

Path Loss Exponent Estimation in a Large Field of Interferers

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 channels, the path loss exponent (PLE) has a strong impact on the quality of links, and hence, it needs to be accurately estimated for the efficient design and operation of wireless networks. In this paper, we address the problem of PLE estimation in large wireless networks, which is relevant to several important issues in networked communications such as localization, energy-efficient routing, and channel access. We consider a large ad hoc network where nodes are distributed as a homogeneous Poisson point process on the plane and the channels are subject to Nakagami- m fading. We propose and discuss three distributed algorithms for estimating the PLE under these settings which explicitly take into account the interference in the network. In addition, we provide simulation results to demonstrate the performance of the algorithms and quantify the estimation errors. We also describe how to estimate the PLE accurately even in networks with spatially varying PLEs and more general node distributions.

I. INTRODUCTION

The wireless channel presents a formidable challenge as a medium for reliable high-rate communication. It is responsible not only for the attenuation of the propagated signal but also causes unpredictable spatial and temporal variations in this loss due to user movement and changes in the environment. In order to capture all these effects, the path loss for RF signals is commonly represented as the product of a deterministic distance component (large-scale path loss) and a randomly-varying component (small-scale fading) [1]. The large-scale path loss model assumes that the received signal strength falls off with distance

according to a power law, at a rate termed the path loss exponent (PLE). Fading describes the deviations of the received signal strength from the power-law decay due to shadowing and the constructive and destructive addition of its multipath components. While the small-scale fading behavior of the wireless channel can be well-represented using stochastic processes¹ [2], it is critical to accurately estimate the PLE for the efficient design and operation of wireless networks.

This estimation problem is non-trivial even for a single link due to the existence of multipath propagation and thermal noise. For large wireless networks without infrastructure, the problem is further complicated due to the following reasons: First, the achievable performance of a typical ad hoc or sensor network is not only susceptible to noise and fading, but also to interference due to the presence of simultaneous transmitters. Dealing with fading and interference simultaneously is a major difficulty in the estimation problem. Second, the distances between nodes themselves are subject to uncertainty. Often, the distribution of the underlying point process can be statistically determined, but precise locations of the nodes are harder to measure. In such cases, we will need to consider the fading and distance ambiguities jointly, i.e., define a spatial point process that incorporates both.

In this paper, we present three distributed algorithms to accurately estimate the channel's PLE for large wireless networks with uniformly randomly distributed nodes in the presence of fading, noise and interference, based solely on received signal strength measurements. We also provide simulation results to illustrate the performance of the methods and study the estimation error. Additionally, we briefly describe how to accurately estimate the PLE in environments with spatially varying PLE values and for more general node distributions. The remainder of the paper is structured as follows. Section II provides a few examples to motivate the importance of knowing the PLE for the analysis and design of communication systems and discusses the prior work on the estimation problem. Section III presents the system and channel models. Section IV describes three distributed algorithms for PLE estimation, each based on a specific network characteristic, and provides the mean squared error (MSE) performance of the algorithms based on simulation. Section V suggests two simple ways to improve the accuracy of the estimation algorithms. Section VI briefly discusses the sensitivity of the algorithms to variations in the system model, and Section VII concludes the paper.

¹While modeling wireless channels, the small-scale fading amplitude is often assumed to be distributed as a Rayleigh, Rician or a Nakagami- m random variable.

II. MOTIVATION AND RELATED WORK

A. Motivation

In this section, we illustrate the importance of knowing the PLE for efficient design and operation of wireless networks. Though it is often assumed in analysis and design problems that the value of the PLE is known a priori, this is not true in practice, and hence, the PLE needs to be accurately estimated during the network initialization phase.

The PLE estimation problem is closely related to that of localization, which is an integral component of network self-configuration. When bestowed with the ability to detect their positions, sensor nodes deployed in an ad hoc fashion can support a rich set of geographically aware protocols and accurately report the regions of detected events. Detailed knowledge of the nodes' locations is also needed for performing energy-efficient routing of packets in the network. An important class of localization algorithms is based on received signal strength (RSS) measurements [3], [4], but it needs accurate estimates of the PLE to perform well. Another fundamental issue in sensor networks is the sensing problem that deals with how well a target area or a phenomenon can be monitored. Studying network characteristics such as connectivity is important for such applications and requires accurate estimates of the PLE [5], [6].

A good knowledge of the PLE is also essential for designing efficient networks. In [7], the authors discuss capacity results for TDMA-based linear networks and show that the optimum number of hops to achieve a desired end-to-end rate strongly depends on the PLE. For example, when the desired (bandwidth-normalized) spectral efficiency exceeds the PLE, single-hop transmission outperforms multihopping. Many of the existing results on capacity scaling for large ad hoc networks strongly depend on the PLE as well. With γ being the PLE, the best known achievability result [8] states that for a network having n uniformly randomly located nodes on the plane, the capacity scales as $n^{2-\gamma/2}$ for $2 \leq \gamma < 3$ and as \sqrt{n} for $\gamma \geq 3$. Depending on the value of the PLE, appropriate routing strategies (nearest-neighbor hopping or hierarchical cooperative schemes) may be implemented to reach the maximally achievable scaling of the system throughput.

Energy consumption in wireless networks is a crucial issue that needs to be addressed at all the layers of the communication system. In [9], the author analyzes the energy consumed for several routing strategies that employ hops of different lengths in a large network with uniformly randomly distributed nodes. Using the results therein, we demonstrate that a good knowledge of the PLE is necessary for efficient routing. Consider the following two simple schemes where communication is assumed to occur only in a sector

ϕ around the source-destination axis.

- 1) Route across n nearest neighbor hops in a sector ϕ , i.e., the next-hop node is the nearest neighbor that lies within $\pm\phi/2$ of the axis to the destination.
- 2) Transmit directly to the n' 'th nearest neighbor in the sector ϕ . Here, n' is chosen in a way that the expected progress is the same for both schemes.

From [9], the ratio of the consumed energies for the two schemes is obtained as

$$\frac{E_1}{E_2} = \frac{n^2\Gamma(1 + \gamma/2)\Gamma(n')}{\Gamma(n' + \gamma/2)},$$

where $\Gamma(\cdot)$ represents the gamma function and $n' = \frac{\pi}{4}(n^2 - 1) + 1$. Observe that the PLE plays an important role in determining energy-efficient routing strategies. In particular, when γ is small, scheme 2 consumes less energy while relaying is more beneficial at high PLE values.

The performance of contention-based systems such as slotted ALOHA is very sensitive to the contention probability p , hence it is critical to choose the optimal operating point of the system. The value of the contention parameter is determined based on various motives such as maximizing the network throughput [10] or optimizing the spatial density of progress [11, Eqn. 5.6]. These quantities also greatly depend on the PLE, and therefore the optimal value of the contention probability can be chosen only after estimating γ .

B. Review of Literature

In this section, we survey some of the existing PLE estimation methods in the literature. Most authors have assumed a simplified channel model consisting only of a large-scale path loss component and a shadowing counterpart, but we are not aware of any prior work that has considered fading, and, most importantly, interference in the system model. Therefore, much of the past work on PLE prediction has focused mainly on RSS-based localization techniques. However, ignoring interference in the system model is not realistic, in particular since PLE estimation needs to be performed before the network is organized.

Estimation based on a known internode distance probability distribution is discussed in [12]. The authors assume that the distance distribution $p_R(r)$ between neighboring nodes is known or can be determined easily. With the transmit power equal to P_0 [dBm] (assume this is a constant for all nodes), the theoretic mean RSS averaged over neighboring node pairs (in the absence of interference) is $\bar{P} =$

$P_0 \mathbb{E}_R [R^{-\gamma}]$. E.g., if the internodal distance distribution is (i.i.d.) Rayleigh² with mean R_0 , then we have $\bar{P} = P_0 (2R_0/\sqrt{\pi})^{-\gamma} \Gamma(1 - \gamma/2)$. The value of γ is estimated by equating \bar{P} to the empirical mean value of the RSS measured over several node pairs.

