Analytical Lower Bounds on the Critical Density in Continuum Percolation

Zhenning Kong and Edmund M. Yeh
Department of Electrical Engineering
Yale University
New Haven, CT 06520, USA
Email: {zhenning.kong, edmund.yeh}@yale.edu

Abstract—Percolation theory has become a useful tool for the analysis of large-scale wireless networks. We investigate the fundamental problem of characterizing the critical density \( \lambda_c^{(d)} \) for \( d \)-dimensional Poisson random geometric graphs in continuum percolation theory. By using a probabilistic analysis which incorporates the clustering effect in random geometric graphs, we develop a new class of analytical lower bounds for the critical density \( \lambda_c^{(d)} \) in \( d \)-dimensional Poisson random geometric graphs. The lower bounds are the tightest known to date. In particular, for the two-dimensional case, the analytical lower bound is improved to \( \lambda_c^{(2)} \geq 0.7698 \).... For the three-dimensional case, we obtain \( \lambda_c^{(3)} \geq 0.4494 \).

I. INTRODUCTION

Recently, percolation theory has become a useful tool for the analysis of large-scale wireless networks [1]–[4]. A percolation process resides in a random graph structure, where nodes or links are randomly designated as either “occupied” or “unoccupied.” When the graph structure resides in continuous space, the resulting model is described by continuum percolation [5]–[10]. A major focus of continuum percolation theory is the \( d \)-dimensional random geometric graph induced by a Poisson point process with constant density \( \lambda \). A fundamental result for continuum percolation concerns a phase transition effect whereby the macroscopic behavior of the system is very different for densities below and above some critical value \( \lambda_c^{(d)} \). For \( \lambda < \lambda_c^{(d)} \) (subcritical), the component containing the origin contains a finite number of points almost surely. For \( \lambda > \lambda_c^{(d)} \) (supercritical), the component containing the origin contains an infinite number of points with a positive probability [7]–[10].

Naturally, the characterization of the critical density \( \lambda_c^{(d)} \) is a central problem in continuum percolation theory. Unfortunately, the exact value of \( \lambda_c^{(d)} \) is very difficult to find. For two-dimensional random geometric graphs, simulation studies show that \( \lambda_c^{(2)} \approx 1.44 \) [11], while the best analytical bounds obtained thus far are \( 0.696 < \lambda_c^{(2)} < 3.372 \) [6], [7]. Recently, in [12], the authors reduce the problem of characterizing \( \lambda_c^{(2)} \) to evaluating numerical integrals using a mapping between continuum percolation and dependent bond percolation on lattices. By Monte Carlo methods, they obtain numerical bounds \( 1.435 < \lambda_c^{(2)} < 1.437 \) with confidence 99.99%. Unfortunately, the bounds obtained in [12] are not in closed-form and cannot be generalized to higher dimensional cases.

In this paper, we give a new mathematical characterization of the critical density \( \lambda_c^{(d)} \) for Poisson random geometric graphs in \( d \)-dimensional Euclidean space, where \( d \geq 2 \). We develop an analytical technique based on probabilistic methods [13] and the clustering effect in random geometric graphs. This analysis yields a new class of lower bounds:

\[
\lambda_c^{(d)} \geq \frac{1}{V^{(d)} (1 - C_t^{(d)})} \quad (1)
\]

where \( V^{(d)} \) is the volume of a \( d \)-dimensional unit sphere and \( C_t^{(d)} \) is the \( t \)-th order cluster coefficient \( (t \geq 3) \) for \( d \)-dimensional Poisson random geometric graphs, which we will define later. This class of analytical lower bounds are the tightest known to date. In particular, by evaluating \( C_3^{(2)} \) in closed-form, the analytical lower bound for two-dimensional Poisson random geometric graphs is improved to \( \lambda_c^{(2)} \geq 0.7698 \). For three-dimensional Poisson random geometric graphs, we obtain the analytical lower bound \( \lambda_c^{(3)} \geq 0.4494 \).

By successively evaluating \( C_t^{(d)} \) for \( t \geq 4 \), we can obtain tighter lower bounds on \( \lambda_c^{(d)} \).

II. RANDOM GEOMETRIC GRAPHS

In wireless networks, a communication link exists between two nodes if the distance between them is sufficiently small, so that the received power is large enough for successful decoding. A mathematical model for this scenario is as follows. Let \( \| \cdot \| \) be the Euclidean norm, and \( f \) be some probability density function (p.d.f.) on \( \mathbb{R}^d \). Let \( X_1, X_2, ..., X_n \) be independent and identically distributed (i.i.d.) \( d \)-dimensional random variables with common density \( f \), where \( X_i \) denotes the random location of node \( i \) in \( \mathbb{R}^d \). The ensemble of all the graphs with undirected links connecting all those pairs \( \{x_i, x_j\} \) with \( \|x_i - x_j\| \leq r \) for \( r > 0 \), is called a random geometric graph [8]. In the following, we focus on random geometric graphs with \( X_1, X_2, ..., X_n \) distributed i.i.d. according to a uniform distribution over a given \( d \)-dimensional box \( A = [0, \sqrt{n}/\lambda]^d \) with \( \lambda > 0 \). We denote such graphs by \( G(\lambda_n^{(d)}; r) \).

