of the bounded region containing all points. As such, multi-dimensional points in the continuum can be thought to belong to the same class if they are within a certain Euclidean distance of one another. As a result, the multi-dimensional points will have representatives belonging to occupied, connected hexagons in the 2-dimensional, bounded, partitioned region. All representatives in connected clusters of hexagons form the members of a class.

5. $K$-Means Classification - A Practical Example

Suppose $n = M^2$. The idea is to partition $B$ into $M^2$ hexagons and find $K = N^2$ contiguous clusters of hexagons such that each of the clusters are mutually disjoint. Into one and only one hexagon of a given cluster will each data point be mapped to form a point in the connected cluster. As such, the connected clusters of hexagons will be the $K = N^2$ classes containing a representative point associated to one and only one data point. We have the following theorem from Murphy, [40, Theorem (1)].

Theorem 49 Assume that there are $M^2$ points and $N^2$ classifications for the points. The minimum number of hexagons required to partition the unit square into $N^2$ disjoint regions such that
$M^2$ is the sum total of all hexagons in the disjoint regions is given by

$$S(M, N) = M^2 + (N - 1)^2 + 2MN.$$  

The idea is to use the result of the theorem to calculate, as a function of $M$ and $N = N(M)$, the exact size of a prototypical hexagon which will be used to partition $B$ into hexagons of equal size. As $K = N^2$ is fixed as the number of classes of data points, $M^2$ is fixed for the initial calculation of $S(M, N)$ and the subsequent classification of the first $M^2$ data points. In [24], it is stated and proven that there is a critical probability of connection between hexagons containing a point of a network such that it is no longer possible to have disjoint clusters of points when this critical probability of connection is exceeded. Hence, all points will be connected into one cluster, which is not what we intend to model, in this case. Since the size of $B$ is fixed, then to decrease the probability of connection while maintaining $K = N^2$ disjoint contiguous clusters of points, the size of each hexagon must decrease while increasing the number of hexagons in the boundaries of the disjoint regions. In this way, the ratio of the total number of occupied hexagons to the total number of hexagons will be less than this critical probability of connection. Note that we used uniformity of the points throughout $B$ so that the approximate number of points in a cluster of hexagons is proportional to the ratio of the number of hexagons in the cluster divided by the number of hexagons in the entire region, $B$. Also, note that the minimum number of hexagons required for separation is given by thm. (49), so that the common radius of the circle that can circumscribe any one of these hexagons is of size

$$R(M, N) = \frac{1}{2\sqrt{S(M, N)}},$$

thereby necessarily indicating that

$$B(M, N) = 2 \cdot R(M, N)$$

is the diameter of the circumscribing circle. We have the following lemma from Murphy, [40, Lemma (2)].

**Lemma 50** $R(M, N)$ is decreasing for increasing $M$ and $N$.

**Theorem 51** Suppose that the node process $X$ generates infinitely many points in $\mathbb{R}^2$. An infinite connected cluster exists across hexagons in $\mathbb{R}^2$ with probability 1 if and only if the probability that any two points connect exceeds $p_c$, where $p_c$ is the critical probability of connection. Otherwise, all connected clusters are disjoint with probability 1.

Theorem (51) is a restatement of [25, Thm. (1.11)]. A direct result of thm. (51) is that, given any bounded region $B$, all points generated within $B$ are almost surely connected into one cluster. Therefore, in order to not exceed the critical probability of connection, which means maintaining the $N^2$ classes of $M^2$ data points, the radial length of each hexagon’s circumscribing circle must be less than or equal to $R(M, N)$. By [24, Thm. (1.11)], the clusters will be disjoint with probability 1. Hence, the following corollary to thm. (49) follows from these statements and lemma (53).

**Corollary 52** Let $h^r$ be a hexagon of size such that it can be inscribed into a circle of radius $r = r(M, N) > 0$ where

$$0 < r \leq R(M, N).$$

If $B$ is partitioned into copies of $h^r$, then with probability 1, $N^2$ is the mean number of disjoint clusters of contiguous hexagons in the region $B$ that are occupied by the $M^2$ points.
With $r_0$ given by cor. (52), the size of the prototypical hexagon can be calculated for repartitioning $\mathcal{B}$. Furthermore, cor. (52) guarantees that the classes will remain distinct, with probability 1, through each new classification. By cor. (52), the expected value of the number of classes to form can be calculated. We have the following lemma from Murphy, [40, Lemma (5)].