If the nearest-neighbor distribution is in a complicated form that is not integrable, an idea similar to the quantile-quantile plot can be used [12]. For cases where it might not be possible to obtain the neighbor distance distribution, the idea of estimating γ using the concept of the Cayley-Menger determinant or the pattern matching technique [12] is useful.

In [13], the authors consider a network where the path loss between a few low-cost sensors is measured and stored for future use. They then propose an algorithm that employs interpolation techniques to estimate the path loss between a sensor and any arbitrary point in the network. In [14], a PLE estimator based on the method of least squares is discussed and used in the design of an efficient handover algorithm. However, as described earlier, the situation is completely different when interference and fading are considered and these purely RSS-based estimators cannot be used. Also, none of the prior estimation algorithms works in a fully distributed nature.

III. SYSTEM MODEL

We consider an infinite planar ad hoc network, where nodes are distributed as a homogeneous Poisson point process (PPP) Φ of density λ . Therefore, the number of points lying in a Borel set B , denoted by $\Phi(B)$, is Poisson-distributed with mean $\lambda \nu_2(B)$, where $\nu_2(\cdot)$ is the two-dimensional Lebesgue measure (area). Also, the number of points in disjoint sets are independent random variables. The PPP model for the node distribution is ubiquitously used and may be justified by claiming that sensor 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.

The attenuation in the channel is modeled as the product of the large-scale path loss with exponent γ and a flat block-fading component. Also, the noise is taken to be AWGN with mean power N_0 . To obtain concrete results, the fading amplitude H is taken to be Nakagami- m distributed. Letting $m = 1$ results in the well-known case of Rayleigh fading, while lower and higher values of m signify stronger and weaker fading scenarios respectively. The case of no fading is modeled by setting $m \rightarrow \infty$. When dealing with

²When nodes are arranged uniformly randomly on the plane, the nearest-neighbor distance is Rayleigh distributed [9].

signal powers, we use the power fading variable denoted by $G = H^2$. The pdf of G is given by [15]

$$p_G(x) = \frac{m^m}{\Gamma(m)} x^{m-1} \exp(-mx), \quad m \geq 1/2, \quad (1)$$

and its moments are

$$\mathbb{E}_G[G^n] = \frac{\Gamma(m+n)}{m^n \Gamma(m)}, \quad n \in \mathbb{R}^+. \quad (2)$$

Note that G captures the random deviation from the large-scale path loss, thus $\mathbb{E}_G[G] = 1$.

Since the PLE estimation is usually performed during network initialization, it is reasonable to assume that the transmissions in the system during this phase are uncoordinated. Therefore, we take the channel access scheme to be ALOHA. We shall see in Section V that even if other MAC schemes were available, ALOHA is a good choice of MAC protocol since employing it minimizes the spatio-temporal correlations in the interference, which helps improve the estimation accuracy. We denote the ALOHA contention probability by a constant p . Therefore, nodes independently decide to transmit with probability p or remain idle with probability $1 - p$ in any time slot³. Consequently, the set of transmitters at any given moment forms a PPP Φ' of density λp . Also, since there is no information available for power control, we assume that all the transmit powers are equal to unity. Then, the interference at node y is given by

$$I_{\Phi}(y) = \sum_{z \in \Phi'} G_{zy} \|z - y\|^{-\gamma},$$

where G_{zy} is the fading gain of the channel from z to y , and $\|\cdot\|$ denotes the Euclidean distance.

We define the communication from the transmitter at x to the receiver at y to be successful if and only if the signal-to-noise and interference ratio (SINR) at y is larger than a threshold Θ , which depends on the modulation and coding scheme and the receiver structure. Mathematically speaking, an outage at a receiver at y occurs if and only if

$$\frac{G_{xy} \|x - y\|^{-\gamma}}{N_0 + I_{\Phi' \setminus \{x\}}(y)} \leq \Theta, \quad (3)$$

where $I_{\Phi' \setminus \{x\}}(y)$ denotes the interference in the network at y and x is the desired transmitter.

IV. PATH LOSS EXPONENT ESTIMATION

This section describes three fully distributed algorithms for PLE estimation, each based on a certain network characteristic, and provides simulation results on the estimation errors. The first algorithm uses

³The beginning and ending times of a slot is based on the individual node's clock cycle. Thus, time slots across different nodes need not (and in general, will not) be synchronized. We will only assume that the duration of the slots are the same.

the mean interference value and assumes the network density λ to be known. Algorithms 2 and 3 are based on outage probabilities and the network's connectivity properties, respectively, and do not require knowledge of λ or the Nakagami parameter m .

The PLE estimation problem is essentially tackled by equating the empirical (observed) values of the aforementioned network characteristics to the theoretically established ones to obtain the estimate $\hat{\gamma}$. In each time slot, nodes either transmit (w.p. p) or listen to record measurements (w.p. $1 - p$). After obtaining the required measurement values over several time slots, the estimation process can be performed at each node in a distributed fashion.

The simulation results provided in this paper are obtained using MATLAB. To analyze the mean error performance of the algorithms, we use estimates resulting from 50000 different realizations of the PPP. Each PPP is generated by distributing a Poisson number of points uniformly randomly in a 50×50 square with density 1. To avoid border effects, we use the measurements recorded at the node lying closest to the center of the network. The accuracy of the algorithms is characterized using the relative MSE, defined as $\mathbb{E} [(\hat{\gamma} - \gamma)^2] / \gamma$. The contention probability is taken to be $p = 0.05$ in each case⁴, and $N_0 = -25$ dBm.

A. Algorithm 1: Estimation Using the Mean Interference

In many situations, the network density is a design parameter and hence known. In other cases, it is possible to estimate the density (see [16, Sec. 2.7] and the references therein for a survey of the estimation methods for a PPP). This section discusses a PLE estimation algorithm that uses the knowledge of the density λ .

A simple technique to infer the PLE γ when the Nakagami parameter m is unknown is based on the mean interference. According to this method, nodes simply need to record the strength of the received power that they observe and use it to estimate γ . We first state existing theoretic results and subsequently describe how the estimation can be performed in practice.

For the PPP network running the slotted ALOHA protocol, the n^{th} cumulant of the interference resulting from transmitters in an annulus of inner radius A and outer radius B around the receiver node is given by [17]

$$C_n = 2\pi\lambda p \mathbb{E}_G[G^m] \frac{B^{2-n\gamma} - A^{2-n\gamma}}{2 - n\gamma}. \quad (4)$$

⁴This value of p was found to be suitable to obtain several quasi-different realizations of the PPP Φ' and helped obtain accurate estimates in a reasonable number of time slots.

In particular, we can consider only the case $\gamma > 2$ (a fair assumption in a wireless scenario) and let B be large (considering the entire network) so that the mean interference is

$$\mu_I = C_1 = 2\pi\lambda p \frac{A^{2-\gamma}}{\gamma-2}, \quad (5)$$

which is independent of m . Consequently, the mean received power is $\mu_R = \mu_I + N_0$. Note from (5) that the mean received power is infinite for $A = 0$. However, the large-scale path loss model is valid only in the far-field of the antenna and breaks down for very small distances. Denote the (known) near-field radius by a positive constant A_0 .

The algorithm based on the interference moments matches the observed and theoretic values of the mean received power and is described as follows.

- Record the strengths of the received powers R_1, \dots, R_N at any arbitrarily chosen node during N time slots and evaluate the empirical mean received power $(1/N) \sum_{i=1}^N R_i$.
- Equate the empirical mean to the theoretical value of μ_R (with $A = A_0$) and estimate γ using a look-up table and the known values of p , N_0 and (known or estimated density) $\hat{\lambda}$.

Fig. 1 depicts the relative MSE values of the estimated PLE $\hat{\gamma}$ for different γ and N values. The estimates are seen to be accurate over a wide range of PLE values, in particular when the PLE is small. Furthermore, the MSE is seen to converge within just about 1000 time slots, i.e., in a few seconds or less, depending on the hardware used.