Let \( A = |A| \) be the \( d \)-dimensional Lebesgue measure (or volume) of \( A \). An event is said to be asymptotic almost sure (abbreviated a.a.s.) if it occurs with a probability converging to 1 as \( n \to \infty \). Consider a graph \( G = (V, E) \), where \( V \)
and $E$ denote the set of nodes and links respectively. Given $u, v \in V$, we say $u$ and $v$ are adjacent if there exists an link between $u$ and $v$, i.e., $(u, v) \in E$. In this case, we also say that $u$ and $v$ are neighbors.

A. Preliminaries

In $G(\lambda n_d ; r)$, let the location $x_j$ of node $j$ be given. A second node $i$ is randomly placed in $A$ according to the uniform distribution $f$. There exists a link between these two nodes if and only if node $i$ lies within a sphere of radius $r$ around $x_j$. Let this spherical region be denoted as $A(x_j)$, then the probability for the existence of a link between $i$ and $j$ is given by

$$P_{\text{link}}(x_j) = \int_{A(x_j)} f(y) dy.$$  \hspace{1cm} (2)

Since the underlying distribution is uniform, the probability $P_{\text{link}}(x_j)$ depends only on the volume of the intersection between $A$ and the node coverage volume $A(x_j)$. Throughout this paper, we ignore border effects $\footnote{More rigorously, we may use a torus instead of a box for $A$. Asymptotically, as $n \to \infty$ and $A \to \infty$ with $n/A = \lambda$ fixed, uniform random geometric graphs on the torus are the same as those in the box.}$. As a consequence, $P_{\text{link}}(x_j)$ is independent of $x_j$, and thus independent among all the links:

$$P_{\text{link}} = \frac{V(d)r^d}{A} = \frac{\lambda V(d)r^d}{n},$$  \hspace{1cm} (3)

where $V(d)$ is the volume of a $d$-dimensional unit sphere $V(d) = \frac{\pi^{d/2}}{\Gamma(d/2)}$, and $\Gamma(\cdot)$ is the Gamma function $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t} dt$.

It follows that in $G(\lambda n_d ; r)$, the probability that the given node $j$ has degree $k$, $0 \leq k \leq n-1$, is given by the binomial distribution:

$$P_k = \binom{n-1}{k} P_{\text{link}}^k [1 - P_{\text{link}}]^{n-1-k}.$$  \hspace{1cm} (4)

Thus, the mean degree for each node is

$$\mu = E[k] = (n-1)P_{\text{link}} = \frac{(n-1)\lambda V(d)r^d}{n}.$$  \hspace{1cm} (5)

Note that as $n$ and $A$ both become large but the ratio $n/A = \lambda$ is kept constant, each node has an approximately Poisson degree distribution $\footnotemark[8, 14]$ with an expected degree

$$\mu = \lim_{n \to \infty} \frac{(n-1)\lambda V(d)r^d}{n} = \lambda V(d)r^d.$$  \hspace{1cm} (6)

As $n \to \infty$ and $A \to \infty$ with $n/A = \lambda$ fixed, $G(\lambda n_d ; r)$ converges in distribution to an (infinite) random geometric graph $G(\mathcal{H}_\lambda^{(d)} ; r)$ induced by a homogeneous Poisson point process with density $\lambda > 0$. For such graphs, we have the following lemma.

Lemma 1: Suppose $G(\mathcal{H}_\lambda^{(d)} ; r)$ is a random geometric graph induced by a homogeneous Poisson point process with density $\lambda > 0$. Let $A'$ be any subset of $\mathbb{R}^d$ with $|A'| < \infty$, then the subgraph $G'$ contained in $A'$ has a finite number of nodes a.a.s.

Proof: The proof is straightforward and omitted here. \hfill \Box

Due to the scaling property of random geometric graphs $\footnote{In the following, we focus on $G(\mathcal{H}_\lambda^{(d)} ; 1)$ and $G(\lambda n_d ; 1)$.}$, $\footnotemark[7, 8]$, the probability that nodes $i$ and $j$ are adjacent given that $i$ and $j$ are both adjacent to node $k$.

B. Cluster Coefficients

An important characteristic of random geometric graphs is the clustering effect. Here, if node $i$ is close to node $j$, and node $j$ is close to node $k$, then $i$ is typically also close to $k$. In the following, we use the cluster coefficients to precisely characterize the clustering property. This turns out to be the key to deriving new bounds for the critical density in continuum percolation.

Definition 1: Given distinct nodes $i, j, k$ in $G(\lambda n_d ; 1)$, the cluster coefficient $C^{(d)}$ is the conditional probability that nodes $i$ and $j$ are adjacent given that $i$ and $j$ are both adjacent to node $k$.

