Distance Distributions in Finite Uniformly Random Networks: Theory and Applications

Sunil Srinivasa and Martin Haenggi

Department of Electrical Engineering

University of Notre Dame

Notre Dame, IN 46556, USA

Email: {*ssriniv1, mhaenggi*}*@nd.edu*

Abstract

In wireless networks, the knowledge of nodal distances is essential for performance analysis and protocol design. When determining distance distributions in random networks, the underlying nodal arrangement is almost universally taken to be a stationary Poisson point process. While this may be a good approximation in some cases, there are also certain shortcomings to this model such as the fact that in practical networks, the number of nodes in disjoint areas are not independent. This paper considers a more realistic network model where a known and fixed number of nodes are independently distributed in a given region and characterizes the distribution of the Euclidean internode distances. The key finding is that when the nodes are uniformly randomly placed inside a ball of arbitrary dimensions, the probability density function of the internode distances follows a generalized beta distribution. This result is applied to study wireless network characteristics such as energy consumption, interference, outage and connectivity.

I. INTRODUCTION

A. Motivation

In wireless channels, the received signal strength falls off with distance according to a power law, at a rate termed the large scale path loss exponent (PLE) [1]. Given a link distance l, the signal power at the receiver is attenuated by a factor of $l^{-\alpha}$, where α is the PLE. Consequently, in wireless networks, distances between nodes strongly impact the signal-to-noise-and-interference ratios (SINRs), and, in turn, the link reliabilities. The knowledge of the nodal distances is therefore essential for the performance analysis and the design of efficient protocols and algorithms.

In many wireless networks, nodes can be assumed to be scattered randomly over an area or volume; the distance distributions then follow from the spatial stochastic process governing the locations of the nodes. For the sake of analytical convenience, the arrangement of nodes in a random network is commonly taken to be a homogeneous (or stationary) Poisson point process (PPP). For the resulting so-called "Poisson network" of density λ, the number of nodes in any given set V of Lebesgue measure $|V|$ is Poisson with mean $\lambda|V|$, and the numbers of nodes in disjoint sets are independent. Even though the PPP assumption can lead to some insightful results, practical networks differ from Poisson networks in certain aspects. First, networks are usually formed by scattering a fixed (and finite) number of nodes in a given area. In this case, the nodal arrangement is a *binomial point process* (BPP), which we define shortly. Secondly, since the area or volume of deployment is necessarily finite, the point process formed is non-stationary and often non-isotropic, meaning that the network characteristics as seen from a node's perspective such as the nearest-neighbor distance or the interference distribution is not the same for all nodes. Furthermore, the numbers of nodes in disjoint sets are not independent; in the case of the BPP, they are governed by a multinomial distribution.

Definition: Formally, a BPP Φ, is formed as a result of distributing N points independently uniformly in a compact set W.

The density of the BPP at any location x is defined to be $\lambda(x) = \frac{N}{|W|}\mathbf{1}(x)$. In this paper, we consider $W \subset \mathbb{R}^d$ (d is an arbitrary positive integer). For any set $V \subset \mathbb{R}^d$, the number of points in V, $\Phi(V)$, is binomial(n, p) with parameters $n = N$ and $p = |V \cap W|/|W|$ [3]. By this property, the number of nodes in disjoint sets are joint via a multinomial distribution. Accordingly, for disjoint sets $V_1, \ldots, V_k$ and $n = n_1 + \ldots + n_k$, we have

$$\Pr(\Phi(V_1) = n_1, \ldots, \Phi(V_k) = n_k) = \frac{n!}{n_1! \cdots n_k!} \frac{|V_1 \cap W|^{n_1} \cdots |V_k \cap W|^{n_k}}{|W|^n}.$$

If the number of nodes or users is known, the PPP is clearly not a good model, since realizations of the process may have more nodes than the number of nodes deployed or no nodes at all. In particular when the number of nodes is small, the Poisson model is inaccurate. The main shortcoming of the Poisson assumption is, however, the independence of the number of nodes in disjoint areas. For example, if all the N nodes are located in a certain part of the network area, the remaining area is necessarily empty. This simple fact is not captured by the Poisson model. This motivates the need to study and accurately characterize finite uniformly random networks, in an attempt to extend the plethora of results for the PPP to the often more realistic case of the BPP. We call this new model a *binomial network*, and it applies to mobile ad hoc and sensor networks and wireless networks with infrastructure, such as cellular telephony networks.

In this paper, we analytically characterize the distribution of internode distances in a binomial network wherein a known number of nodes are independently distributed in a compact set. As a special case, we derive the Euclidean distance properties in a d-dimensional isotropic[1] BPP, and use it to study relevant problems in wireless networks such as energy consumption, design of efficient forwarding and localization algorithms, interference characterization, and outage and

[1]A point process is said to be isotropic if its distribution is invariant to rotations.

connection probability evaluation.

B. Related Work

Even though the knowledge of the statistics of the node locations in wireless networks is crucial, relatively few results are available in the literature in this area. Moreover, much of the existing work deals only with moments of the distances (means and variances) or characterizes the exact distribution only for very specific system models.

In [4], the probability density function (pdf) and cumulative distribution function (cdf) of the distances between nodes are derived for networks with uniformly random and Gaussian distributed nodes over a rectangular area. [5] studies mean internodal distance properties for several kinds of multihop systems such as ring networks, Manhattan street networks, hypercubes and shufflenets. [6] provides closed-form expressions for the distributions in d-dimensional homogeneous PPPs and describes several applications of the results for large networks. [7] considers one-dimensional Poisson networks and analyzes the distribution and moments of the single-hop distance, which is defined as the maximum possible distance between two nodes that can communicate with each other. [8] derives the joint distribution of distances of nodes from a common reference point for networks with a finite number of nodes randomly distributed on a square and [9] determines the pdf and cdf of the distance between two randomly selected nodes in square random networks.

II. Distribution of Internode Distances

In this section, we determine the distribution of the Euclidean distance to the n^{th} nearest point from an arbitrary reference point for a general BPP. In the special case of a d-dimensional isotropic BPP, we establish that this random variable (r.v.) follows a generalized beta distribution. We also derive the distances to the nearest and farthest nodes and the void probabilities.

Consider the BPP Φ with N points uniformly randomly distributed in a compact set $W \subset \mathbb{R}^d$ (see Fig. 1). Let R_n denote the r.v. representing the Euclidean distance from an arbitrary reference

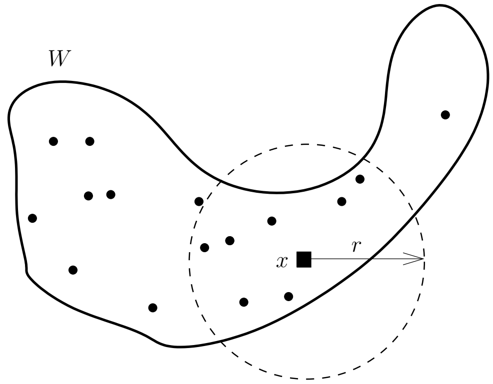

Fig. 1. A BPP with $N = 16$ points uniformly randomly distributed in an arbitrary compact set W. We wish to determine the distribution of the distances to the other points from the reference point x. The dashed circle represents the ball $b_d(x, r)$.

point x to the n^{th} nearest node of the BPP[2] and let $b_d(x, r)$ denote the d-dimensional ball of radius r centered at x.