**Lemma 53** For $M^2$ uniformly distributed data points in $\mathcal{B}$ and for any $\rho \in (0, p_c]$, with $p_c = 1 - 2 \sin (\pi/18)$,

$$\frac{M^2}{S(M, N)} = \frac{M^2}{M^2 + (N - 1)^2 + 2MN} = \rho$$

(37)

determines the expected number $K = N^2$ of disjoint classes to form such that $M^2$ is the total of all occupied hexagons across all classes.

**Lemma 54** For fixed $\rho \in (\frac{1}{2}, 1)$ and $r > 0$ there exists $\delta = \delta(\rho) \in (0, \frac{1}{2})$, such that

$$\left\{ \frac{|C|}{S(M, N)} < \frac{1}{2} \right\} = \left( A_{[S(M, N), \rho-\delta]}^H \right)^c$$

upto sets of $P$-measure zero.

**Proof** By definition, $\left( A_{[S(M, N), \rho-\delta]}^H \right)^c = \left\{ \frac{|C|}{S(M, N)} < \rho - \delta \right\}$. Take $\delta = \rho - \frac{1}{2}$.

By choosing $\delta$ as in lemma (54), continuity in $r > 0$ and the non-decreasing property of $P\left( A_{[S(M, N), \rho-\delta]}^H \right)$ for increasing $r > 0$ granted by cor. (12) and prop. (62), respectively, then by ineq. (10), it follows that

$$R(M, N) < r_0^* = r_0^*(M, N)$$

for the property $\left( A_{[S(M, N), \rho-\delta]}^H \right)^c$, since

$$P\left( \left( A_{[S(M, N), \rho-\delta]}^H \right)^c \right) = 1$$

$$> \frac{1}{2}$$

$$= P\left( \left( A_{[S(M, N), \rho-\delta]}^H \right)^c \right)$$

and the probability of $\left( A_{[S(M, N), \rho-\delta]}^H \right)^c$ is non-decreasing for decreasing $r \leq r_0^*$, a reversal.

Let $\epsilon \in (0, \frac{1}{2})$ be given and let $r_1^* > 0$ and $r_2^* > 0$, guaranteed by cor. (12), be such that

$$P\left( \left( A_{[S(M, N), \rho-\delta]}^H \right)^c \right) = 1 - \epsilon$$

and $P\left( \left( A_{[S(M, N), \rho-\delta]}^H \right)^c \right) = \epsilon$, respectively. Then, again by cor. (12), it follows that

$$R(M, N) < r_1^* < r_0^* = r_0^*(M, N) < r_2^*.$$

By symmetry, it follows that

$$R(M, N) < r_1^* < r_0^* = r_0^*(M, N) < r_2^* < 2r_0^* - R(M, N).$$

(38)
Note that by cor. (52) and by symmetry,

\[ P \left( \left( A_{[S(M,N),\rho-\delta]}^H \right)^c \right) = 0 \]

when \( r \geq 2r_0^* - R(M,N) \). Therefore, if \( \left( A_{[S(M,N),\rho-\delta]}^H \right)^c \) occurs with probability 0, then the property \( \left\{ \frac{M^2}{S(M,N)} < \frac{1}{2} \right\} \) occurs with probability 0. Otherwise, \( \left( A_{[S(M,N),\rho-\delta]}^H \right)^c \) would occur with positive probability, since \( \left\{ \frac{M^2}{S(M,N)} < \frac{1}{2} \right\} \subseteq \left\{ \frac{|<C>_{H_r}|}{S(M,N)} < \frac{1}{2} \right\} = \left( A_{[S(M,N),\rho-\delta]}^H \right)^c \), up to sets of \( P \)-measure zero, by lemma (54). Hence, \( \left\{ \frac{M^2}{S(M,N)} \geq \frac{1}{2} \right\} \) occurs with probability 1. As a result,

\[
\frac{M^2}{M^2 + 2MN + (N-1)^2} \geq \frac{1}{2}
\]

with probability 1. Therefore, with probability 1 for \( M \), it follows that \( N \) has the solution

\[
N \geq 1 - M^2 + \sqrt{2M^2(M^2 + 1)}.
\]