The estimate $\hat{\gamma}$ can be used along with $\hat{\lambda}$ to also estimate the Nakagami parameter m . Indeed, from (4), the variance of the interference is

$$\sigma_I^2 = C_2 = 2\pi\lambda p \left(1 + \frac{1}{m}\right) \frac{A^{2-2\gamma}}{2\gamma-2}. \quad (6)$$

Since noise and interference are independent of each other, the variance of the received power is $\sigma_R^2 = \sigma_I^2 + \sigma_N^2$, where σ_N^2 denotes the (known) variance of the noise power. Therefore, an estimate of the fading parameter m is obtained by inverting (6) as

$$\hat{m} = \left(\frac{(\sigma_R^2 - \sigma_N^2)(\hat{\gamma} - 1)}{\pi \hat{\lambda} p A_0^{2-2\hat{\gamma}}} - 1 \right)^{-1}. \quad (7)$$

B. Algorithm 2: Estimation Based on Virtual Outage Probabilities

We now describe an estimation method based on outage probabilities that requires the knowledge of neither the network density nor the Nakagami fading parameter. We first review some theoretical results

and then present a practical scheme to estimate γ .

In [11], it is shown that when the signal power is exponentially distributed, the success probability p_s across any link in the Poisson network is equal to the product of the Laplace transforms of noise and interference. For the rest of this subsection, we assume that the system is interference-limited, i.e., $N_0 \ll I$. In particular, when the transceiver pair separation is unity⁵, we can express the probability of a successful transmission [17] as

$$p_s \approx \exp(-c_1 \Theta^{2/\gamma}), \quad (8)$$

where

$$c_1 = \lambda p \pi \mathbb{E}[G^{2/\gamma}] \Gamma(1 - 2/\gamma) = \frac{\lambda p \pi \Gamma(m + 2/\gamma) \Gamma(1 - 2/\gamma)}{\Gamma(m) m^{2/\gamma}}.$$

To estimate γ , the nodes are required to measure the SIR values during several time slots and use it to compute the empirical success probability, which matches the theoretical value (8). However, it is impractical to place transmitters for each receiver node where a SIR measurement is taken. Instead, nodes can simply measure the received powers, and compute the (virtual) SIRs taking the signal powers to be independent realizations of an exponential random variable. This algorithm is implemented at each node as follows.

- Record the values of the received powers R_1, \dots, R_N at the node during N time slots. Take the signal powers S_i , $1 \leq i \leq N$, to be N independent realizations of an exponential random variable with unit mean. Using the values S_i/R_i , $\forall i$, a histogram of the observed SIR values is obtained.
- Evaluate the empirical success probabilities at two different thresholds, i.e., compute $p_{s,j} = (1/N) \sum_{i=1}^N \mathbf{1}_{\{S_i/R_i > \theta_j\}}$, $j = 1, 2$.
- Match the empirically observed values with the theoretical values: From (8), we obtain $\ln(p_{s,1})/\ln(p_{s,2}) = (\Theta_1/\Theta_2)^{2/\gamma}$. Solving for γ yields the estimate

$$\hat{\gamma} = \frac{2 \ln(\Theta_1/\Theta_2)}{\ln(\ln(p_{s,1})/\ln(p_{s,2}))}, \quad (9)$$

which is independent of both λ and m .

Fig. 2 plots the relative MSE of $\hat{\gamma}$ for $\Theta_1 = 10$ dB and $\Theta_2 = 0$ dB for different γ and N values. We see that the error is small when the PLE is small, but increases at larger values of the PLE.

⁵When the transmitter node is unit distance away from the receiver node, the PLE will not affect the received power strength. This case is particularly helpful for the implementation of this PLE estimation algorithm.

Fig. 3 plots the relative MSE of $\hat{\gamma}$ versus m for various PLE values computed at the end of $N = 10000$ time slots. Note that the algorithm performs more accurately at lower values of m . We provide an intuitive explanation for this behavior in Section IV-D.

We have seen that γ can be estimated by measuring the outages at two specific values of the threshold. We now describe how to improve the estimation accuracy, at the cost of additional complexity. The idea is to find the best fit of the empirical distribution of the SIR to the theoretical value given by (8). Denote the observed values of the SIRs during N time slots by β_1, \dots, β_N . Let the empirical complementary cdf of the SIR at β_i be $\bar{F}(\beta_i)$. We use the Kolmogorov-Smirnov statistic [18] to define the goodness-of-fit. Accordingly, the estimate of the PLE is given by

$$\hat{\gamma} = \arg \min_{\gamma} \max_{1 \leq i \leq N} \left(\bar{F}(\beta_i) - \exp(-c_1 \beta_i^{2/\gamma}) \right). \quad (10)$$

The distribution (or curve) fitting method works effectively because of the difference in the behavior of the outage probability with respect to the parameters λ , m and γ . Specifically, consider the exponent in (8), and let $f(\Theta) = c_1 \Theta^{2/\gamma}$. In the plot of $f(\Theta)$ versus Θ , a change in the value of λ or m only scales the function f , while changing γ skews it. Thus, the value of γ that fits the distribution can be efficiently estimated by fitting the curve, even when λ and m are unknown.

Fig. 4 plots the relative MSE of $\hat{\gamma}$ versus the number of time slots, for $\gamma = 3, 4$. To find the solution to (10), we used the in-built function 'fminsearch' in MATLAB. The dashed lines indicate the MSE for the case where the estimation is performed using only two specific values of the threshold (as in (9)) and depict the improvement in performance obtained by employing distribution fitting. Evidently, when λ or m (or both) is known, the estimation accuracy can be improved further.

C. Algorithm 3: Estimation Based on the Cardinality of the Transmitting Set

Without the knowledge of the network density λ or the Nakagami parameter m , the PLE can be accurately estimated also based on the connectivity properties of the network. In this subsection, we derive the average number of nodes that are connected to an arbitrary node in the network and describe a PLE estimation algorithm based on our analysis.

For any node y , define its *transmitting set* $T(y)$ as the group of transmitting nodes whom it receives a (correctly decodable) packet from, in a given time slot. More formally, for receiver y , transmitter node x is in its transmitting set if they are connected, i.e., the SINR at y is greater than a certain threshold Θ .

Note that this set changes from time slot to time slot. Also note that for $\Theta = 0$ dB, the condition for a transceiver pair to be connected is that the received signal strength is greater than the interference power. Thus, for $\Theta \geq 1$, the cardinality of the transmitting set, $|T(y)|$, can at most be one, and that transmitter is the one with the best channel to the receiver. The estimation algorithm is based on matching the theoretical and empirical values of the mean number of elements in the transmitting set. The following proposition forms the basis of this estimation scheme.

Proposition 4.1: Under the conditions of $m \in \mathbb{N}$ and $N_0 \ll I$, the mean cardinality of the transmitting set of any arbitrary node in the network, \bar{N}_T , is proportional to $\Theta^{-2/\gamma}$.

Proof: For $N_0 \ll I$, the success probability for a transceiver pair at an arbitrary distance R units apart can be expressed as

$$\begin{aligned} p_s(R) &= \mathbb{E}_I [\Pr(GR^{-\gamma} > I\Theta \mid I)] \\ &\stackrel{(a)}{=} \mathbb{E}_I \left[\int_{I\Theta R^\gamma}^{\infty} \frac{m^m}{\Gamma(m)} x^{m-1} \exp(-mx) dx \right] \\ &\stackrel{(b)}{=} \frac{1}{\Gamma(m)} \int_0^{\infty} \Gamma(m, x\Theta R^\gamma m) p_I(x) dx, \end{aligned} \quad (11)$$

where (a) is obtained using (1) and (b) using the definition of the upper incomplete gamma function⁶, $\Gamma(\cdot, \cdot)$. Here, $p_I(x)$ denotes the pdf of the interference.