The calculation of $C^{(d)}$ for two-dimensional random geometric graphs is illustrated in Figure 1. To determine $C^{(d)}$, assume both nodes $i$ and $j$ lie within $A(x_i)$, then the conditional probability that nodes $i$ and $j$ are also adjacent is equal to the probability that two randomly chosen points in a circle with radius $1$ is at most distance $1$ apart. In other words, given the coordinates of $x_k$ and $x_i$, the probability that there is an link between $i$ and $j$ is equal to the fraction of $A(x_k)$ that intersects $A(x_i)$. By averaging $x_i$ over all points in $A(x_k)$, the cluster coefficient can be found as $C^{(d)} = 1 - \frac{3\sqrt{3}}{2\pi} \approx 0.5865, \ldots$ $\footnotemark[15]$.\footnotemark[15]

The cluster coefficient $C^{(d)}$ reflects the triangle effect in random geometric graphs. To further capture the cluster effect, we now generalize the notion of cluster coefficients for more than three nodes.

Definition 2: For $t \geq 3$, suppose $v_1, \ldots, v_{t-1} \in G(\lambda n_d ; 1)$ form a single chain, i.e., they satisfy the following properties:

i) For each $j = 1, 2, \ldots, t-2$, $(v_j, v_{j+1}) \in E$;

ii) For all $1 \leq j, k \leq t$, $(v_j, v_k) \notin E$ for $|j - k| > 1$, where $E$ denotes the set of links in $G(\lambda n_d ; 1)$. Then the $t$-th order cluster coefficient $C_t^{(d)}$ is defined to be the conditional probability that a node $v_t$ is adjacent to at least one of the

**Fig. 1. Calculation of cluster coefficient $C^{(2)}$**
Finally, we say that authors reduce the problem of characterizing $f$ to a lower bound for the origin can be arbitrary. For $t$ \geq 4, we are able to compute them by numerical integration. For example, we obtain $C_4(1) \approx 0.6012$ and $C_2(2) \approx 0.6179$. We also note that $0 < C_t(1) < 1$, for all $t \geq 3$, since $C_t(1)$ is a (nonzero) conditional probability.

C. Critical Density for Random Geometric Graphs

Let $\mathcal{H}_{\lambda_0} = \mathcal{H}_{\lambda}, \mathcal{H}_{\lambda_0} \cup \{0\}$, i.e., the union of the origin and the infinite homogeneous Poisson point spatial process with density $\lambda$ in $\mathbb{R}^d$. Note that in a random geometric graph induced by a homogeneous Poisson point process, the choice of the origin can be arbitrary.

Definition 3: For $G(\mathcal{H}_{\lambda_0}, 1)$, the percolation probability $p_{\infty}(\lambda)$ is the probability that the component containing the origin has an infinite number of nodes of the graph.

Definition 4: For $G(\mathcal{H}_{\lambda}, 1)$, the critical density (continuum percolation threshold) $\lambda_c(\lambda)$ is defined as $\lambda_c(\lambda) = \inf\{\lambda > 0 : p_{\infty}(\lambda) > 0\}$. It is known from continuum percolation theory that if $\lambda > \lambda_c(\lambda)$, then there exists a unique connected component containing $\Theta(n)$ nodes in $G(\mathcal{H}_{\lambda}, 1)$ a.a.s.\footnote{This large connected component is called the giant component.}

III. NEW LOWER BOUNDS ON THE CRITICAL DENSITY

A fundamental result of continuum percolation states that $0 < \lambda_c(\lambda) < \infty$ for all $d \geq 2$. Exact values for $\lambda_c(\lambda)$ and $p_{\infty}(\lambda)$ are not yet known. For $d = 2$, simulation studies [11] show that $\lambda_c(2) \approx 1.44$, while the best analytical bounds obtained thus far are $0.696 < \lambda_c(2) < 3.372$ [6], [7]. Recently, in [12], the authors reduce the problem of characterizing $\lambda_c(\lambda)$ to evaluating numerical integrals, and they obtain numerical bounds $1.435 < \lambda_c(2) < 1.437$ with confidence $99.99\%$. Unfortunately, these bounds are not in closed form and are restricted to the two dimensional case. In the following, we present an analysis which combines a technique used in [13] (for random graphs) and the clustering effect in random geometric graphs to obtain a new mathematical characterization of the critical density $\lambda_c(\lambda)$ for $d \geq 2$. This analysis yields a new class of improved lower bounds for $\lambda_c(\lambda)$. In particular, they yield the tightest analytical lower bounds known to date.

Theorem 1: Let $\mu$ be the mean degree of $G(\lambda_n(\lambda), 1)$, where $d \geq 2$. For any given integer $t \geq 3$, if

$$\mu < \frac{1}{1 - C_t(\lambda)},$$

where $C_t(\lambda)$ is defined by Definition 2, then the largest component of $G(\lambda_n(\lambda), 1)$ has at most $\alpha$ in $n$ nodes a.a.s., where $\alpha$ is a positive constant.

Before giving the proof, we define the diameter of a node with respect to a subgraph.

Definition 5: Given a graph $G = (V, E)$, for any subgraph $G' = (V', E'') \subseteq G$ and a node $u \in V'$, define the diameter of node $u$ with respect to $G' = (V', E')$ as

$$d(u, G') \equiv \max_{v \in V'} \{d(u, v)\},$$

where $d(u, v)$ is the distance between $u$ and $v$, measured by the length of the shortest path between $u$ and $v$ in terms of the number of links.