**Lemma 55** If \( r \geq 1/(2N) \), then \( P \left( \left( A_{[S(M,N),\rho-\delta]}^H \right)^c \right) = 0. \)

**Proof** Without loss of generality, suppose area\( (B) = 1 \) and further suppose that \( B \) is divided into squares with sides of length \( 2r = 1/N \). By hypothesis, \( B \) contains \( M^2 \) data points and it is to be divided into \( N^2 \) regions. Clearly then, there are no boundary hexagons separating each of the \( N^2 \) regions since the sides of \( B \) have length \( 2rN = 1 \) which gives \( B \) an area of 1. Let each square be inscribed with a circle of radius \( r \), which itself may be inscribed within a hexagon. By hypothesis, each of the \( N^2 \) hexagons in \( B \) contains at least one of the \( M^2 \) data points. Hence, each of the \( N^2 \) (occupied) hexagons is connected in a cluster to every other hexagon in \( B \) so that\( P \left( A_{[S(M,N),\rho-\delta]}^H \right) = 1. \) Since \( P \left( A_{[S(M,N),\rho-\delta]}^H \right) = 1 \) for \( r = 1/(2N) \), then \( P \left( A_{[S(M,N),\rho-\delta]}^H \right) = 1 \) for \( r \geq 1/(2N) \) by prop. (62).

As a result of lemma (55) and by using ineq. (38), a conservative estimate for \( r_0^* \) is given by a solution to

\[
2r_0^* - R(M,N) \geq \frac{1}{2N}
\]

(40)

that maximizes \( 1/(2N) \) as a function of \( M \). The value of \( N \) satisfies \( N \geq 1 - M^2 + \sqrt{2M^2(M^2 + 1)} \). A maximal solution is found when \( N = 1 - M^2 + \sqrt{2M^2(M^2 + 1)} \). As such, for \( \epsilon \in (0, 1/2) \), since \((r_1^*, r_2^*) \subset (R(M,N), 2r_0^* - R(M,N)) \), then by ineq. (40),

\[
r_2^* - r_1^* \approx 2r_0^* - 2R(M,N)
\]

(41)

\[
= \frac{1}{2N} - R(M,N)
\]

is an estimate of the length of the sharp threshold interval \( r_2^* - r_1^* \) about \( r_0^* \).

Using the value of \( r_0^* \) given by eq. (40) and by using the estimate for the length of the sharp threshold interval about \( r_0^* \) given by eq. (41), an estimate for the value of \( r_1^* \) can be obtained. Thus, when \( r \leq r_1^* \), the property \( \left( A_{[S(M,N),\rho-\delta]}^H \right)^c \) occurs with probability at least \( 1 - \epsilon \) and falls sharply to a probability of occurrence of \( \epsilon \) as \( r \to r_2^* \).
By cor. (48) and thm. (46), the left half of the sharp threshold interval about \( r_0 \) is given by \([r_1^*, r_0]\). Using lemma (30), there exists \( r_2 \leq r_2^* \) such that \([r_0, r_2] \) is the right half of the sharp threshold interval for \( \epsilon \) \( > \) 0 given. Thus, when \( r \leq r_1^* \), the property \( \left( A_r^{S(M,N),\rho-\delta} \right)^c \) occurs with probability at least \( 1 - \epsilon \) and falls sharply to a probability of occurrence of (no greater than) \( \epsilon \) as \( r \to r_2^* \). As such, the sharp threshold interval for classifying \( M^2 \) data points into \( N^2 \) classes, in the mean continuum case, is of length (no greater than) \( r_2^* - r_1^* \).

**Theorem 56** Let \( \Delta^*(M,N) \) denote the sharp threshold interval length for the event of classifying \( M^2 \) random data points into \( N^2 \) classes. Then,

\[
\Delta^*(M,N) = O(N^{-1}).
\]

**Proof** Follows directly from eq. (41), eq. (36) and thm. (49).

6. Conclusions

For a node process which generates points in a bounded region, the points connect in clusters when they are within a certain range of each other. It was shown that the set of graphs over connected points undergoes a phase transition such that the probability of the occurrence of at least half of all points being connected into a single cluster rises sharply from some small positive value to some value close to one on a short interval. The probability measure over the graph property in the continuum was shown to be equivalent to the probability measure over the graph property produced after the bounded region was partitioned by hexagons of a certain random size, which depends upon the number of points generated by the node process. With the given partition and probability measure, a framework was provided such that the length of the short interval could be estimated and a theoretical estimate was provided. Upon providing a more practical estimate of the prototypical hexagon, copies of which could be used to partition the bounded region, the necessary framework was provided to be able to give a practical estimate for the critical radius, the length of the short interval and its end points when performing \( K \)-means classification of a given set of data points under certain probability distribution requirements.