The expressions can be further simplified when m is an integer. For $m \in \mathbb{N}$, we have

$$\begin{aligned} p_s(R) &\stackrel{(a)}{=} \sum_{k=0}^{m-1} \frac{1}{k!} \int_0^{\infty} (x\Theta R^\gamma m)^k \exp(-x\Theta R^\gamma m) p_I(x) dx \\ &\stackrel{(b)}{=} \sum_{k=0}^{m-1} \frac{(-\Theta R^\gamma m)^k}{k!} \frac{d^k}{ds^k} M_I(s) \Big|_{s=\Theta R^\gamma m}, \end{aligned} \quad (12)$$

where $M_I(s)$ is the moment generating function (MGF) of I . Here, (a) is obtained from the series expansion of the upper incomplete gamma function and (b) using the definition of the MGF. When the node distribution is Poisson, we have the following closed-form expression for the MGF [17, Eqn. 20]: $M_I(s) = \exp(-\lambda p \pi \mathbb{E}_G[G^{2/\gamma}] \Gamma(1 - 2/\gamma) s^{2/\gamma})$, for $\gamma > 2$.

Using this, we get

$$p_s(R) = \exp(-c_2 R^2) \sum_{k=0}^{m-1} \frac{(c_2 R^2)^k}{k!} \left(\frac{2}{\gamma} \right)^k, \quad m \in \mathbb{N} \quad (13)$$

where $c_2 = \lambda p \pi \mathbb{E}_G[G^{2/\gamma}] \Gamma(1 - 2/\gamma) (\Theta m)^{2/\gamma} = c_1 (\Theta m)^{2/\gamma}$.

⁶Mathematica: Gamma[a,z]

Now, we consider an arbitrary receiver node y , shift it to the origin and analyze the transmitting set for this “typical” node. Consider a disk of radius a centered at the origin. Let E denote the event that an arbitrarily chosen transmitter inside this disk is in y ’s transmitting set. Since the nodes in the disk are uniformly randomly distributed, we have

$$\begin{aligned}
\Pr(E) &= \mathbb{E}_R[p_s(R) \mid R] \\
&= \frac{2\pi}{\pi a^2} \int_0^a \sum_{k=0}^{m-1} \frac{\exp(-c_2 r^2) r^{2k}}{k!} \left(\frac{2c_2}{\gamma}\right)^k r dr \\
&= \frac{1}{a^2} \sum_{k=0}^{m-1} \left(\frac{2c_2}{\gamma}\right)^k \int_0^a \frac{\exp(-c_2 r^2)}{k!} r^{2k} 2r dr \\
&\stackrel{(a)}{=} \frac{1}{a^2 c_2} \sum_{k=0}^{m-1} \left(\frac{2}{\gamma}\right)^k \frac{1}{k!} \int_0^{c_2 a^2} t^k \exp(-t) dt \\
&\stackrel{(b)}{=} \frac{1}{a^2 c_2} \sum_{k=0}^{m-1} \left(\frac{2}{\gamma}\right)^k \frac{1}{k!} (\Gamma(k+1) - \Gamma(k+1, c_2 a^2)),
\end{aligned}$$

where (a) is obtained by a simple change of variables ($c_2 r^2 = t$) and (b) using the definition of the upper incomplete gamma function.

Denote the mean number of transmitters in the disk of radius a by $N_a = \lambda p \pi a^2$. Then, we can write

$$\begin{aligned}
\bar{N}_T = \mathbb{E}|T(y)| &= \lim_{a \rightarrow \infty} N_a \Pr(E) \\
&\stackrel{(a)}{=} \frac{\lambda p \pi}{c_2} \sum_{k=0}^{m-1} \left(\frac{2}{\gamma}\right)^k = \frac{\lambda p \pi}{c_2} \frac{1 - \left(\frac{2}{\gamma}\right)^m}{1 - \frac{2}{\gamma}} \\
&\stackrel{(b)}{=} \frac{\Gamma(m) \left(1 - \left(\frac{2}{\gamma}\right)^m\right)}{\Gamma\left(m + \frac{2}{\gamma}\right) \Gamma\left(2 - \frac{2}{\gamma}\right) \Theta^{2/\gamma}}.
\end{aligned} \tag{14}$$

Here, (a) is obtained using the fact that $\lim_{z \rightarrow \infty} \Gamma(a, z) = 0$ and (b) using the definition of c_2 and (2). ■

The analytical value of the mean cardinality of the transmitting set when $m \in \mathbb{N}$ can be evaluated from (14). It is plotted in Fig. 5 for two different thresholds and $\gamma = \{2.5, 3, 3.5, 4, 4, 5\}$. Interestingly, since $\Gamma(2 + 2/\gamma) = (1 + 2/\gamma)\Gamma(1 + 2/\gamma)$, the values of \bar{N}_T at $m = 1$ and $m = 2$ are the same.

From (14), we see that \bar{N}_T is inversely proportional to $\Theta^{2/\gamma}$. Therefore, when m is a positive integer, the ratio of the mean cardinalities of the transmitting set at two different threshold values is independent of m . This forms the main idea behind the estimation algorithm, and we surmise that this behavior holds at arbitrary $m \in \mathbb{R}^+$.

The algorithm based on the cardinality of the transmitting works at each node in the network as follows.

- For a known threshold value $\Theta_1 \geq 1$, set $N_{T,1}(i) = 1$ in time slot i , $1 \leq i \leq N$, if the condition $\text{SIR} > \Theta_1$ holds, i.e., the node can correctly decode a packet and $N_{T,1}(i) = 0$ otherwise. Compute the empirical mean cardinality averaged over several time slots, $\bar{N}_{T,1} = (1/N) \sum_{i=1}^N N_{T,1}(i)$.
- Likewise, evaluate $\bar{N}_{T,2} = (1/N) \sum_{i=1}^N N_{T,2}(i)$ for another threshold value, $\Theta_2 \geq 1$.
- Equate the mean cardinalities of the transmitting set for the two different threshold values to obtain $\bar{N}_{T,1}/\bar{N}_{T,2} = (\Theta_2/\Theta_1)^{2/\gamma}$. Following this, γ is estimated as

$$\hat{\gamma} = \frac{2 \ln(\Theta_2/\Theta_1)}{\ln(\bar{N}_{T,1}/\bar{N}_{T,2})}. \quad (15)$$

Thus, this algorithm requires the knowledge of neither λ nor m .

Since the performance of this algorithm depends on knowing whether packets are correctly decoded or not, it is advisable to keep time slots across different nodes synchronized so that the SIR at any node remains the same throughout a slot of time (or the packet transmission time). However, this assumption may be relaxed by making the duration of a time slot large relative to the packet transmission time, and setting $N_T(i) = 1$ even if at least one of the received packets is decoded successfully.

Fig. 6 plots the empirical relative MSE of $\hat{\gamma}$ for algorithm 3 versus the number of time slots N for various PLE values, while Fig. 7 shows the relative MSE of $\hat{\gamma}$ versus m for $N = 10000$. Unlike Algo. 1 and 2, the relative MSE decreases with γ in this case. Also, we observe that the MSE is low at lower values of m and increases with m .

D. Discussion

The problem of PLE estimation is fundamental and non-trivial. Each of the three algorithms we have described is fully distributed and can be performed at each node in the network. There is no need for coordination among nodes, and they do not require any information on the locations of nodes in the network or the Nakagami parameter m . Simulation results validate that the estimates are quite accurate over a large range of the system parameters γ and m . Based on the relative MSE values, we conclude that at low values of γ , Algo. 1 performs the best (though it requires the network density to be known), while when γ is high, Algo. 3 is preferred. If time slots across nodes are not synchronized, Algo. 2 is useful. Also, the convergence of the MSE is seen to occur within about 2000 time slots for all of the algorithms. For time slots of the order of milliseconds, it takes only a few seconds for the PLE to be estimated in practice.