Note that the diameter of graph $G$ is the maximum of the diameters of all nodes with respect to graph $G$, i.e., $\text{diam}(G) = \max_{u \in G} \{\text{diam}(u, G)\}$. Another useful fact for the following proof is that for a random geometric graph $G(\lambda_n(\lambda), 1)$, if $\text{diam}(u, G') \leq c$, then the Euclidean distance between $u$ and any node $v$ in $G' = (V', E')$ is no more than $c$, i.e., $\|X_u - X_v\| \leq c$, for all $v \in V'$.

Proof of Theorem 7: Let $\mu = \frac{1 - C_t(\lambda)}{1 - C_t(\lambda)}$, where $0 < \epsilon < 1$. For simplicity, let $p$ denote the probability that there is an link between two nodes, i.e., $p = P_{\text{link}}$ given by (3), so that $(n - 1)p = \mu$.

We consider an arbitrary node (with fixed label and random position) $v \in G(\lambda_n(\lambda), 1)$ and study the following “active-saturated” process. For $i = 0, 1, 2, \ldots$, let $A_i$ denote the set of “active” nodes, and $S_i$ denote the set of “saturated” nodes, starting with $A_0 = \{v\}$, $S_0 = \emptyset$. At $(i + 1)$-th step, we select an arbitrary node $u$ from $A_i$ and update the active and saturated sets as follows:

$$A_{i+1} = (A_i \setminus u) \cup (N_i \cap (A_i \cup S_i'))', \quad S_{i+1} = S_i \cup u,$$

where $N_i$ is the set of neighbors of $u$. In other words, at each step we move a node $u$ from the active set to the saturated set, and at the same time move to the active set all the neighbors of $u$ which do not currently belong to the active or saturated...
Therefore, we can go through all the nodes in $v$’s component, represented by $\Gamma_v$, until $A_i = \emptyset$.

Let $Y_{i+1}$ be the number of nodes added to $A_i$ at step $i+1$:

$$Y_{i+1} = |N_i \cap (A_i \cup S_i)|.$$  

(10)

Note that $|S_i| = i$ for $i \leq |\Gamma_v|$.

We say a sample graph $G_n$ of $G(\lambda_n; 1)$ is good, if there is no component having size strictly larger than $\frac{3-2\epsilon}{\epsilon^2}\ln n$, or for any node $v \in G_n$ with $|\Gamma_v| > \frac{3-2\epsilon}{\epsilon^2}\ln n$ and any sequence of the “active-saturated” steps starting at $v$, there exists a bounded $k'$ such that

$$\forall j \geq k' \text{ and } \forall u \in A_j, \quad \text{diam}(u, S_j \cup u) \geq t - 2,$$

(11)

$$c_{k'} \equiv |A_{k'} \cup S_{k'}| < (\ln n)^{1/3}.$$  

(12)

Note that $k'$ depends on $n$, the sample graph $G_n$, the node $v \in G_n$ and the sequence of the “active-saturated” steps starting at $v$. Let $\mathcal{T}_n$ be the collection of all good sample graphs with $n$ nodes. We will show that with probability 1, there exists a uniform bound $k < \infty$ such that $k' \leq k_0$ for any $n$, any $G_n \in \mathcal{T}_n$ and any $v \in G_n \in \mathcal{T}_n$ with $|\Gamma_v| > \frac{3-2\epsilon}{\epsilon^2}\ln n$, and any sequence of the “active-saturated” steps starting at $v$. We assume that this holds for the moment.

Now given $G(\lambda_n^d; 1) \in \mathcal{T}_n$, consider an arbitrary node (with fixed label) $v \in G(\lambda_n^d; 1)$, if $|\Gamma_v| > \frac{3-2\epsilon}{\epsilon^2}\ln n$, the “active-saturated” process can sustain at least $\frac{3-2\epsilon}{\epsilon^2}\ln n$ steps. Let

$$k = \frac{3-2\epsilon}{\epsilon^2}\ln n - k' \geq 0.$$  

Because $|S_i| = i$, we have

$$|\Gamma_v| \geq k + k' \iff |A_{k+k'} \cup S_{k+k'}| \geq k + k'.$$  

(13)

Since $G(\lambda_n^d; 1) \in \mathcal{T}_n$, by probability 1, there exist constants $k_0$ and $k' \leq k_0$ satisfying condition (11) and (12). By the definition of $c_{k'}$, (13) is equivalent to

$$|\Gamma_v| \geq k + k' \iff \sum_{i=1}^{k+k'} Y_i \geq k + k' - c_{k'}.$$  

(14)

Therefore,

$$\Pr\{|\Gamma_v| \geq k + k'\} = \Pr\left\{\sum_{i=1}^{k+k'} Y_i \geq k + k' - c_{k'}\right\}.$$  

(15)