The complementary cumulative distribution function (ccdf) of R_n is the probability that there are less than n points in $b_d(x, r)$:

$$\bar{F}_{R_n}(r) = \sum_{k=0}^{n-1} \binom{N}{k} p^k (1-p)^{N-k}, \quad 0 \leq r \leq R, \tag{1}$$

where $p = |b_d(x, r) \cap W|/|W|$. In the case of a *non-homogeneous* BPP with a general density function $\lambda(x)$, $p = \int_{b_d(x,r)\cap W} \lambda(x)\mathrm{d}x$.

$\bar{F}_{R_n}$ can be written in terms of the regularized incomplete beta function as

$$\bar{F}_{R_n}(r) = I_{1-p}(N - n + 1, n), \quad 0 \leq r \leq R, \tag{2}$$

where

$$I_x(a, b) = \frac{\int_0^x t^{a-1}(1-t)^{b-1}\mathrm{d}t}{B(a, b)}.$$

Here, $B(a, b)$ denotes the beta function, which is expressible in terms of gamma functions as

[2]For the rest of the paper, we assume that x is not a point of the BPP. However, if $x \in \Phi$, the remaining point process simply becomes a BPP with $N - 1$ points.

$B(a,b)=\Gamma(a)\Gamma(b)/\Gamma(a+b)$.

The pdf of the distance function is given by

$$\begin{aligned} f_{R_n} &= -\mathrm{d}\bar{F}_{R_n}/\mathrm{d}r \\ &= \frac{\mathrm{d}p}{\mathrm{d}r}\frac{(1-p)^{N-n}p^{n-1}}{B(N-n+1,n)}. \end{aligned} \tag{3}$$

We now analytically derive the pdf of the Euclidean distance between points in a d-dimensional isotropic BPP, and later, in Section III, compute its moments. In Section IV, we derive the pdf of the distances when W is a general l-sided regular polygon. In Section V, we apply our findings to the study of wireless networks.

Theorem 2.1: In a point process consisting of N points uniformly randomly distributed in a d-dimensional ball of radius R centered at the origin, the Euclidean distance R_n from the origin to its n^{th} nearest point follows a generalized beta distribution, i.e.,

$$f_{R_n}(r) \;=\; \frac{d}{R}\frac{B(n-1/d+1,N-n+1)}{B(N-n+1,n)}\beta\left(\left(\frac{r}{R}\right)^d; n-\frac{1}{d}+1, N-n+1\right), \quad r\in[0,R],$$

where $\beta(x;a,b)$ is the beta density function[3] defined as $\beta(x;a,b)=\frac{1}{B(a,b)}x^{a-1}(1-x)^{b-1}$.

Proof: For the isotropic d-dimensional BPP, we have $W=b_d(o,R)$. The volume of this ball $|W|$ is equal to $c_d R^d$, where

$$c_d = |b_d(0,1)| = \frac{\pi^{d/2}}{\Gamma(1+d/2)}$$

is the volume of the unit ball in $\mathbb{R}^d$ [3]. Important cases include $c_1=2$, $c_2=\pi$ and $c_3=4\pi/3$. The density of this process is equal to $N/(c_d R^d)$ inside the ball.

With the reference point being the origin, note that $p=c_d r^d/c_d R^d=(r/R)^d$ and from (3),

[3]Mathematica: PDF[BetaDistribution[a, b],x].

we have

$$\begin{aligned} f_{R_n}(r) &= \frac{d}{R}\left(\frac{r}{R}\right)^{d-1} \frac{(1-p)^{N-n}p^{n-1}}{B(N-n+1,n)} \\ &= \frac{d}{R}\frac{(1-p)^{N-n}p^{n-1/d}}{B(N-n+1,n)} \\ &= \frac{d}{R}\frac{B(n-1/d+1,N-n+1)}{B(N-n+1,n)}\beta\left(\left(\frac{r}{R}\right)^d; n-\frac{1}{d}+1, N-n+1\right) \end{aligned} \tag{4}$$

for $0 \leq r \leq R$. The final equality casts R_n as a generalized beta-distributed variable. ■

Corollary 2.2: For the practical cases of $d = 1$ and $d = 2$, we have

$$f_{R_n}(r) = \frac{1}{R}\beta\left(\frac{r}{R}; n, N-n+1\right)$$

and

$$f_{R_n}(r) = \frac{2}{R}\frac{\Gamma(n+\frac{1}{2})\Gamma(N+1)}{\Gamma(n)\Gamma(N+\frac{3}{2})}\beta\left(\frac{r^2}{R^2}; n+\frac{1}{2}, N-n+1\right)$$

respectively.

Fig. 2 plots the distance pdfs for the cases of $d = 1$ and $d = 2$.

Remarks:

1) The void probability p_B^0 of the point process is defined as the probability of there being no point of the process in an arbitrary test set B [3]. For a BPP with N points distributed over a set W, it is easy to see that

$$p_B^0 = (1 - |B \cap W|/|W|)^N . \tag{5}$$

For the isotropic BPP considered above, when the test set is $B = b_d(o, r)$, we have $p_B^0 = \left(1 - (r/R)^d\right)^N$.

2) Of interest in particular are the nearest- and farthest-node distances. The nearest-node

distance pdf is given by

$$f_{R_1}(r) = \frac{dN}{r}\left(1-\left(\frac{r}{R}\right)^d\right)^{N-1}\left(\frac{r}{R}\right)^d, \tag{6}$$

and the distance to the farthest point from the origin is distributed as

$$f_{R_N}(r) = \frac{dN}{r}\left(\frac{r}{R}\right)^{Nd}, \quad 0 \le r \le R. \tag{7}$$

Both are generalized Kumaraswamy distributions [10].

3) For a one-dimensional BPP, $f_{R_n}(r) = f_{R_{N-n+1}}(R-r)$, and therefore knowledge of the distance pdfs for the nearest $\lceil N/2 \rceil$ nodes gives complete information on the distance distributions to the other points.
4) If a point of the BPP, x, is located at the origin, the remaining $N-1$ points are uniformly distributed in $b_d(0,R)$. Thus, the pdf of the Euclidean distance from x to its neighbors is identical to (4), with N replaced by $N-1$. Also note that (4) also holds for any reference point x for $0 \le r \le R - \|x\|$.

We wish to compare the distance distributions from the origin for an isotropic BPP and a PPP with the same density. However, note that in general, the PPP may have fewer points than the number dropped. In order to make a fair comparison, we condition on the fact that there are at least N points present in the PPP model. The following corollary establishes the distance pdfs for such a conditioned PPP. Also note that conditioned on there being exactly N points present, the PPP is equivalent to a BPP [3].

Corollary 2.3: Consider a PPP of density λ over a finite volume $b_d(o,R)$. Conditioned on there being at least N points in the ball, the distance distribution from the origin to the n^{th} nearest node ($n \le N$) is given by

$$f'_{R_n}(r) = \frac{\lambda d c_d r^{d-1}\left(A_{n-1}(r)\left(\sum_{k=N-n}^{\infty} B_k(r)\right)\right)}{\sum_{k=N}^{\infty} A_k(R)}, \quad r \in [0,R], \tag{8}$$

where $A_k(r) := e^{-\lambda c_d r^d}\left(\lambda c_d r^d\right)^k/k!$ and $B_k(r) := e^{-\lambda c_d\left(R^d-r^d\right)}\left(\lambda c_d\left(R^d-r^d\right)\right)^k/k!$.