Each of the estimation algorithms works by equating empirically measured values of certain network characteristics with their corresponding theoretical values. There is a caveat though, that needs to be addressed. The theoretical results used are for the “average network” (they are obtained by averaging over all possible realizations of the transmitter locations and channel states). However, in practice we have only a single realization of the node distribution at hand. Thus, even though the set of transmitters and the fading component of the channel change independently in different slots, the node locations remain the same and the interference at the nodes are spatio-temporally correlated [19]. This means that in general, the empirically computed values only approximate the theoretic results and the estimates are biased. The bias (and the MSE) can be significantly lowered if the nodes that record measurements have access to several independent realizations of the PPP and we use this idea later to improve the estimation accuracy (see Section V). The ALOHA MAC scheme turns out to be particularly helpful in this regard.

The fact that nodes have access to just a single realization of the PPP also intuitively explains why the relative MSE decreases with m for Algorithms 2 and 3. Indeed, the variance of G is $1 + 1/m$ (obtained from (2)), which increases with decreasing m . Considering the fading and link distance ambiguities jointly, a lower m corresponds to having greater randomness in the location of the nodes (upon taking the fading component to be a constant). Thus, this condition is equivalent to nodes being able to see several diverse realizations of the process over different time slots, and leads to a lower MSE.

V. IMPROVING THE ESTIMATION ACCURACY

As mentioned, the values measured at the nodes match the theoretical values more closely if each node has access to a larger number of realizations of the process. Fortunately, in the scenario where nodes are distributed as a homogeneous PPP, we can employ two simple ways based on this principle to improve the estimation accuracy. We describe them in this section and also provide simulation results on the MSE that validate the significant improvement in the performance of the estimation algorithms.

A. Mobile nodes

Assume that nodes are mobile and that in each time slot, they each move with a constant velocity v in a randomly chosen direction ϕ , that is uniformly distributed in $[0, 2\pi)$ (random waypoint mobility model). Since the nodes move independently, each node observes a different PPP realization (with the same density) in each different time slot. By recording measurements over different time slots, γ can be estimated more accurately. Fig. 8 plots the relative MSE of $\hat{\gamma}$ for each of the three methods when $v = 0.1$

m/slot. The dashed lines represent the relative MSE when mobility is not considered, and are also plotted to depict the improvement in performance.

B. Coordinating Nodes

Alternatively, if nodes in a neighborhood can coordinate and exchange information, then their measurements can be combined to yield a more accurate estimate of the PLE. Since the homogeneous PPP is ergodic, for any measurable function defined on Φ , its statistical average (obtained over different PPP realizations), and its spatial average (obtained over different nodes in a single realization) are equal almost everywhere in the limit [16, pp. 172]. Based on this result, the estimation process can be performed more accurately on a single realization of the network by collecting the recorded measurements over several nodes.

Fig. 9 plots the relative MSE of the estimate versus the number of coordinating nodes K . From the figure, we see that the MSE sharply reduces with larger K . As mentioned earlier, with $K \rightarrow \infty$, the relative MSE $\rightarrow 0$. To obtain the simulation results, we used the collective measurements recorded at the K nodes located closest to the network center over $N = 2000$ slots. The estimates are based on the averaged value of these measurements.

VI. SENSITIVITY OF THE ALGORITHMS

We have formalized our algorithms based on the homogeneous PPP model and a spatially invariant path loss assumption. However, in reality, it is more likely that the nodal arrangement is not completely spatially random but takes on other forms such as being clustered or more regular. Also, the PLE value changes depending on the terrain type and the environmental conditions and hence cannot always be taken to be a constant over the entire network. In this section, we briefly comment on the sensitivity of our algorithms to these issues and illustrate how the PLE may be accurately estimated well even when some assumptions are relaxed.

A. Spatial Invariance of the PLE

In this subsection, we address the case where the PLE is not spatially invariant. For illustration purposes, we consider part of a network consisting of a square subregion A of side l centered at the origin with PLE γ_1 , and an outer region B of PLE γ_2 as shown in Fig. 10. To model the path loss, we use the multi-slope

piecewise linear model [1]. Accordingly, for transmitter node n_1 and receiver node n_2 , the path loss over a distance $r_1 + r_2$ (see Fig. 10) is $(r_1/r_0)^{-\gamma_1} \cdot (1 + r_2/r_1)^{-\gamma_2}$, for $r_1 > r_0$, where r_0 is the near-field radius.

Under this setting, we study the error performance of Algorithm 3 for different locations along the x -axis. Fig. 11 plots the relative MSE of $\hat{\gamma}$ when $l = 10$ with $\gamma_1 = 4$ and $\gamma_2 = 3$ and shows that Algorithm 3, by itself, works quite accurately even in a network with two different values of γ . The same qualitative behavior can be expected from Algorithms 1 and 2.

For cases where the PLE varies more rapidly or when the network is sparse, nodes can coordinate to obtain better estimates. To accurately do this, it is helpful if nodes have a general idea of the *PLE coherence length*, i.e., the distance over which the PLE can be assumed to be invariant. This may vary from about a mile if the network terrain changes rapidly to as much as hundreds of miles if the network extends from an urban to a suburban to a rural area. It can be assumed that the network operator has a general idea about variations in the PLE and based on this, the network is divided into sub-areas with constant PLEs, each of which is estimated separately. For instance, if the PLE coherence length is d , each node can estimate the PLE based on measurements recorded by other nodes that lie inside a disk of radius d around it.

B. Other Point Process Models

Even though all the algorithms are formulated for the case of the homogeneous PPP, they may also be used to estimate the PLE in different spatial point process patterns. The idea is to artificially make the arrangement of nodes appear more “spatially random”. This can be effectively performed upon simply employing randomized power control, wherein instead of having all the nodes transmit at unit power, we let nodes transmit at power levels drawn from a certain distribution. In fact, with independent thinning and appropriate rescaling, every point process is transformed into a stationary PPP (in the limit $p \rightarrow 0$) [21, Prop. 11.3.I]. A good choice for the distribution of transmit power levels is the exponential distribution since it is also the maximum non-negative entropy distribution [20], i.e., among all continuous pdfs supported on $[0, \infty)$ with a given mean, the exponential distribution has the maximum entropy.

Upon employing power control, the algorithms designed for the PPP case may also be used to estimate the PLE for other point processes. Fig. 12 plots the relative MSE values for Algorithm 3 for three non-PPP network models: the regular lattice grid, the Matern hard core process and the Thomas cluster process (for details on these point processes, please refer to [16]), and shows that the algorithm’s performance is quite accurate when power control (based on the exponential distribution) is employed. Similar qualitative

behavior may be expected of Algorithms 2 and 3. Using power control improves the estimation accuracy even in the case that the nodal arrangement is a PPP since it helps realize diverse realizations of the PPP.

VII. SUMMARY

In wireless systems, knowing the value of the PLE is critical, thus an accurate estimate is essential for their analysis and design. We offer a novel look at the issue of PLE estimation in a large wireless network, by taking into account Nakagami- m fading, the underlying node distribution and the network interference. We assume that nodes are arranged as a homogeneous PPP on the plane and the MAC scheme is slotted ALOHA (at least during the PLE estimation phase). For such a system, we present three distributed algorithms for PLE estimation, and provide simulation results to quantify the estimation errors. By incorporating mobility or coordination among nodes, the accuracy of the estimators can be greatly improved. The proposed algorithms perform well even when the PLE is spatially varying, and upon employing power control, in networks with more general node distributions. To the best of our knowledge, they are the first fully distributed PLE estimation algorithms in wireless networks with interference. This work is extensible to one- or three-dimensional networks in a straightforward manner.