We now bound the RHS probability. Consider the “active-saturated” process after $k'$ steps. For all $j \geq k'$, we have $\text{diam}(u, S_j \cup u) \geq t - 2, \forall u \in A_j$. At each step $i$, we move $Y_i$ nodes to the active set. Suppose at the $(j+1)$-th step, $\sum_{i=1}^{j} Y_i = m$, and there are $n - 1 - m$ nodes remaining. Now suppose we move node $u_{j+1}$ from the active set to the saturated set, then $Y_{j+1}$ is the number of nodes adjacent to $u_{j+1}$ but not in $A_j$ or $S_j$. Since all nodes in $A_j$ are adjacent to some node in $S_j$, $Y_{j+1}$ is also the number of nodes adjacent to $u_{j+1}$, not in $S_j$ and not adjacent to any node in $S_j$. Since $\text{diam}(u_{j+1}, S_j \cup u_{j+1}) \geq t - 2$, there exists a sequence of nodes $w_j, w_{j+1}, \ldots, w_{j+1}$ in $S_j$ that forms a single chain with node $u_{j+1}$ (i.e., satisfies condition (i)-(ii) of Definition 2). Let

$$C_t(n, u_{j+1}, w_{j+1}, \ldots, w_{j+1})$$

be the conditional probability that one of the remaining $n - 1 - m$ nodes, $w$, is adjacent to at least one of the nodes $w_{j+1}, \ldots, w_{j+1}$ given that $w$ is adjacent to $u_{j+1}$. Then the probability that a node is adjacent to $u_{j+1}$ and not adjacent to any $w_{j+1}, i = 1, 2, \ldots, t - 2$, is $q_n(u_{j+1}) \equiv p_n(u_{j+1})(1 - \tilde{C}_t(n, u_{j+1}, w_{j+1}, \ldots, w_{j+1}))$, where $p_n(u_{j+1})$ is the average probability that there is a link between node $u_{j+1}$ and any other node. Since $\text{diam}(u_{j+1}, S_j \cup u_{j+1})$ may be larger than $t - 2$, and there are other geometric constraints for each of the $n - 1 - m$ remaining nodes, the probability of any one of the remaining $n - m - 1$ nodes is adjacent to $u_{j+1}$, not in $S_j$ and not adjacent to any node in $S_j$ is less than or equal to $q_n(u_{j+1})$.

Now $Y_{j+1} = \sum_{i=1}^{n-1-m} B_i$, where $B_i = 1$ if node $i$ is adjacent to $u_{j+1}$, not in $S_j$ and not adjacent to any node in $S_j$, and $B_i = 0$ otherwise. Note that these $B_i$’s are not independent. Nevertheless, by the argument above, we have

$$\forall i = 1, \ldots, n - 1 - m, \text{let } B_i \equiv \text{Binom}(n-1, q_n(u_{j+1}))\text{ random variables.}$$

By Proposition 1 in Appendix I, $Y_{j+1} \equiv \sum_{i=1}^{n-1-m} B_i \sim \text{Binom}(n-1, q_n(u_{j+1}))$, which is further stochastically upper bounded by $Y_{j+1} \sim \text{Binom}(n-1, q_n(u_{j+1}))$. Therefore, conditional on $Y_{j+1} = m$, for any $m \leq n - 1$, $Y_{j+1}$ is stochastically upper bounded by a random variable $Y'_{j+1}$ with distribution $\text{Binom}(n-1, q_n(u_{j+1}))$.

Using the same argument, we see that

$$\sum_{i=1}^{k+k'} Y_i \text{ is stochastically upper bounded by } \sum_{i=1}^{k+k'} Y'_i \equiv Z_k \sim \text{Binom}(k(n-1), q_n(u_m)),$$

where $q_n(u_m) = \text{sup}_{1+k \leq i \leq k+k'} q_n(u_i)$. Thus,

$$\Pr\left\{\sum_{i=1}^{k+k'} Y_i \geq k + k' - c_{k'}\right\} \leq \Pr\{Z_k \geq k + k' - c_{k'}\}.$$  

(16)

Let $\mu_n(u_m) \equiv (n-1)p_n(u_m)$ be the mean degree of node $u_m$. Since $E[Z_k] = k(n-1)q_n(u_m) = k\mu_n(u_m)(1 - C_t(n, u_m, w_1, \ldots, w_n))$, we have

$$\Pr\{Z_k \geq k + k' - c_{k'}\} = \Pr\{Z_k \geq E[Z_k] + k + k' - c_{k'} - E[Z_k]\} = \Pr\{Z_k \geq E[Z_k] + k\delta_n(u_m) + k' - c_{k'}\},$$

where $\delta_n(u_m) = 1 - \mu_n(u_m)(1 - C_t(n, u_m, w_1, \ldots, w_n))$. Note that conditioned on $G(\lambda_n^d; 1) \in \mathcal{T}_n$, the node distribution may not be uniform. Nevertheless, we will show that

$$\Pr\{G(\lambda_n^d; 1) \in \mathcal{T}_n \rightarrow 1 \text{ as } n \rightarrow \infty\}.$$
By the Chernoff bound [13], for $\delta > 0$,
\[
\Pr\{Z \geq E[Z] + \delta\} \leq \exp\left\{-\frac{\delta^2}{2E[Z] + 2\delta/3}\right\}. \tag{17}
\]