**Appendix A: Appendix**

A.1. **Graph**

**Proposition 57** If \( r < r' \), then \( G(X_n; r) \subseteq G(X_n; r') \).

**Proof** Suppose \( r < r' \). If \( < x, y >_r \in G(X_n; r) \), then \( d(x, y) \leq r < r' \) so that \( < x, y >_r \in G(X_n; r') \). Hence, \( G(X_n; r) \subseteq G(X_n; r') \).

A.2. **Increasing Property**

**Lemma 58** \( |A_{[n,\rho]}| \leq 1 \).

**Proof** If \( A_{[n,\rho]} = \emptyset \), then there is nothing to prove. Thus, suppose that \( A_{[n,\rho]} \) occurs and \( < C >_r \in A_{[n,\rho]} \). Since \( \rho_n(C) \geq \rho > 1/2 \), then all other connected components are of order strictly less than half of all points. Therefore, \( |A_{[n,\rho]}| = 1 \).
Proposition 59 \( A^r_{[n, \rho]} \) is an increasing property in \( r \).

Proof Suppose \( < C >_r \in A^r_{[n, \rho]} \) and fix arbitrary \( r' > r \). Then, \( d(x, y) \leq r < r' \) for all \( x, y \in < C >_r \). Thus, \( < C >_r \subseteq < C >_{r'} \) implies \( N = |< C >_r| \leq |< C >_{r'}| \). Hence, \( < C >_r \in A^r_{[n, \rho]} \) implies \( < C >_{r'} \in A^r_{[n, \rho]} \). Since \( r' > r \) is arbitrary, then \( A^r_{[n, \rho]} \) is an increasing property in \( r \).

Proposition 60 \( A^r_{[n, \rho]} \) is a decreasing property in \( n \).

Proof Suppose \( < C >_r \in A^r_{[n, \rho], r} \). If \( n' < n \), then \( |< C >_r|/n' > |< C >_r|/n \geq \rho \) so that \( < C >_r \in A^r_{[n', \rho], r} \). Hence, \( A^r_{[n, \rho]} \subseteq A^r_{[n', \rho], r} \). Since \( n' < n \), then \( A^r_{[n, \rho]} \) is decreasing in \( n \).

A.3. Probability Measure

Proposition 61 The property \( A^r_{[n, \rho]} \) is \( P \)-measurable.

Proof For \( x, y \in X_n \) and \( S \subseteq X_n \), define the state on \( < x, y >_r \) to be 1 if and only if \( < x, y >_r \in G(S; r) \) and -1 otherwise. Then, \( S \) mutually determines an element \( \omega_S \in \Omega = \{-1, 1\}^{X_n} \) so that \( S \) is \( P \)-measurable. Since \( A^r_{[n, \rho]} \) is the property that there exists \( \omega_S \in \Omega \) mutually determined by \( S \subseteq X_n \) such that \( (\max_{y \in S} |< C >_r|)/n \geq \rho \), then \( A^r_{[n, \rho]} \) is \( P \)-measurable.

Proposition 62 \( P(A^r_{[n, \rho]}(r)) \) is a non-decreasing function of \( r \).

Proof Suppose \( r_1 \leq r_2 \). Since \( A^r_{[n, \rho]} \) is an increasing property in \( r \) by prop. (59), then \( A^{r_1}_{[n, \rho]} \subseteq A^{r_2}_{[n, \rho]} \) so that \( P(A^{r_1}_{[n, \rho]}) \leq P(A^{r_2}_{[n, \rho]}) \) by properties of probability measures. Thus, \( P(A^r_{[n, \rho]}(r)) \) is non-decreasing in \( r \).

Proposition 63 \( P(A^r_{[n, \rho]}(n)) \) is a non-increasing function of \( n \).

Proof Suppose \( n' < n \). Since \( A^r_{[n, \rho]} \) is a decreasing property in \( n \) by prop. (60), then \( A^{r}_{[n, \rho]} \subseteq A^{r}_{[n', \rho], r} \) so that \( P(A^{r}_{[n, \rho]}(r)) \leq P(A^{r}_{[n', \rho], r}) \) by properties of probability measures. Thus, \( P(A^r_{[n, \rho]}(r)) \) is non-increasing in \( n \).

A.4. Connection Radius

Proposition 64 \( r(n, \rho, \epsilon) \) is a non-decreasing function of \( \epsilon \).