REFERENCES

- [1] T. S. Rappaport, *Wireless Communications - Principles and Practice*, Prentice Hall, 1991.
- [2] W. C. Jakes, *Microwave Mobile Communications*, Wiley-IEEE Press, May 1994.
- [3] A. Savvides, W. L. Garber, R. L. Moses and M. B. Srivastava, "An Analysis of Error Inducing Parameters in Multihop Sensor Node Localization," *IEEE Transactions on Mobile Computing*, Vol. 4, No. 6, pp. 567-577, Nov./Dec. 2005.
- [4] N. Patwari, I. Hero, A. O. M. Perkins, N. Correal, and R. O'Dea, "Relative Location Estimation in Wireless Sensor Networks," *IEEE Transactions on Signal Processing*, Vol. 51, No. 8, pp. 2137-2148, Aug. 2003.
- [5] C. Bettstetter, "On the Connectivity of Ad Hoc Networks," *The Computer Journal, Special Issue on Mobile and Pervasive Computing*, Vol. 47, No. 4, pp. 432-447, July 2004.
- [6] D. Miorandi and E. Altman, "Coverage and Connectivity of Ad Hoc Networks in the Presence of Channel Randomness," *IEEE INFOCOM*, Mar. 2005.
- [7] M. Sikora, J. N. Laneman, M. Haenggi, D. J. Costello, Jr., and T. E. Fuja, "Bandwidth- and Power-Efficient Routing in Linear Wireless Networks," *IEEE Transactions on Information Theory*, Vol. 52, No. 6, pp. 2624-2633, June 2006.
- [8] A. Özgür, O. Lévêque and D. Tse, "Hierarchical Cooperation Achieves Optimal Capacity Scaling in Ad Hoc Networks," *IEEE Transactions on Information Theory*, Vol. 53, No. 10, pp. 3549-3572, Oct. 2007.
- [9] M. Haenggi, "On Routing in Random Rayleigh Fading Networks," *IEEE Transactions on Wireless Communications*, Vol. 4, No. 4, pp. 1553-1562, July 2005.
- [10] M. Haenggi, "Outage, Local Throughput, and Capacity of Random Wireless Networks," accepted to *IEEE Transactions on Wireless Communications*, 2009. Available at <http://www.nd.edu/~mhaenggi/pubs/twc09.pdf>.

- [11] 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.
- [12] G. Mao, B. D. O. Anderson and B. Fidan, "Path Loss Exponent Estimation for Wireless Sensor Network Localization," *Computer Networks*, Vol. 51, Iss. 10, pp. 2467-2483, July 2007.
- [13] X. Zhao, L. Razoumov and L. J Greenstein, "Path Loss Estimation Algorithms and Results for RF Sensor Networks," *Vehicular Technology Conference*, Vol. 7, pp. 4593- 4596, Sep. 2004.
- [14] N. Benvenuto and F. Santucci, "A Least Squares Path-Loss Estimation Approach to Handover Algorithms," *IEEE Transactions on Vehicular Technology*, Vol. 48, No. 2, pp. 437-447, Mar. 1999.
- [15] M. Nakagami, "The m-distribution : A General Formula for Intensity Distribution of Rapid Fading," in W. G. Hoffman, *Statistical Methods in Radiowave Propagation*, Pergamon Press, Oxford, U. K., 1960.
- [16] D. Stoyan, W. S. Kendall and J. Mecke, *Stochastic Geometry and its Applications*, 2nd Edition, Wiley & Sons, 1996.
- [17] S. Srinivasa and M. Haenggi, "Modeling Interference in Finite Uniformly Random Networks," *International Workshop on Information Theory for Sensor Networks (WITS '07)*, Santa Fe, June 2007.
- [18] M. A. Stephens, "EDF Statistics for Goodness of Fit and Some Comparisons," *Journal of the American Statistical Association*, Vol. 69, No. 347, pp. 730-737, Sep. 1974.
- [19] R. K. Ganti and M. Haenggi, "Spatial and Temporal Correlation of the Interference in ALOHA Ad Hoc Networks," submitted to *IEEE Communication Letters*, 2009. Available at <http://www.nd.edu/~mhaenggi/pubs/commletter09.pdf>.
- [20] A. M. Kagan, Y. V. Linnik and C. R. Rao, *Characterization Problems in Mathematical Statistics*, Wiley, 1973.
- [21] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*, 2nd Edition, Springer, 2007.

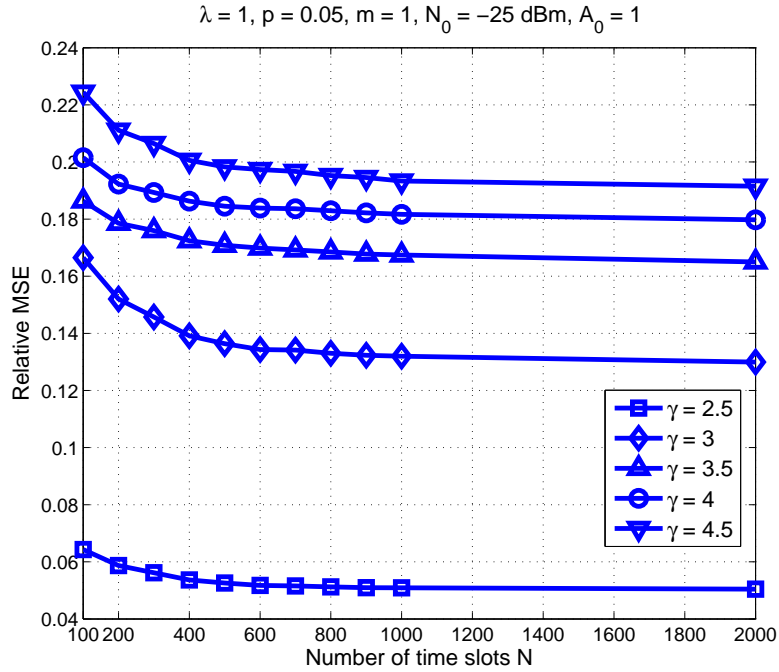


Fig. 1. Relative MSE of $\hat{\gamma}$ versus the number of time slots for different PLE values, for the estimation method based on the mean interference. The error is small, in particular when the PLE is small.

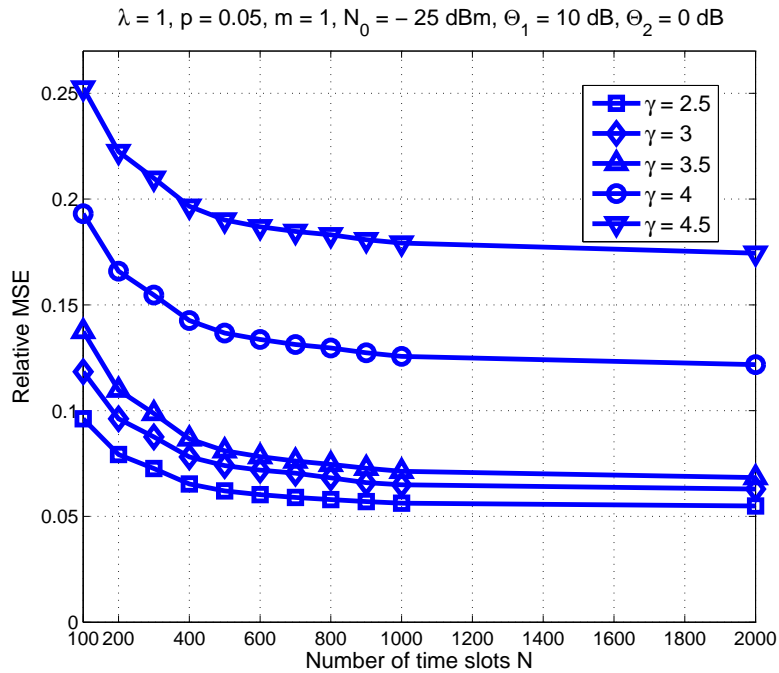


Fig. 2. Relative MSE of $\hat{\gamma}$ versus the number of time slots for the estimation method based on virtual outage probabilities. Alike Algo. 1, the estimation error increases with larger $\hat{\gamma}$.