Proof: The complementary conditional cdf of R_n is given by

$$\begin{aligned}
\bar{F}'_{R_n}(r) &= \Pr\left(\Phi\left(b_d(o,r)\right) < n \mid \Phi\left(b_d(o,R)\right) \geq N\right) \\
&= \frac{\Pr\left(\Phi\left(b_d(o,r)\right) < n, \Phi\left(b_d(o,R)\right) \geq N\right)}{\Pr\left(\Phi\left(b_d(o,R)\right) \geq N\right)} \\
&\overset{(a)}{=} \frac{\sum_{k=0}^{n-1}\Pr\left(\Phi\left(b_d(o,r)\right) = k\right)\Pr\left(\Phi\left(b_d(o,R)\setminus b_d(o,r)\right) \geq N-k\right)}{\Pr\left(\Phi\left(b_d(o,R)\right) \geq N\right)} \\
&= \frac{\sum_{k=0}^{n-1} A_k(r)\left(1-\sum_{l=0}^{N-k-1} B_l(r)\right)}{\sum_{k=N}^{\infty} A_k(R)}, \qquad (9)
\end{aligned}$$

where (a) is obtained from the property that the number of points of the PPP in disjoint sets are independent of each other. It is easy to see that

$$\frac{\mathrm{d}}{\mathrm{d}r}A_k(r) = \begin{cases} \lambda d c_d r^{d-1}\left(A_{k-1}(r) - A_k(r)\right) & k > 0 \\ -\lambda d c_d r^{d-1} A_0(r) & k = 0 \end{cases} \qquad \text{(i)}$$

and

$$\frac{\mathrm{d}}{\mathrm{d}r}B_l(r) = \begin{cases} \lambda d c_d r^{d-1}\left(B_l(r) - B_{l-1}(r)\right) & l > 0 \\ \lambda d c_d r^{d-1} B_0(r) & l = 0. \end{cases} \qquad \text{(ii)}$$

Therefore, we have

$$\frac{\mathrm{d}}{\mathrm{d}r}\sum_{l=0}^{N-k-1} B_l(r) = \lambda d c_d r^{d-1} B_{N-k-1}(r). \qquad \text{(iii)}$$

The details of the remainder of the proof are straightforward but tedious and are omitted here. Since the pdf of the conditional distance distribution is $f'_{R_n} = -\mathrm{d}\bar{F}'_{R_n}/\mathrm{d}r$, one basically has to differentiate the numerator in (9), and after some simplifications using (i)-(iii), it will be seen that the conditional distance pdf is identical to (8). ■

Fig. 2 depicts the pdfs of the distances for one- and two-dimensional BPPs (from (4)) and compares it with the distance pdfs for a conditioned PPP with the same density.

When a large number of points are distributed randomly over a large area, their arrangement

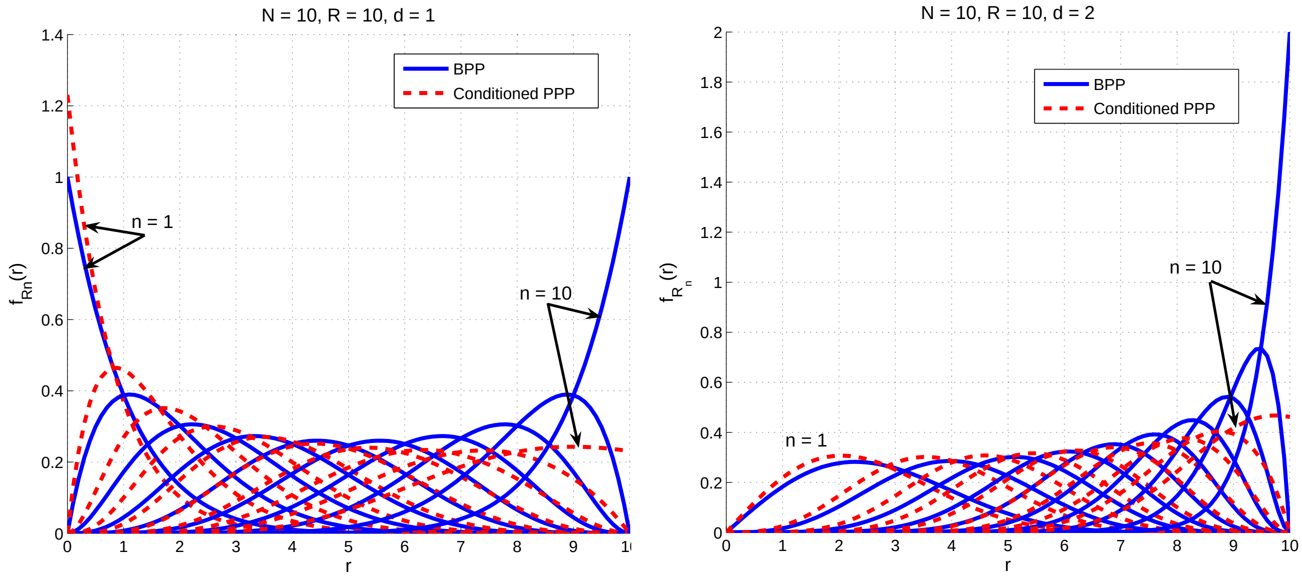

Fig. 2. Distance pdfs for each of the neighbors for one- and two- dimensional binomial and conditioned Poisson networks.

can be well approximated by an infinite homogeneous PPP. The PPP model for the nodal distribution is ubiquitously used for wireless networks and may be justified by claiming that nodes are dropped from an aircraft in large numbers; for mobile ad hoc networks, it may be argued that terminals move independently of each other. We now present a corollary to the earlier theorem, that reproduces a result from [6].

Corollary 2.4: In an infinite PPP with density λ on $\mathbb{R}^d$, the distance R_n, between a point and its n^{th} neighbor is distributed according to the generalized gamma distribution.

$$f_{R_n}(r) = e^{-\lambda c_d r^d} \frac{d(\lambda c_d r^d)^n}{r\Gamma(n)}, \quad r \in \mathbb{R}. \tag{10}$$

Proof: If the total number of points N tends to infinity in such a way that the density $\lambda = N/(c_d R^d)$ remains constant, then the BPP asymptotically (as $R \to \infty$) behaves as a PPP

[3]. Taking $R = \sqrt[d]{N/c_d\lambda}$ and applying the limit as $N \to \infty$, we obtain for a PPP,

$$\begin{aligned} f_{R_n}(r) &= \lim_{N\to\infty} \frac{d}{R} \frac{(1-p)^{N-n}\, p^{n-1/d}\Gamma(N+1)}{\Gamma(N-n+1)\Gamma(n)} \\ &= \frac{d}{r\Gamma(n)}(\lambda c_d r^d)^n \lim_{N\to\infty}\left(1-\frac{\lambda c_d r^d}{N}\right)^N \frac{N(N-1)\ldots(N-n+1)}{N^n} \\ &= e^{-\lambda c_d r^d}\frac{d(\lambda c_d r^d)^n}{r\Gamma(n)}. \end{aligned}$$

for $r \in \mathbb{R}$. This is an alternate proof to the one provided in [6]. ■

III. Moments of the Internode Distances

We now consider the isotropic d-dimensional BPP and use the internode distance pdf (4) to compute its moments. The γ^{th} moment of R_n is calculated as follows[4].