Proof Suppose \( \epsilon_1, \epsilon_2 \in (0, \frac{1}{2}) \) such that \( \epsilon_1 \leq \epsilon_2 \). Define \( r_1 = r(n, \rho, \epsilon_1) \) and \( r_2 = r(n, \rho, \epsilon_2) \) and suppose \( r_1 > r_2 \). Since \( P(A^r_{[n, \rho]}(r)) \) is non-decreasing in \( r \) by prop. (62), then \( P(A^{r_1}_{[n, \rho]}) \geq P(A^{r_2}_{[n, \rho]}) \geq \epsilon_2 \geq \epsilon_1 \). Hence, \( r_2 \in \{ r > 0 : P(A^r_{[n, \rho]}) \geq \epsilon_1 \} \) and \( r_2 < r_1 = \inf \{ r > 0 : P(A^r_{[n, \rho]}) \geq \epsilon_1 \} \). Contradiction. Thus, \( r_1 \leq r_2 \) so that \( r(n, \rho, \epsilon) \) is non-decreasing in \( \epsilon \).

Lemma 65 If \( R = 2 \ast \max \{ d(x, y) : x, y \in X_n \} \), then \( X_n = \{ x \in X_n : d(x, y) \leq R \} \) for all fixed \( y \in X_n \).

Proof Clearly, \( \{ x \in X_n : d(x, y) \leq R \} \subseteq X_n \). Conversely, fix any \( y \in X_n \). For every \( x \in X_n \), it is true that \( d(x, y) \leq 2 \ast \max \{ d(x, y) : x, y \in X_n \} = R \). Hence, \( X_n \subseteq \{ x \in X_n : d(x, y) \leq R \} \) for all fixed \( y \in X_n \). Thus, \( X_n = \{ x \in X_n : d(x, y) \leq R \} \) for all fixed \( y \in X_n \).

Corollary 66 If \( R = 2 \ast \max \{ d(x, y) : x, y \in X_n \} \), then \( C_y > R \in A^R_{[n, \rho]}(r) \) for all \( y \in X_n \) and \( n \geq 1 \).
**Proof** Fix an arbitrary \( y \in X_n \). By lemma (65), if \( C_y > R = \{ x \in X_n : d(x, y) \leq R \} \), then \( C_y > R = X_n \) so that \( |C_y > R| = |X_n| = n \). Therefore, since \( y \in X_n \) is arbitrary, then \( C_y > R = A^R_{\{n, \rho\}} \) for all \( y \in X_n \) and \( n \geq 1 \).

**Corollary 67** If \( R = 2 \times \max\{d(x, y) : x, y \in X_n\} \), then \( P(A^R_{\{n, \rho\}}) = 1 \) for all \( n \geq 1 \).

**Proof** By lemma (65) and cor. (66), it is true that \( X_n \in A^R_{\{n, \rho\}} \) for all \( n \geq 1 \) and \( \rho \in (\frac{1}{2}, 1) \). Thus, \( A^R_{\{n, \rho\}} \neq \emptyset \) for all \( n \geq 1 \) and \( \rho \in (\frac{1}{2}, 1) \). Hence, \( P(A^R_{\{n, \rho\}}) = 1 \) for all \( n \geq 1 \).

**Lemma 68** If \( R = 2 \times \max\{d(x, y) : x, y \in X_n\} \), then \( 0 < r(n, \rho, \epsilon) \leq R \) for all \( \epsilon \in (0, \frac{1}{2}) \).

**Proof** By lemma (65), it is true that \( X_n = \{ x \in X_n : d(x, y) \leq R \} \) for all fixed \( y \in X_n \). Therefore, \( P(A^R_{\{n, \rho\}}) = 1 \) \( \geq \epsilon \) for all \( \epsilon \in (0, \frac{1}{2}) \). Suppose that \( \epsilon_0 \in (0, \frac{1}{2}) \) exists such that \( r_0 = r(n, \rho, \epsilon_0) > R \). Thus, \( A^R_{\{n, \rho\}} \subseteq A^{\epsilon_0}_{\{n, \rho\}} \) so that

\[
1 = P(A^R_{\{n, \rho\}}) \leq P(A^{\epsilon_0}_{\{n, \rho\}})
\]

since \( P(A^{\epsilon_0}_{\{n, \rho\}}) \) is non-increasing in \( n \) by prop. (66), non-decreasing in \( r \) by prop. (62) and by properties of probability measures. Hence, \( P(A^{\epsilon_0}_{\{n, \rho\}}) = 1 \). But, then \( R \in \{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_0 \} \) and \( R < r_0 = \inf\{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_0 \} \). Contradiction. Thus, \( 0 < r_0 \leq R \). Therefore, \( 0 < r(n, \rho, \epsilon) \leq R \) for all \( \epsilon \in (0, \frac{1}{2}) \).