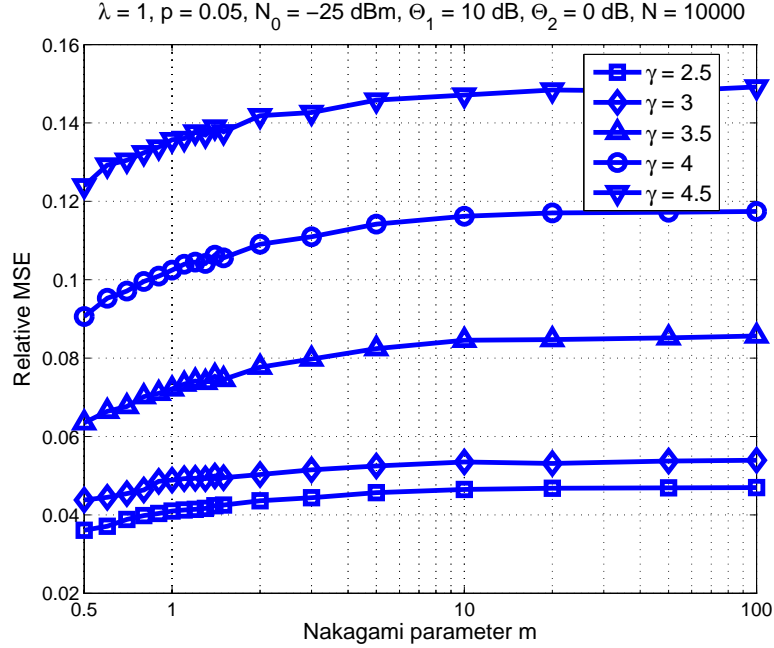


Fig. 3. Relative MSE of $\hat{\gamma}$ versus the Nakagami parameter m for different PLE values, for the estimation method based on virtual outage probabilities. Note that this algorithm performs more accurately at lower values of m .

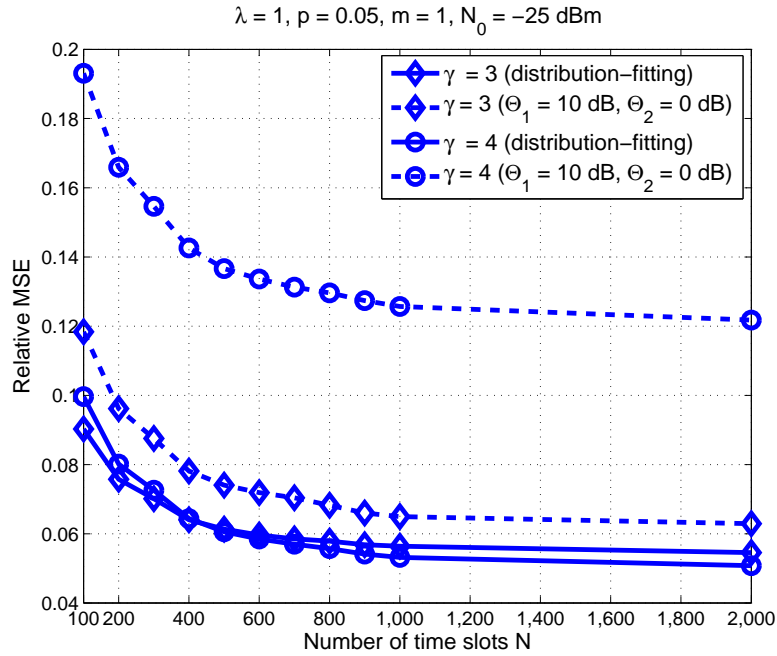


Fig. 4. Relative MSE of $\hat{\gamma}$ versus N for Algorithm 2 for $\gamma = 3, 4$ (dashed lines) and the same upon employing the distribution-fitting based technique (solid lines). The latter method is seen to drastically improve the estimation accuracy, in particular when γ is large.

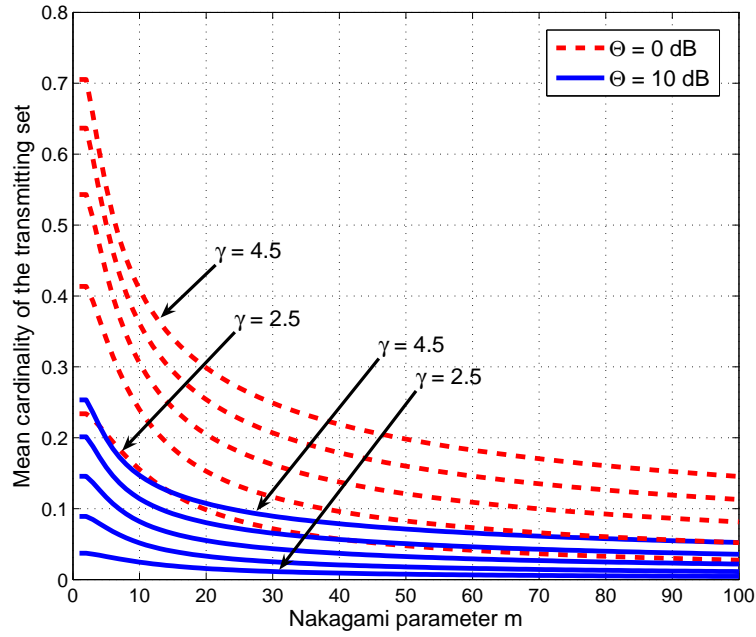


Fig. 5. The theoretical expected cardinality of the transmitting set for various values of Θ , m and γ .

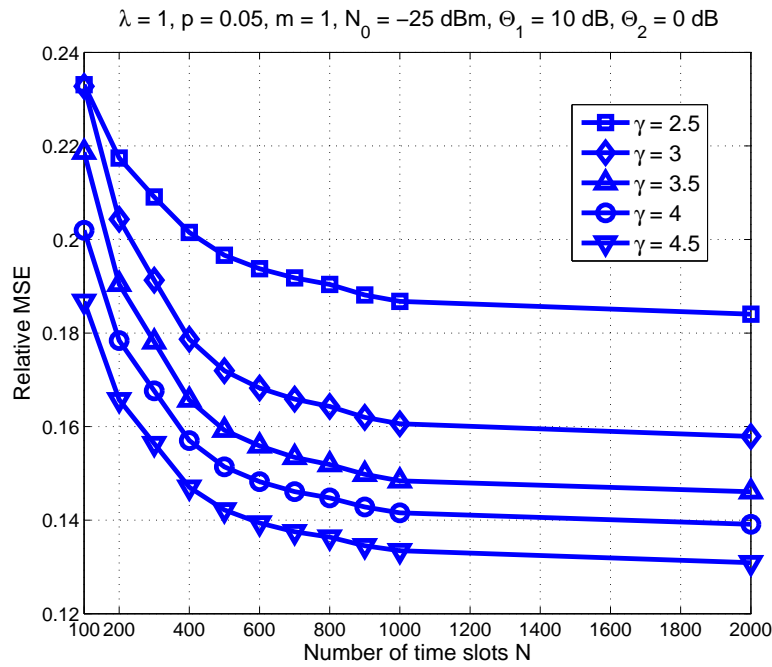


Fig. 6. Relative MSE of $\hat{\gamma}$ versus the number of time slots for different PLE values, for the estimation method based on the mean cardinality of the transmitting set. In contrast to Algo. 1 and 2, the relative MSE decreases with increasing γ .