$$\begin{aligned} \mathbb{E}[R_n^\gamma] &= \frac{d}{R}\frac{1}{B(N-n+1,n)}\int_0^R \left[r^\gamma\left(\frac{r}{R}\right)^{nd-1}\left(1-\left(\frac{r}{R}\right)^d\right)^{N-n}\right]\mathrm{d}r. \\ &\overset{(a)}{=} \frac{R^\gamma}{B(N-n+1,n)}\int_0^1 t^{n+\gamma/d-1}(1-t)^{N-n}\mathrm{d}t \\ &= \frac{R^\gamma}{B(N-n+1,n)}B_x(n+\gamma/d, N-n+1)\big|_0^1 \\ &\overset{(b)}{=} \begin{cases} \frac{R^\gamma\Gamma(N+1)\Gamma(\gamma/d+n)}{\Gamma(n)\Gamma(\gamma/d+N+1)} & \text{if } n+\gamma/d>0 \\ \infty & \text{otherwise} \end{cases} \\ &= \begin{cases} R^\gamma n^{[\gamma/d]}/(N+1)^{[\gamma/d]} & \text{if } n+\gamma/d>0 \\ \infty & \text{otherwise,} \end{cases} \end{aligned} \tag{11}$$

where $B_x[a,b]$ is the incomplete beta function[5] and $x^{[n]} = \Gamma(x+n)/\Gamma(x)$ denotes the rising Pochhammer symbol notation. Here, (a) is obtained by making the substitution $r = Rt^{1/d}$ and

[4]Note that $\gamma \in \mathbb{R}$ in general, and is not restricted to being an integer.

[5]Mathematica: Beta[x, a, b].

(b) using the following identities:

$$B_0(a,b) \;=\; \begin{cases} 0 & \mathcal{R}e(a) > 0 \\ -\infty & \mathcal{R}e(a) \le 0, \end{cases}$$

and $B_1(a,b) = B(a,b)$ if $\mathcal{R}e(b) > 0$.

The expected distance to the n^{th} nearest node is thus

$$\mathbb{E}(R_n) = \frac{Rn^{[1/d]}}{(N+1)^{[1/d]}}, \tag{12}$$

and the variance of R_n is easily calculated as

$$\text{Var}[R_n] = \frac{R^2 n^{[2/d]}}{(N+1)^{[2/d]}} - \left(\frac{Rn^{[1/d]}}{(N+1)^{[1/d]}}\right)^2. \tag{13}$$

Remarks:

1) For one-dimensional networks, $\mathbb{E}[R_n] = Rn/(N+1)$. Thus, on an average, it is as if the points are arranged on a regular lattice. In particular, when N is odd, the middle point is located exactly at the center on average.
2) On the other hand, as $d \to \infty$, $\mathbb{E}[R_n] \to R$ and it is as if all the points are equidistant at maximum distance R from the origin.
3) In the general case, the mean distance to the n^{th} nearest node varies as $n^{1/d}$ for large n. This follows from the series expansion of the Pochhammer sequence [11]

$$n^{[q]} = n^q \left(1 - \mathcal{O}\left(1/n\right)\right).$$

   Also, for $d > 2$, the variance goes to 0 as n increases. This is also observed in the case of a Poisson network [6].
4) By the triangle inequality, the mean internodal distance between the i^{th} and j^{th} nearest

nodes from the origin, D_{ij}, is bounded as (assuming $i < j$)

$$\frac{R\left(j^{[1/d]} - i^{[1/d]}\right)}{(N+1)^{[1/d]}} < \mathbb{E}[D_{ij}] < \frac{R\left(i^{[1/d]} + j^{[1/d]}\right)}{(N+1)^{[1/d]}}.$$

5) For the special case of $\gamma/d \in \mathbb{Z}$, we obtain

$$\mathbb{E}[R_n^\gamma] = R^\gamma \binom{n+\gamma/d-1}{\gamma/d} \Big/ \binom{N+\gamma/d}{\gamma/d}.$$

IV. Distance Distributions in Regular Polygonal BPPs

In this section, we derive the pdf of the distance to the n^{th} nearest node from the origin, in BPPs distributed on a l-sided regular polygon W. Assume that the polygon is centered at the origin and $|W| = A$. Then, its inradius and circumradius are respectively given by

$$R_{\mathrm{i}} = \sqrt{\frac{A}{l}\cot\left(\frac{\pi}{l}\right)} \text{ and } R_{\mathrm{c}} = \sqrt{\frac{2A}{l}\csc\left(\frac{2\pi}{l}\right)}.$$

Also, let the total number of nodes be N and assume that no point of the process is at the origin.

Clearly, when $r \le R_{\mathrm{i}}$, $b_2(o,r)$ lies completely within the polygon and the number of points lying in it, $\Phi(b_2(o,r))$, is binomial distributed with parameters $n = N$ and $p = \pi r^2/A$.

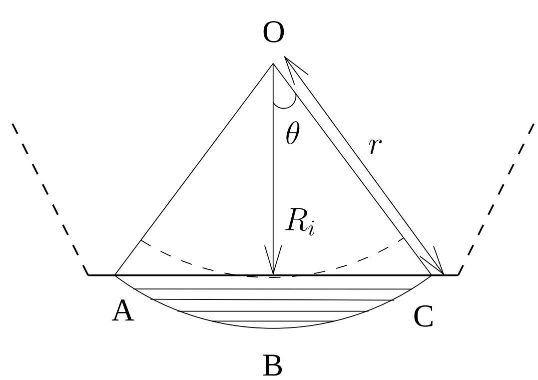

Fig. 3. Section of a l-sided regular polygon depicting one of its sides. O is the origin. For $R_{\mathrm{i}} < r \le R_{\mathrm{c}}$, the area of the shaded segment ABC is $r^2\theta - R_{\mathrm{i}}\sqrt{r^2 - R_{\mathrm{i}}^2}$.

When $R_{\mathrm{i}} < r \le R_{\mathrm{c}}$, $|W \cap b_2(o,r)|$ can be evaluated by considering the regions of the circle lying outside the polygon (see the shaded segment in Fig. 3). It is easy to see that $\Phi(b_2(o,r))$ follows a binomial distribution with parameters $n = N$ and

$$q = \frac{\pi r^2 - lr^2\theta + lR_{\mathrm{i}}\sqrt{r^2 - R_{\mathrm{i}}^2}}{A},$$

where $\theta = \cos^{-1}(R_{\mathrm{i}}/r)$. Following (3), we can write

$$f_{R_n}(r) = \begin{cases} \frac{2r\pi}{A}\frac{(1-p)^{N-n}p^{n-1}}{B(N-n+1,n)} & 0 < r \leq R_{\mathrm{i}} \\ \frac{2r(\pi-l\theta)}{A}\frac{(1-q)^{N-n}q^{n-1}}{B(N-n+1,n)} & R_{\mathrm{i}} < r \leq R_{\mathrm{c}} \\ 0 & R_{\mathrm{c}} < r. \end{cases} \tag{14}$$

Fig. 4 plots the pdf of the farthest neighbors in a BPP with 10 nodes, distributed on a l-sided regular polygon with $A = 100$, for $l = 3, 4, 5$ and $l \to \infty$.