**Proposition 69** Suppose \( \{ \epsilon_k \in (0, \frac{1}{2}) \}_{k \geq 1} \) is any convergent sequence such that \( \epsilon_k \to \epsilon_0 \). Define \( r_k = r(n, \rho, \epsilon_k) \) and \( r_0 = r(n, \rho, \epsilon_0) \). For arbitrary \( \xi > 0 \), if \( \{ k \geq 1 : |P(A^r_{\{n, \rho\}}) - P(A^{\epsilon_0}_{\{n, \rho\}})| \geq \xi \} \) is a set of measure zero, then \( r_k \to r_0 \) as \( k \to \infty \).

**Proof** If \( \xi > 0 \) is arbitrary and \( \{ k \geq 1 : |P(A^r_{\{n, \rho\}}) - P(A^{\epsilon_0}_{\{n, \rho\}})| \geq \xi \} \) is a set of measure zero, then

\[
P(A^r_{\{n, \rho\}}) = P(A^{\epsilon_0}_{\{n, \rho\}}) \geq \epsilon_0
\]

for all \( k \geq 1 \). Hence, \( r_k \in \{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_0 \} \) for all \( k \geq 1 \). Thus,

\[
\lim_{k \to \infty} r_k = \lim_{k \to \infty} r(n, \rho, \epsilon_k) = \lim_{k \to \infty} \inf\{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_k \} = \inf\{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_0 \} = r(n, \rho, \epsilon_0) = r_0
\]

where eq. (42) and eq. (43) follow since \( r_k \in \{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_k \} \cap \{ r > 0 : P(A^r_{\{n, \rho\}}) \geq \epsilon_0 \} \) for all \( k \geq 1 \) and \( \epsilon_k \to \epsilon_0 \) as \( k \to \infty \).

**References**

[1] Bai, X.; Kumar, S.; Xuan, D.; Yun, Z.; Lai, T. (2006), “Deploying Wireless Sensors to Achieve Both Coverage and Connectivity”, *ACM MobiHoc’06, May 22 - 25, 2006, Florence, Italy*
[2] Bai, H.; Chen, X.; Guan, X. (2006), “Preserving Coverage for Wireless Sensor Networks of Nodes with Various Sensing Ranges”, Proceedings of 2006 IEEE International Conference on Networking, Sensing and Control, ICNSC 2006, Ft. Lauderdale, Florida, USA, April 23-25, 2006, 5459.

[3] Balister, A.; Bollobas, B.; Sarkar, A.; Walters, M. (2005), “Connectivity of Random k-Nearest Neighbour Graphs”, Advances in Applied Probability, Vol. 37, no. 1, pp. 1 - 24, 2005.

[4] Balister, A.; Bollobas, B.; Sarkar, A.; Kumar, S. (2007), “Reliable Density Estimates for Coverage and Connectivity in Thin Strips of Finite Length”, ACM MobiCom’07, Sep. 9 - 14, 2007

[5] Benes, V.; Rataj, J. (2004), Stochastic Geometry: Selected Topics, Kluwer Academic Publishers.

[6] Bettstetter, C. (2002), “On the Minimum Node Degree and Connectivity of a Wireless Multihop Network”, ACM MobiHoc’02

[7] Bhondekar, A.P.; Vig, R.; Singla, M.L.; Ghanshyam, C.; Kapur, P. (2009), “Genetic Algorithm Based Node Placement Methodology for Wireless Sensor Networks”, Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol. 1, IMECS 2009, Mar. 18 - 20, 2009, Hong Kong

[8] Bollobas, B. (1991), “Random Walks”, Proceedings of Symposia in Applied Mathematics, Vol. 9, 1991, pp. 1 - 20

[9] Bourgain, J.; Kahn, J.; Kalai, G.; Katznelson, Y.; Linial, N. (1992), “The Influence of Variables in Product Spaces”, Israel Journal of Mathematics, Vol. 77, No. 1 - 2, 1992, pp. 55 - 64

[10] Cai, Haiyan; Jia, Xiaohua; Mo, Sha (2010), “Critical Sensor Density for Partial Connectivity in Large Area Wireless Sensor Networks”, Proceedings IEEE InfoCom ’10, 2010

[11] Chen, Y.; Chuah, C.; Zhao, Q. (2005), “Sensor Placement for Maximizing Per Unit Cost in Wireless Sensor Networks”, Proc. IEEE Military Communications Conf., Oct. 2005

[12] Coffman Jr., E.G.; Shor P.W. (2005), “A Simple Proof of the O(\sqrt{n} \log^{3/4}(n)) Upright Matching Bound”, SIAM J. Disc Math, Vol. 4, No. 1, pp. 48 - 57, 2005.