Thus, for $n$ sufficiently large,
\[
\Pr\{Z_u \geq k + k' - c_k'\} \leq \exp\left\{-\frac{2k(1 - \delta_n(u_m)) + 2\delta_n(u_m) + k' - c_k'}{2} \right\}
\]
\[
= \exp\left\{-\frac{3}{2} \left[ k^2\delta_n(u_m)^2 + 2k\delta_n(u_m)k' - c_k' + (k' - c_k')^2 \right] \right\}, \tag{18}
\]

Since $k = \frac{3 - 2\epsilon}{\epsilon^2} \ln n - k'$, with $c_k' < (\ln n)^{\frac{1}{3}}$, as $n \to \infty$, the RHS of (18) has the same order as
\[
\exp\left\{-\frac{3}{2} \left(3\ln n - k'\right)^2\delta_n(u_m)^2 \right\} \ln n
\]
\[
+ \frac{2\delta_n(u_m)}{3 - 2\delta_n(u_m)} \left(k' \left(1 - \delta_n(u_m)\right) - c_k'\right)\right\} \right\] \leq \exp\left\{-\frac{3\delta_n(u_m)(2 - \delta_n(u_m))}{2} \left(3 - 2\delta_n(u_m)\right) \right\}
\]
\[
\exp\left\{-\frac{3(3 - 2\epsilon)\delta_n(u_m)^2}{2(3 - 2\delta_n(u_m))} \ln n + \frac{3\delta_n(u_m)}{3 - 2\delta_n(u_m)}c_k'\right\}. \tag{19}
\]

Now choose $\gamma > 0$ such that $\frac{(\epsilon - \gamma)^2}{2(\epsilon - \gamma) - 3} > \frac{3}{2}$. Because $\delta_n(u_m) \to \epsilon$ as $n \to \infty$, there exists $0 < n_1 < \infty$, such that for $n \geq n_1$, $|\delta_n(u_m) - \epsilon| \leq \gamma$. Then using $\epsilon - \gamma \leq \delta_n(u_m) \leq \epsilon + \gamma$, we can bound (19) by
\[
c' \exp\left\{-\frac{3(3 - 2\epsilon)(\epsilon - \gamma)^2}{2(\epsilon - \gamma) - 3} \ln n + \frac{3(\epsilon + \gamma)}{3 - 2(\epsilon + \gamma)}c_k'\right\}
\]
\[
\leq c' \exp\left\{-\frac{3(3 - 2\epsilon)(\epsilon - \gamma)^2}{2(\epsilon - \gamma) - 3} \ln n + \frac{3(\epsilon + \gamma)}{3 - 2(\epsilon + \gamma)}c_k'\right\}
\]
\[
= O(n^{-\frac{1}{3}}), \tag{20}
\]

where
\[
c' = \exp\left\{-\frac{3(3 - 2\epsilon)(\epsilon - \gamma)^2}{2(\epsilon - \gamma) - 3}\right\}.
\]

By (15)-(20), for any arbitrary node $v \in G(X_n^{(d)}; 1) \in \mathcal{T}_n$,
\[
\Pr\{\Gamma_v \geq 3\ln n\} = O(n^{-\frac{1}{3}}). \tag{21}
\]

Set $\alpha = \frac{3 - 2\epsilon}{\epsilon^2}$. The probability that random geometric graph $G(X_n^{(d)}; 1)$ has at least one component whose size is no smaller than $\alpha \ln n$ is
\[
\Pr\{\exists v \in G(X_n^{(d)}; 1) : |\Gamma_v| \geq \alpha \ln n\}
\]
\[
= \Pr\{\exists v \in G(X_n^{(d)}; 1) : |\Gamma_v| \geq \alpha \ln n\} | G(X_n^{(d)}; 1) \in \mathcal{T}_n\}
\]
\[
\cdot \Pr\{G(X_n^{(d)}; 1) \in \mathcal{T}_n\}
\]
\[
+ \Pr\{\exists v \in G(X_n^{(d)}; 1) : |\Gamma_v| \geq \alpha \ln n\} | G(X_n^{(d)}; 1) \notin \mathcal{T}_n\}
\]
\[
\cdot \Pr\{G(X_n^{(d)}; 1) \notin \mathcal{T}_n\}
\]
\[
\leq \alpha \Pr\{\Gamma_v \geq \alpha \ln n\} | G(X_n^{(d)}; 1) \in \mathcal{T}_n\} \Pr\{G(X_n^{(d)}; 1) \in \mathcal{T}_n\}
\]
\[
+ \Pr\{G(X_n^{(d)}; 1) \notin \mathcal{T}_n\}.
\]

To complete the proof, we show two facts:
(i) $\Pr\{G(X_n^{(d)}; 1) \in \mathcal{T}_n\} \to 1$ as $n \to \infty$.
(ii) With probability 1, there exists $k_0 < \infty$ such that $k' \leq k_0$ for any $n$, any $G(X_n^{(d)}; 1) \in \mathcal{T}_n$, and any $v \in G(X_n^{(d)}; 1) \in \mathcal{T}_n$, and any realization of the “active-saturated” process starting at $v$.