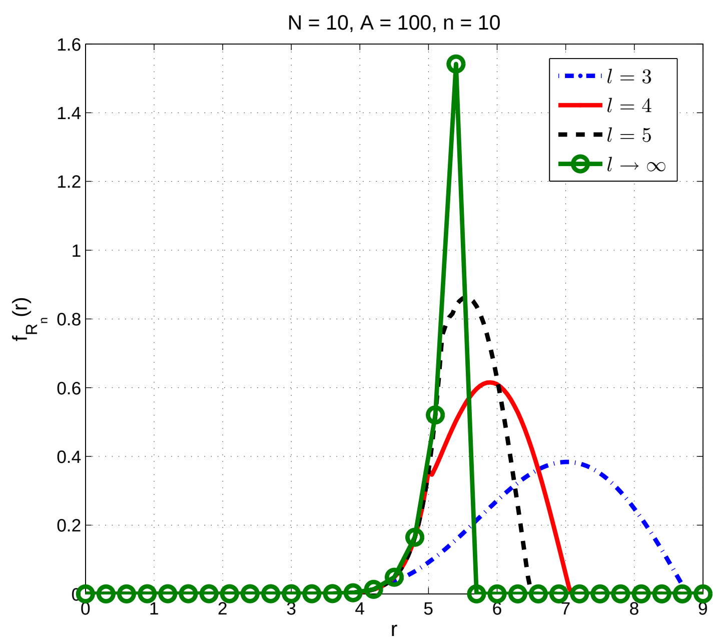

Fig. 4. The pdf of the distances to the farthest nodes from the origin in a BPP with 10 nodes and area 100 units, distributed on a l-sided polygon for $l = 3, 4$ and 5. The dotted line depicts the farthest neighbor distance in a circle ($l \to \infty$), for which $R_{\mathrm{i}} = R_{\mathrm{c}} = 10/\sqrt{\pi}$.

V. Applications to wireless networks

We now apply the results obtained in the previous section to wireless networks. For the system model, we assume a d-dimensional network over a ball $b_d(o, R)$, where N nodes are uniformly randomly distributed. Nodes are assumed to communicate with a base station (BS) positioned at the origin o. The attenuation in the channel is modeled by the large scale path loss function g with PLE α, i.e., $g(x) = \|x\|^{-\alpha}$. The channel access scheme is taken to be slotted ALOHA with contention parameter δ.

A. Energy Consumption

The energy that is required to successfully deliver a packet over a distance r in a medium with PLE α is proportional to r^α. Therefore, the average energy required to deliver a packet from the n^{th} nearest neighbor to the BS is given by (11), with $\gamma = \alpha$. This approximately scales as $n^{\alpha/d}$ when the routing is taken over single hops. When $\alpha < d$, it is more energy-efficient to use longer hops than when the PLE is greater than the number of dimensions.

B. Design of Routing Algorithms

The knowledge of nodal distances is also useful for the analysis and design of routing schemes for wireless networks. We illustrate this via an example wherein a greedy forwarding strategy that maximizes the expected progress of a packet towards its destination needs to be designed.

Consider the scenario where N nodes are uniformly distributed in a disk of radius R. Assume that several packets need to be forwarded from the BS to an arbitrarily chosen destination node D, which lies far away from the BS. We also assume that each node has a peak (transmit) power constraint of $P \ll R^\alpha$. Let us suppose that the nodes adopt a greedy forwarding strategy wherein each relay node $\mathrm{X_i}$ that gets a packet relays it to its farthest neighbor in a sector of angle ϕ ($0 \leq \phi \leq \pi$), i.e., along $\pm\phi/2$ around the $\mathrm{X_i}$-D axis (see Fig. 5). Evidently, for large ϕ, the direction of the farthest neighbor in the sector may be off the $\mathrm{X_i}$-D axis, while for small ϕ, there may not be enough nodes inside the sector. The natural question to ask is: what value of ϕ maximizes the expected progress of packets[6] towards the destination?

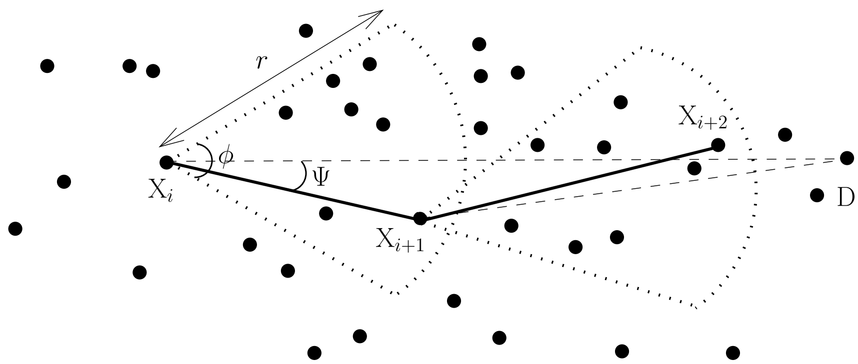

Fig. 5. The greedy forwarding strategy. Each relay X_i forwards the packet to its farthest neighbor lying inside the sector of radius ϕ around the X_i-D axis. The thick lines represent the path taken by the packet (through three arbitrary relays) for this particular realization.

A problem of similar flavor is studied in [12] for an interference-limited PPP, wherein the authors evaluate the optimal density of transmitters that maximizes the expected progress of a packet. In [13], the author determines the energy required to deliver a packet over a certain distance for various routing strategies in a PPP. In [14], the optimal transmission radius that maximizes the expected progress of a packet is determined for different transmission protocols in Poisson packet radio networks.

In order to evaluate the progress of a packet in the binomial network, we first note that if there are exactly k nodes in an arbitrary sector of angle ϕ and radius $r = P^{1/\alpha}$ (which is the range of transmission), the average distance to the farthest (kth) neighbor in that sector is the same as (12)[7], with $n = k$, $R = r$ and $d = 2$. We also know that the number of nodes lying in that sector is binomial with parameters N and $(r^2\phi/2\pi R^2)$. Thus, the mean distance to the

[6]We define the progress of a packet from a relay node X_i as the effective distance travelled along the X_i-D axis.

[7]This follows from the observation that in (1), the distance distributions depend only on $p = |b_2(x, r) \cap W|/|W|$, and the values of p for the sector and the circle are the same.

farthest neighbor in the considered sector can be written as

$$\sum_{k=1}^{N} \binom{N}{k} \left(\frac{r^2\phi}{2\pi R^2}\right)^k \left(1 - \frac{r^2\phi}{2\pi R^2}\right)^{N-k} \frac{2rk}{2k+1}. \tag{15}$$

Note that the sectors emanating from nodes $\mathrm{X_i}$ and $\mathrm{X_{i+1}}$ overlap partially, and also, the total number of nodes is fixed; therefore the mean distance to the farthest neighbor, $\mathbb{E}[X']$, is actually upper-bounded by (15). However, since we consider the farthest neighbors, the sectoral overlap is small.