[13] Consul, P.C. (1989), Generalized Poisson Distributions Properties and Applications, Marcel Dekker.

[14] Csaki, E. (1997), “Some Results for Two Dimensional Random Walks”, Advances in Combinatorial Methods and Applications to Probability and Statistics, 1997

[15] Durrett, R. (1991), The TEX Probability : Theory and Examples., Duxbury.

[16] Efrat, A.; Itai, A.; Katz, M. (1996), “Improvements on Bottleneck Matching and Related Problems Using Geometry”, Proceedings of the 12th Annual Symposium on Computational Geometry, pp. 301 - 310, 1996

[17] Franceschetti, M.; Booth, L.; Cook, M.; Meester, R.; Bruck, J. (2005), “Continuum Percolation with Unreliable and Spread-Out Connections”, Journal of Statistical Physics, Vol. 118, No. 3/4, Feb. 2005.

[18] Friedgut, E.; Kalai, G. (1996), “Every Monotone Graph Property has a Sharp Threshold”, Proceedings of the American Mathematical Society, Vol. 124, No. 10, 1996, pp. 2993 - 3002

[19] Geman, S.; Geman, D. (1984), “Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, No. 6 (Nov. 1984)

[20] Georgii, Hans-Otto (1988), Gibbs Measures and Phase Transitions, Walter de Gruyter.
[21] Gilbert, E.N. (1961), “Random Plane Networks”, *J. Soc. Indust. Appl. Math*, Vol. 9, No. 4, Dec. 1961, pp. 533 - 543

[22] Glauche, I.; Krause, W.; Sollacher, R.; Greiner, M. (2003), “Continuum Percolation of Wireless Ad Hoc Communication Networks”, *Physica A*, 325, 2003, pp. 577 - 600

[23] Goel, A.; Rai, S.; Krishnamachari, B. (2005), “Monotone Properties of Random Geometric Graphs have Sharp Thresholds”, *The Annals of Applied Probability*, Vol. 15, No. 4 (Nov. 2005), pp. 2535 - 2552

[24] Grimmett, Geoffrey (1999), *Percolation*, Springer-Verlag.

[25] Grimmett, Geoffrey (2006), *The Random-Cluster Model*, Springer-Verlag.

[26] Guyon, Xavier. (1995), *Random Fields on a Network: Modeling, Statistics and Applications*, Springer.

[27] Han, X.; Cao, X.; Zhang, Y.; Lloyd, E.; Shen, C. (2007), “Deploying Directional Wireless Sensors with Guaranteed Coverage and Connectivity”, *IEEE InfoCom 2007*

[28] Hogg, R.V.; McKean, J.W.; Craig, A.T. (2005), *Introduction to Mathematical Statistics*, Pearson Prentice Hall.

[29] Hou, T.; Li, V. (1986), “Transmission Range Control in Multi-hop Packet Radio Networks”, *IEEE Trans. on Communications*, Vol. 34, No. 1, 1986, pp. 38 - 44

[30] Kirilyuk, K.P. (2004), “Complex Dynamics of Autonomous Communication Networks and the Intelligent Communication Paradigm”, *Report at the International Workshop on Autonomic Communication, Berlin, 1819 October 2004*.

[31] Kleinrock, L.; Silvester, J. (1978), “Optimal Transmission Radii for Packet Radio Networks or Why Six is a Magic Number”, *NTC’78*

[32] Kolmogorov, A.N.; Fomin, S.V. (1970), *Introductory Real Analysis*, Dover Publications.

[33] Kuperberg, W. (1989), “Covering the Plane with Congruent Copies of a Convex Body”, *Bulletin of the London Mathematical Society*, Vol. 21, pp. 82 - 86, 1989

[34] Leighton, T.; Shor, P. (1989), “Tight Bounds for Minimax Grid Matching with Applications to the Average Case Analysis of Algorithms”, *Combinatorica*, Vol. 9, No. 2, 1989, pp. 161 - 187

[35] Liu, J.; Adler, M.; Towsley, D.; Yun, Z.; Zhang, C. (2006), “On Optimal Communication Cost for Gathering Correlated Data through Wireless Sensor Networks”, *ACM MobiCom’06*, Sep. 24 - 29, 2006

[36] Mameli, V.; Musio, M. (2013), *A Generalization of the Skew-Normal Distribution: The Beta Skew-Normal, Communications in Statistics - Theory and Methods*, Volume 42, pp. 2229-2244, 2013

[37] Meester, Ronald; Roy, Rahul (1996), *Continuum Percolation*, Cambridge University Press.