To show (i), note that in $G(X_n^{(d)}; 1)$, if there is no component having size strictly larger than $\alpha \ln n$, then it is good; otherwise, we prove that for any node $v \in G(X_n^{(d)}; 1)$ with $|\Gamma_v| > \frac{3 - 2\epsilon}{\epsilon^2} \ln n$ and any realization of the “active-saturated” process starting at $v$, there exists a bounded $k'$ such that (11) holds a.a.s. Suppose for some $v \in G(X_n^{(d)}; 1)$ with $|\Gamma_v| > \frac{3 - 2\epsilon}{\epsilon^2} \ln n$ and a realization of the “active-saturated” process starting at $v$ such that for any $k \leq |\Gamma_v| - 1$, there exists a step $j$, $k \leq j \leq |\Gamma_v| - 1$, and $w \in A_j$, such that $\text{diam}(w, S_j \cup w) < t - 2$. Since $|\Gamma_v| > \frac{3 - 2\epsilon}{\epsilon^2} \ln n$, as $n \to \infty$, $|\Gamma_v| \to \infty$, the “active-saturated” process can go on forever. Thus, $S_j$ asymptotically contains an infinite number of nodes. Since $\text{diam}(w, S_j \cup w) < t - 2$, all the nodes of $S_j$ lie in a ball centered at $w$ with radius $t - 2$ (by the argument immediately following Definition 5), which occurs with probability approaching 0 as $n \to \infty$, by Lemma 1. Now suppose $k'$ satisfying (11) exists but is unbounded as $n \to \infty$. Then let $k'$ be the smallest $k'$ satisfying (11). For step $j = k' - 1$, there exists at least one node $w \in A_j$ such that $\text{diam}(w, S_j \cup w) < t - 2$. Then, arguing as before, we can show this happens with probability approaching 0 as $n \to \infty$.

Next we show that $c_k' < (\ln n)^{\frac{1}{3}}$ a.a.s. We know $c_k' = |A_k \cup S_k| = \sum_{i=1}^{k'} Y_i$. Using an argument similar to that leading to (16), we see that the RHS is stochastically upper bounded by a random variable with distribution Binom($k' (n - 1), p$). Assuming fact (ii), which is shown below, the RHS is further stochastically upper bounded by a random variable $\tilde{c}_k'$ with distribution Binom(k_0 n, p). Applying the Chernoff bound (17), we have, for sufficiently large $n$,
\[
\Pr\{\tilde{c}_k' \geq (\ln n)^{\frac{1}{3}}\} \leq \Pr\{c_k' \geq (\ln n)^{\frac{1}{3}}\}
\]
\[
= \Pr\{c_k' \geq k_0 \mu + (\ln n)^{\frac{1}{3}} - k_0 \mu\}
\]
\[
\leq \exp\left\{-\frac{\left[(\ln n)^{\frac{1}{3}} - k_0 \mu\right]^2}{2k_0 \mu + \frac{3}{2}((\ln n)^{\frac{1}{3}} - k_0 \mu)}\right\}
\]
\[
= 0 \text{ a.a.s.} \tag{22}
\]

Finally, we show fact (ii). Suppose that for each $i = 1, 2, \ldots$, there exists $n_i$ and $v \in G(X_n^{(d)}; 1) \in \mathcal{T}_n$, with $n_i \geq |\Gamma_v| > \frac{3 - 2\epsilon}{\epsilon^2} \ln n_i$, and a realization of the “active-saturated” process starting at $v$ such that $|\Gamma_v| - 1 \geq k' \geq i$. As $i \to \infty$, $k' \to \infty$, $|\Gamma_v| \to \infty$, and $n_i \to \infty$. However, this holds with probability 0. This completes our proof. □

Note that by either ignoring border effects or by taking $n \to \infty$, $C_t^{(d)}$ (defined for $G(X_n^{(d)}; 1)$) is also the $t$th-order cluster coefficient for the infinite Poisson random geometric
graph $G(H_{\lambda}^{(d)}; 1)$. Thus by Theorem 1 we obtain the following important corollaries giving new improved lower bounds on the critical density for $d$-dimensional Poisson random geometric graphs.

**Corollary 1:** Let $\mu_c^{(d)}$ and $\lambda_c^{(d)}$ be the critical mean degree and critical density for $G(H_{\lambda}^{(d)}; 1)$, respectively, where $d \geq 2$. Then for all $t \geq 3$,

$$\mu_c^{(d)} \geq \frac{1}{1 - C_t^{(d)}}, \quad \text{and} \quad \lambda_c^{(d)} \geq \frac{1}{V(d) \left(1 - C_t^{(d)}\right)}, \quad (23)$$

**Proof:** Follows immediately from Theorem 1 and (6). $\square$

In particular, for two-dimensional Poisson random geometric graphs, substituting $V(2) = \pi$ and $C_3^{(2)} = 1 - \frac{2\sqrt{2}}{\pi} = 0.5865...$ into (23), we have the following corollary.

**Corollary 2:** The critical mean degree $\mu_c^{(2)}$ and the critical density $\lambda_c^{(2)}$ for $G(H_{\lambda}^{(2)}; 1)$ satisfy $\mu_c^{(2)} \geq 2.419...$ and $\lambda_c^{(2)} \geq 0.7698...$.