Next, let Ψ denote the angle between the line connecting $\mathrm{X_i}$ to its farthest neighbor ($\mathrm{X_{i+1}}$) and the $\mathrm{X_i}$-D axis. Since the nodal distribution is uniformly random, Ψ is uniformly distributed on $[-\phi/2, \phi/2]$. The expected progress of a packet is $\mathbb{E}[X] = \mathbb{E}[X']\mathbb{E}[\cos(\Psi)]$ since Ψ and X' are independent of each other, and is upper-bounded as

$$\mathbb{E}[X] \leq \frac{2}{\phi} \sin\left(\frac{\phi}{2}\right) \sum_{k=1}^{N} \frac{2P^{1/\alpha}k}{2k+1} \binom{N}{k} \left(\frac{P^{2/\alpha}\phi}{2\pi R^2}\right)^k \left(1 - \frac{P^{2/\alpha}\phi}{2\pi R^2}\right)^{N-k}. \tag{16}$$

The optimum value of ϕ that maximizes the progress of packets can be numerically determined from (16).

Fig. 6 plots the expected progress of a packet (upper bound) versus ϕ for several values of N using (16), and compares it with the empirical value, obtained via simulation. We see that the bound is reasonably tight, in particular at lower N. The optimum values of ϕ are also marked in the figure.

C. Localization

In wireless networks, localization is an integral component of network self-configuration. Nodes that are able to accurately estimate their positions can support a rich set of geographically aware protocols and report the regions of detected events. Localization is also useful for performing energy-efficient routing in a decentralized fashion.

In this section, we investigate conditional distance distributions and study their usefulness to

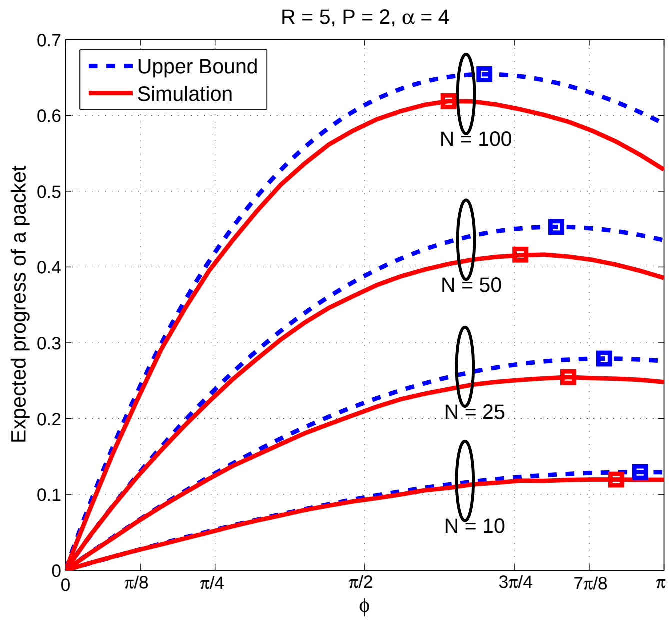

Fig. 6. The expected progress of a packet (empirical and upper bound) for various values of N. The square markers correspond to the optimum values of ϕ that maximize the packet's progress.

localization algorithms. We consider the scenario wherein a few nodes can estimate or even precisely measure their distances from the BS. What can be said about the distance statistics of the other nodes given this information?

Suppose we know that the k^{th} nearest neighbor is at distance s from the center[8]. Then, clearly, the first $k-1$ nodes are uniformly randomly distributed in $b_d(o,s)$ while the more distant nodes are uniformly randomly distributed in $b_d(o,R)\setminus b_d(o,s)$. Following (4), the distance distributions of the first $k-1$ nearest neighbors from the origin can be written as

$$f_{R_n}(r \mid R_k = s) \;=\; \frac{d}{s}\frac{B(n-1/d+1,k-n)}{B(k-n,n)}\beta\left(\left(\frac{r}{s}\right)^d; n-\frac{1}{d}+1, k-n\right), \quad n<k$$

for $0 \le r \le s$, which again follows a generalized beta distribution.

[8]Based on the RSS from the base station, perhaps averaged over a period of time to eliminate the variations due to fading, nodes can determine how many other nodes are closer to the transmitter than they are. This way, a node would find out that it is the k^{th} nearest neighbor to the base station.

For the remaining nodes i.e., for $n > k$, we have in $r \in [s, R]$,

$$\begin{aligned} f_{R_n}(r \mid R_k = s) &= -\frac{\mathrm{d}}{\mathrm{d}r} I_{1-q}(N-n+1, n-k) \\ &= \frac{dr^{d-1}}{R^d - s^d} \frac{(1-q)^{N-n} q^{n-k-1}}{B(N-n+1, n-k)} \end{aligned}$$

where $q = (r^d - s^d)/(R^d - s^d)$.

The moments of R_n are also straightforward to obtain. Following (11), we see that for $n < k$ and $n + \alpha/d > 0$,

$$\mathbb{E}[R_n^\alpha \mid R_k = s] = \frac{s^\alpha n^{[\alpha/d]}}{(k+1)^{[\alpha/d]}}. \tag{17}$$

For $n > k$, we have

$$\begin{aligned} \mathbb{E}[R_n^\alpha \mid R_k = s] &= \int_s^R \frac{dr^{\alpha+d-1}}{R^d - s^d} \frac{(1-q)^{N-n} q^{n-k-1}}{B(N-n+1, n-k)} \mathrm{d}r \\ &= \frac{1}{B(N-n+1, n-k)} \int_0^1 q^{n-k-1} (1-q)^{N-n} \left(q\left(R^d - s^d\right) + s^d\right)^{\alpha/d} \mathrm{d}q \\ &= \frac{s^\alpha}{(n-k)B(N-n+1, n-k)} F_1\left(n-k; n-N, -\frac{\alpha}{d}; n-k+1; 1, 1 - \frac{R^d}{s^d}\right), \end{aligned}$$

where $F_1[a; b_1, b_2; c; x, y]$ is the Appell hypergeometric function of two variables[9].

Often, it is easiest to measure the nearest-neighbor distance. Give this distance as s, we have for $n > 1$,

$$f_{R_n}(r \mid R_1 = s) = \frac{dr^{d-1}}{R^d - s^d} \frac{\left(1 - \left(\frac{r^d - s^d}{R^d - s^d}\right)\right)^{N-n} \left(\frac{r^d - s^d}{R^d - s^d}\right)^{n-2}}{B(N-n+1, n-1)}$$

for $r \in [s, R]$. Also, the mean conditional distances to the remaining neighbors are

$$\mathbb{E}[R_n \mid R_1 = s] = \frac{s}{(n-1)B(N-n+1, n-1)} F_1\left(n-1; n-N, -\frac{1}{d}; n-1; 1, 1 - \frac{R^d}{s^d}\right).$$

Fig. 7 plots the mean conditional distances in a network with 10 nodes when the nearest-neighbor

[9] Mathematica: AppellF1[a, b1, b2, c, x, y].

distance is unity.