[38] Moise, E. (1990), *Elementary Geometry from an Advanced Standpoint*, Addison Wesley.

[39] Murphy, R. (2011), “A Data Classification Methodology Using the Random Cluster Model”, *eq Technologic White Paper (Used with Permission), Unpublished, 2011*

[40] Murphy, R. (2015), “Estimating the Mean Number of K-Means Clusters to Form”, *arXiv, ID 1503.03488*, 2015

[41] Rai, A. (2004), “The Spectrum of a Random Geometric Graph is Concentrated”, *http://arxiv.org/PS_cache/math/pdf/0408/0408103.pdf*, Sep. 2004

[42] Saligram, V.; Zhao, M. (2012), *Local Anomaly Detection, Journal of Machine Learning Research, WE/CP, Volume 22*, pp. 969 - 983, 2012

[43] Santi, P.; Blough, D. (2003), “The Critical Transmitting Range for Connectivity in Sparse Wireless Ad-Hoc Networks”, *IEEE Trans. on Mobile Computing, Vol. 2, No. 1, 2003*, pp.
25 - 39

[44] Schechter, E. (1996), *Handbook of Analysis and Its Foundations.*, Academic Press.

[45] Shiryaev, A.N. (1996), *Probability*, Springer Verlag.

[46] Shor, P.W.; Yukich, J.E. (1991), “Minimax Grid Matching and Empirical Measures”, *The Annals of Probability, Vol. 19, No. 3* (1991), pp. 1338 - 1348

[47] Song, X.; Wu, M.; Jermaine, C.; Ranka, S. (2007), *Conditional Anomaly Detection, IEEE Transactions on Knowledge and Data Engineering, Volume 19, Issue 5, pp. 631 - 635, 2007*

[48] Steinwart, I.; Hush, D.; Scovel, C. (2005), *A Classification Framework for Anomaly Detection, Journal of Machine Learning Research, Volume 6, pp. 211 - 232, 2005*

[49] Takagi, H.; Kleinrock, L. (1984), “Optimal Transmission Ranges for Randomly Distributed Packet Radio Terminals”, *IEEE Trans. on Communications, Vol. 32, No. 3, 1984, pp. 246 - 257*

[50] Wan, P.; Yi, C. (2004), “Asymptotic Critical Transmission Radius and Critical Neighbor Number for k-Connectivity in Wireless Ad-Hoc Networks”, *ACM MobiHoc ’04*

[51] Wang, X.; Guoliang, X.; Zhang, Y.; Lu, C.; Pless, R.; Gill, C. (2003), “Integrated Coverage and Connectivity Configuration in Wireless Sensor Networks”, *SenSys ’03, Nov. 5 - 7, 2003, Los Angeles, CA, USA*

[52] Xing, G.; Wang, X.; Zhang, Y.; Lu, C.; Pless, R.; Gill, C. (2005), “Integrated Coverage and Connectivity Configuration for Energy Conservation in Sensor Networks”, *ACM Trans. on Sensor Networks, Vol. 1, No. 1, 2005, pp. 36 - 72*

[53] Xue, F.; Kumar, P.R. (2006), “On the $\theta$-Coverage and Connectivity of Large Random Networks”, *Joint Special Issue of the IEEE Trans. on Information Theory and the IEEE/ACM Trans. on Networking on ’Networking and Information Theory’, (May 2006)*

[54] Yu, Y.; Hong, B.; Prasanna, V.K. (2005), “Communication Models for Algorithm Design Automation in Wireless Sensor Networks”, *Proceedings of the IEEE International Parallel and Distributed Processing Symposium, 2005*

[55] Zhang, H.; Hou, J. (2005), “Maintaining Sensing Coverage and Connectivity in Large Sensor Networks”, *Ad Hoc & Sensor Wireless Networks, Vol. 1, No. 1-2, 2005 pp. 89 - 124*

[56] Zhang, Y.; Chong, E.K.P.; Hannig, J.; Estep, D. (2010), “Continuum Limits of Markov Chains with Application to Network Modeling”, *Proceedings of the 49th IEEE Conference on Decision and Control, CDC 2010, December 15-17, 2010, pp. 6779 - 6784, 2010*