Note that if we use high-order cluster coefficients $C_{t-2}^{(2)}$, $t \geq 4$, computed by numerical methods, e.g., $C_{4}^{(2)} \approx 0.6012$ and $C_{5}^{(2)} \approx 0.6179$, we can obtain further improved (approximate) lower bounds: $\mu_c^{(2)} \geq 2.617,$ and $\lambda_c^{(2)} \geq 0.885$. By applying $C_3^{(3)} = 0.4688...$ given by (8), we have

**Corollary 3:** The critical mean degree $\mu_c^{(3)}$ and the critical density $\lambda_c^{(3)}$ for $G(H_{\lambda}^{(3)}; 1)$ satisfy $\mu_c^{(3)} \geq 1.412...$ and $\lambda_c^{(3)} \geq 0.4494...$.

This lower bound is close to the known results obtained by simulation—0.65 [16], and it is the best known analytical lower bound on $\lambda_c^{(3)}$.

IV. CONCLUSION

We have established a new class of analytical lower bounds on the critical density $\lambda_c^{(d)}$ for percolation in $d$-dimensional Poisson random geometric graphs. These analytical lower bounds are the tightest known to date, and reveal a deep underlying relationship between the cluster coefficient and the critical density in continuum percolation.

ACKNOWLEDGEMENT

We would like to thank the anonymous referees for their helpful comments and suggestions, which improved the paper. This research is supported in part by Army Research Office (ARO) grant W911NF-06-1-0403.

REFERENCES

[1] L. Booth, J. Bruck, M. Franceschetti, and R. Meester, “Covering algorithms, continuum percolation and the geometry of wireless networks,” *Annals of Applied Probability*, vol. 13, pp. 722–741, May 2003.

[2] M. Franceschetti, L. Booth, M. Cook, J. Bruck, and R. Meester, “Continuum percolation with unreliable and spread out connections,” *Journal of Statistical Physics*, vol. 118, pp. 721–734, Feb. 2005.

[3] O. Dousse, M. Franceschetti, and P. Thiran, “Information theoretic bounds on the throughput scaling of wireless relay networks,” in *Proc. IEEE INFOCOM’05*, Mar. 2005.

[4] O. Dousse, F. Baccelli, and P. Thiran, “Impact of interferences on connectivity in ad hoc networks,” *IEEE Trans. Network.*, vol. 13, pp. 425–436, April 2005.

[5] E. N. Gilbert, “Random plane networks,” *J. Soc. Indust. Appl. Math.*, vol. 9, pp. 533–543, 1961.

[6] P. Hall, “On continuum percolation,” *Annals of Prob.*, vol. 13, pp. 1250–1266, 1985.

[7] R. Meester and R. Roy, *Continuum Percolation*. New York: Cambridge University Press, 1996.

[8] M. Penrose, *Random Geometric Graphs*. New York: Oxford University Press, 2003.

[9] G. Grimmett, *Percolation*. New York: Springer, second ed., 1999.

[10] B. Bollobás and O. Riordan, *Percolation*. New York: Cambridge University Press, 2006.

[11] J. Quintanilla, S. Torquato, and R. M. Ziff, “Efficient measurement of the percolation threshold for fully penetrable discs,” *Physica A*, vol. 86, pp. 399–407, 2000.

APPENDIX I

**Proposition 1:** Suppose random variables $X_i, i = 1, ..., m$ satisfy the following conditions: (i) $\Pr\{X_i \geq x\} \leq \Pr\{X_i^+ \geq x\}, \forall x$; (ii) $\Pr\{X_i \geq x|X_1 = x, ..., X_{i-1} = x_{i-1}\} \leq \Pr\{X_i^+ \geq x, \forall x, x_1, ..., x_{i-1}\}$; (iii) $X_i^+, i = 1, ..., m$ are independent of each other and of $X_1, i = 1, ..., m$. Then

$$\Pr\{\sum_{i=1}^{m} X_i \geq z\} \leq \Pr\{\sum_{i=1}^{m} X_i^+ \geq z\}, \forall z$$

**Proof:** It suffices to show the result for $m = 2$. Since

$$\Pr\{X_2 \geq y|X_1 = x_1\} \leq \Pr\{X_2^+ \geq y|X_1 = x_1\}, \forall x_1, y,$$

$$\Pr\{X_1 + X_2 \geq z|X_1 = x_1\} \leq \Pr\{X_1 + X_2^+ \geq z|X_1 = x_1\}$$

$$= \Pr\{X_1 \geq z - X_2^+\}$$

$$= \sum_{x_2} \Pr\{X_2^+ = x_2\} \Pr\{X_1 + X_2^+ \geq z|X_2^+ = x_2\}$$

$$= \sum_{x_2} \Pr\{X_2^+ = x_2\} \Pr\{X_1 \geq z - x_2\}$$

$$\leq \sum_{x_2} \Pr\{X_2 \geq y|X_1 = x_1\} \Pr\{X_2^+ \geq z - x_2\}$$

$$= \sum_{x_2} \Pr\{X_2^+ = x_2\} \Pr\{X_2^+ \geq z - x_2\}$$

$$= \sum_{x_2} \Pr\{X_2^+ = x_2\} \Pr\{X_2^+ \geq z - x_2\}$$

$$= \Pr\{X_1 + X_2 \geq z\}$$.

$\square$