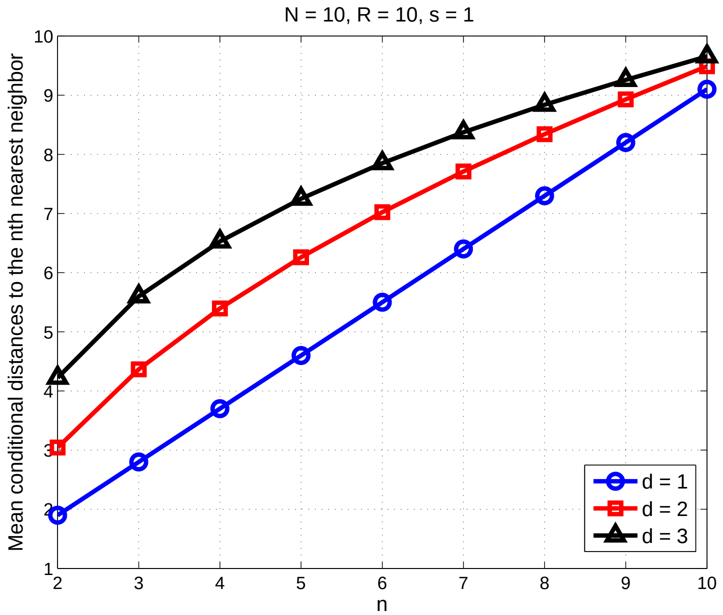

Fig. 7. The mean conditional distances of the higher-order neighbors in a binomial network with 10 nodes and $d = 1, 2, 3$, when it is known that the nearest neighbor is at unit distance away from the base station.

D. *Interference*

In order to accurately determine network parameters such as outage, throughput or transmission capacity, the interference in the system I needs to be known.

Let $T_n \in \{0, 1\}$, $1 \leq n \leq N$ denote the random variable representing whether the n^{th} nearest node to the BS transmits or not, in a particular time slot. With the channel access scheme being ALOHA, these are i.i.d. Bernoulli variables (with parameter δ).

The mean interference as seen at the center of the network is given by

$$\begin{aligned}\mu_I &= \mathbb{E}\left[\sum_{n=1}^{N}(T_n R_n^{-\alpha})\right] = \sum_{n=1}^{N}\mathbb{E}[T_n]\mathbb{E}\left[R_n^{-\alpha}\right],\\ &= \delta\sum_{n=1}^{N}\mathbb{E}\left[R_n^{-\alpha}\right].\end{aligned}$$

Setting $\gamma = -\alpha$ and $n = 1$ in (11), we can conclude that the mean interference is infinite for $d \leq \alpha$. This is due to the nearest interferer. Even the mean interference from just the n^{th} nearest transmitter is infinite if $\alpha \geq nd$. When the number of dimensions is greater than the PLE, we have

$$\mu_I = \frac{\delta R^{-\alpha}\Gamma(N+1)}{\Gamma(N+1-\alpha/d)} \sum_{n=1}^{N} \frac{\Gamma(n-\alpha/d)}{\Gamma(n)}.$$

One can inductively verify that

$$\sum_{n=1}^{k} \frac{\Gamma(n-\alpha/d)}{\Gamma(n)} = \frac{\Gamma(k-\alpha/d)}{\Gamma(k)} \frac{k-\alpha/d}{1-\alpha/d} \quad \forall k \in \mathbb{Z}, \tag{18}$$

and we obtain after some simplifications,

$$\mu_I = \frac{N\delta d R^{-\alpha}}{d-\alpha}, \quad d > \alpha. \tag{19}$$

The unboundedness of the mean interference at practical values of d and α (i.e., $d < \alpha$) actually occurs due to the fact that the path loss model we employ breaks down for very small distances, i.e., it exhibits a singularity at $x = 0$. One way to overcome this issue is to impose a guard zone of radius ϵ around every receiver. In other words, every receiver has an exclusion zone of radius ϵ around it and the nodes lying within it are not allowed to transmit.

Since the average number of nodes in the ball $b(o,\epsilon)$ is $N\epsilon^d/R^d$, we obtain the mean interference in this case to be

$$\begin{aligned} \mu_I &= \frac{NpdR^{-\alpha}}{d-\alpha} - \frac{N\epsilon^d \delta d \epsilon^{-\alpha}}{R^d(d-\alpha)} \\ &= \frac{N\delta d(R^{d-\alpha} - \epsilon^{d-\alpha})}{R^d(d-\alpha)}, \quad \forall d \neq \alpha. \end{aligned} \tag{20}$$

Taking limits, we obtain $\mu_I = N\delta d \ln(R/\epsilon)/R^d$ when $d = \alpha$.

E. Outage Probability and Connectivity

Assuming that the system is interference-limited, an outage $\mathcal{O}$ is defined to occur if the SIR at the BS is lower than a certain threshold Θ. Let the desired transmitter be located at unit distance from the origin, transmit at unit power and also not be a part of the original point process. Then, the outage probability is $\Pr(\mathcal{O}) = \Pr[1/I < \Theta]$.
Considering only the interference contribution from the nearest neighbor to the origin, a simple lower bound is established on the outage probability as

$$\begin{aligned} \Pr(\mathcal{O}) &\geq \Pr\left(T_1 R_1^{-\alpha} > 1/\Theta\right) \\ &= \delta \Pr\left(R_1 < \Theta^{1/\alpha}\right) \\ &= \begin{cases} \delta\left(1 - \left(1 - \frac{\Theta^{d/\alpha}}{R^d}\right)^N\right) & \Theta \leq R^\alpha \\ \delta & \Theta > R^\alpha. \end{cases} \end{aligned} \tag{21}$$

The empirical values of success probabilities and their upper bounds (21) are plotted for different values of N in Fig. 8. As the plot depicts, the bounds are tight for lower values of N and Θ, and therefore we conclude that the nearest neighbor contributes most of the network interference. However, as α decreases, the bound gets looser since the contributions from the farther neighbors are also increased.

Next we study the connectivity properties of the binomial network, assuming that interference can be controlled such that the system is noise-limited. Define a node to be connected to the origin if the SNR at the BS is greater than a threshold Θ. Let the nodes transmit at unit power and assume noise to be AWGN with variance N_0. In the absence of interference, the probability

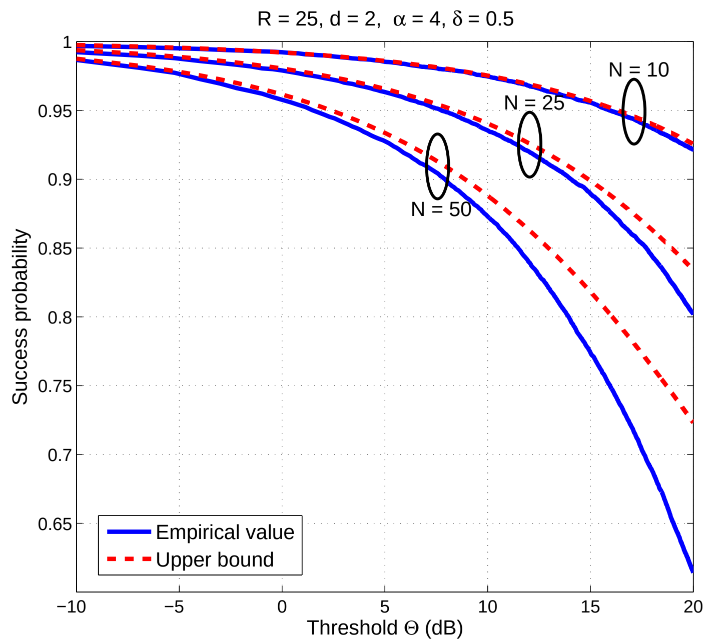

Fig. 8. Comparison of exact success probabilities versus their upper bounds for different values of the system parameters.

that the BS is connected to its n^{th} nearest neighbor is

$$\begin{aligned} &= \Pr(R_n^{-\alpha} > N_0\Theta) \\ &= 1 - \Pr(R_n > (N_0\Theta)^{-1/\alpha}) \\ &= \begin{cases} 1 - I_{1-p'}(N-n+1, n) & \Theta > R^{-\alpha}/N_0 \\ 1 & \Theta \le R^{-\alpha}/N_0, \end{cases} \end{aligned} \tag{22}$$

where $p' = \left((N_0\Theta)^{-1/\alpha}/R\right)^d$. Fig. 9 plots the connection probability in a two-dimensional binomial network with 25 nodes.

The mean number of nodes that are connected to the BS is $N \min\{1, \left(\frac{(N_0\Theta)^{-1/\alpha}}{R}\right)^d\}$.

F. *Other Applications*

We now list a few other areas where knowledge of the distance distributions is useful.

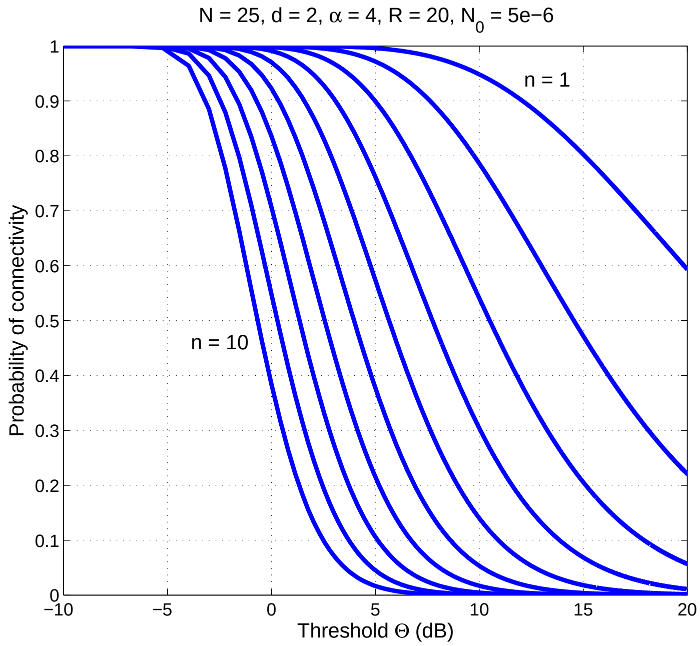

Fig. 9. The probability of the n^{th} nearest neighbor, $n = 1, 2, \ldots, 10$, being connected to the BS for a binomial network with 25 nodes.

- *Routing*: The question of whether to route over smaller or longer hops is an important, yet a nontrivial issue [15], [16], and it gets more complicated in the presence of interference in the network. The knowledge of internodal distances is necessary for evaluating the optimum hop distance and maximizing the progress of a packet towards its destination.
- *Path loss exponent estimation*: The issue of PLE estimation is a very important and relevant problem [17]. Several PLE estimation algorithms are based on received signal strength techniques, which require the knowledge of distances between nodes.

VI. Concluding Remarks

We argue that the Poisson model for nodal distributions in wireless networks is not accurate in many practical situations and instead consider the often more realistic binomial network model. We derive exact analytical expressions for the pdfs of the internodal distances in a network where a known number of nodes are independently distributed in a compact set. Specializing to the case of an isotropic random network, we show that the distances between nodes follow a generalized beta distribution and express the moments of these random variables in closed-form. We also derive the distribution of the internodal distances for the BPP distributed on a regular polygon. Our findings have applications in several problems relating to wireless networks such as energy consumption, design of efficient routing and localization algorithms, connectivity, interference characterization and outage evaluation.

REFERENCES

[1] T. S. Rappaport, *Wireless Communications - Principles and Practice*, Prentice Hall, 1991.

[2] I. F Akyildiz, W. Su, Y. Sankarasubramaniam and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, Vol. 40, Iss. 8, pp. 102-114, Aug. 2002.

[3] D. Stoyan, W. S. Kendall and J. Mecke, “Stochastic Geometry and its Applications,” Wiley & Sons, 1978.

[4] L. E. Miller, ”Distribution of link distance in a wireless network,” *Journal of Research of National Institute of Standards and Technology*, Vol. 106, pp. 401-412, Mar./Apr. 2001.

[5] C. Rose, “Mean internodal distance in regular and random multihop networks,” *IEEE Transactions on Communications*, Vol. 40, No. 8, pp. 1310-1318, Aug. 1992.

[6] M. Haenggi, “On distances in uniformly random networks,” *IEEE Transactions on Information Theory*, Vol. 51, No. 10, pp. 3584-3586, Oct 2005.

[7] S. Vural and E. Ekici, Probability distribution of multihop distance in one-dimensional sensor networks, *Computer Networks Journal (Elsevier)*, Vol. 51, No. 13, pp. 3727-3749, Sep. 2007.

[8] L. E. Miller, “Joint distribution of link distances,” *Conference on Information Sciences and Systems*, Mar. 2003.

[9] C-C. Tseng, H-T. Chen and K-C. Chen, “On the distance distributions of the wireless ad hoc networks,” *IEEE Vehicular Technology Conference*, Vol. 2, pp. 772-776, May 2006.

[10] P. Kumaraswamy, “A generalized probability density function for double-bounded random processes,” *Journal of Hydrology*, Vol. 46, No. 1/2, pp. 79-88, 1980.

[11] R. L. Graham, D. E. Knuth and O. Patashnik, “Concrete mathematics: A foundation for computer science,” 2nd ed., Addison-Wesley, 1994.

[12] F. Baccelli, B. Błaszczyszyn and P. Mühlethaler, “An ALOHA protocol for multihop mobile wireless networks,” *IEEE Transactions on Information Theory*, Vol. 52, No. 2, pp. 421-436, Feb. 2006.

[13] M. Haenggi, “On routing in random Rayleigh fading networks,” *IEEE Transactions on Wireless Communications*, Vol. 4, No. 4, pp. 1553-1562, July 2005.

[14] H. Takagi and L. Kleinrock, “Optimal transmission ranges for randomly distributed packet radio terminals,” *IEEE Transactions on Communications*, Vol. COM-32, No. 3, pp. 246-257, Mar. 1984.

[15] M. Haenggi and D. Puccinelli, “Routing in ad hoc networks: A case for long hops,” *IEEE Communications Magazine*, Vol. 43, Iss. 10, pp. 93-101, Oct. 2005.

[16] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, and T. Fuja, “Bandwidth- and power-efficient routing in linear wireless networks,” *Joint Special Issue of IEEE Transactions on Information Theory and IEEE Transactions on Networking*, Vol. 52, Iss. 6, pp. 2624-2633, June 2006.

[17] S. Srinivasa and M. Haenggi, “Path loss estimation in large wireless networks,” in preparation for submission to *IEEE Transactions on Wireless Communications*. Available at `http://aps.arxiv.org/abs/0802.0351`.