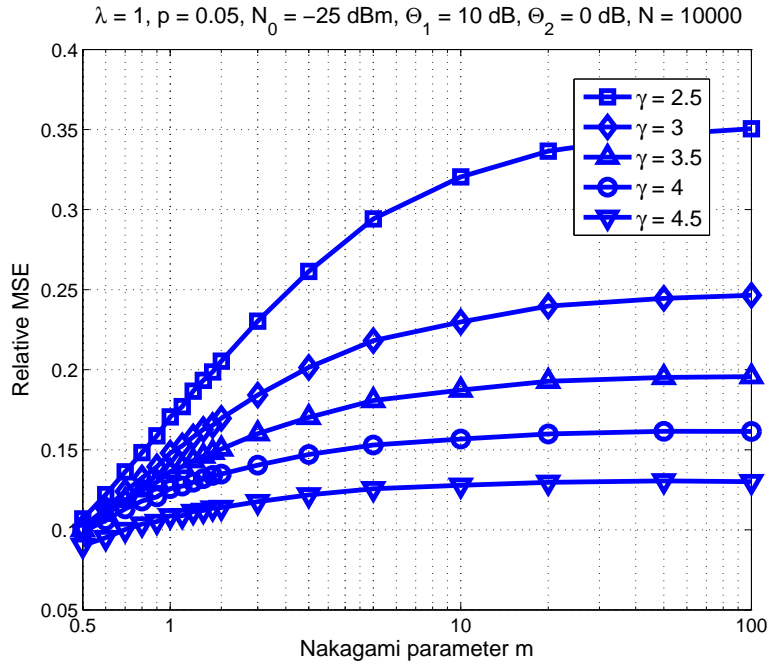


Fig. 7. Relative MSE of $\hat{\gamma}$ versus the Nakagami parameter m for the estimation algorithm based on the mean cardinality of the transmitting set. The estimates are more accurate at lower m .

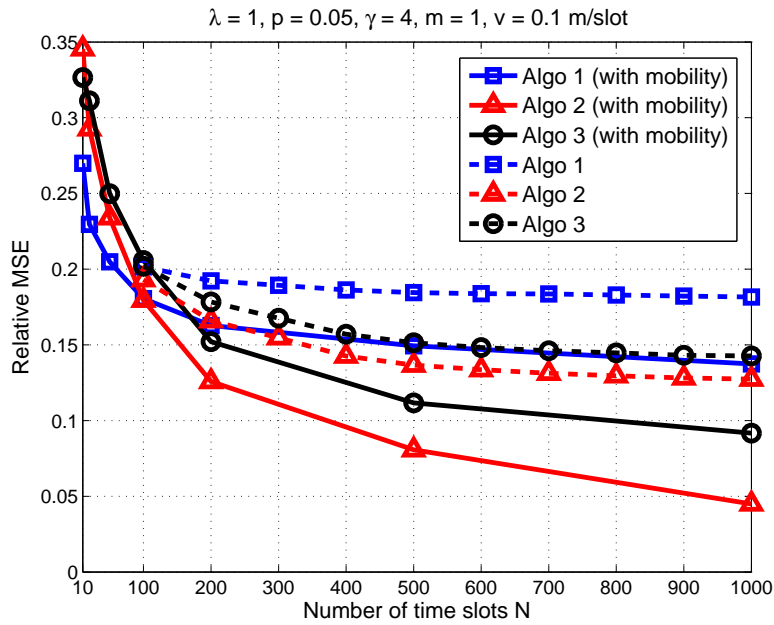


Fig. 8. Relative MSE of $\hat{\gamma}$ versus the number of time slots, for the three estimation algorithms, with and without consideration of mobility.

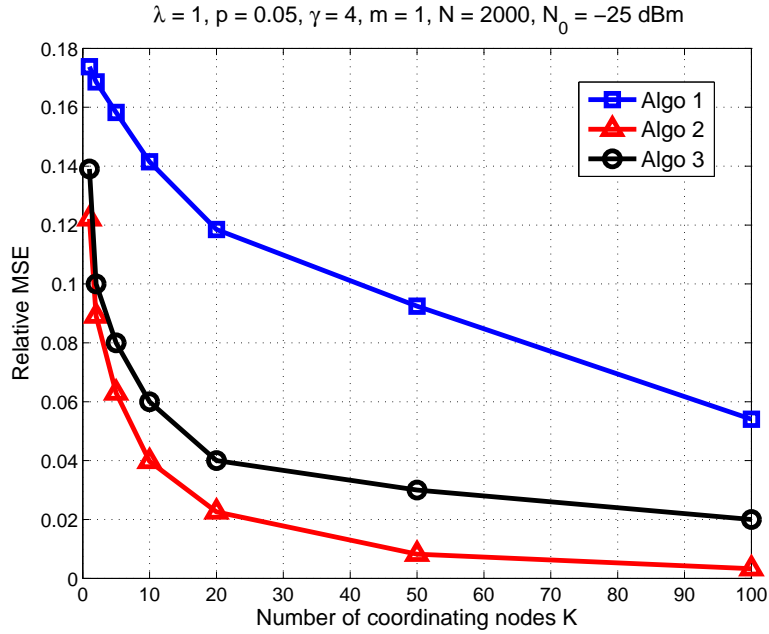


Fig. 9. Relative MSE of $\hat{\gamma}$ versus the number of coordinating nodes, for each of the estimation algorithms.

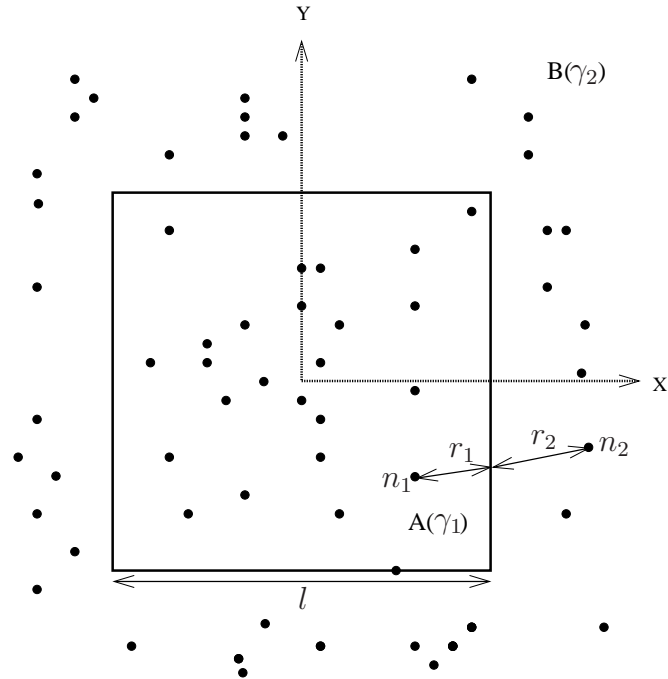


Fig. 10. The considered Poisson network model with the square subregion A having a different value of PLE, γ_1 compared to the rest of the network B. The attenuation between nodes n_1 and n_2 is modeled by a piecewise linear path loss model.

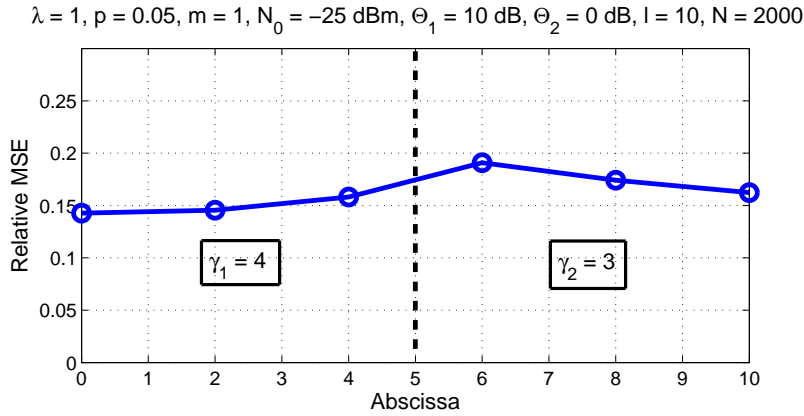


Fig. 11. Relative MSE $\hat{\gamma}$ for algorithm 3 at different locations $(x, 0)$. The true values of the PLE are $\gamma_1 = 4$ for $x \leq 5$ and $\gamma_2 = 3$ for $x > 5$.

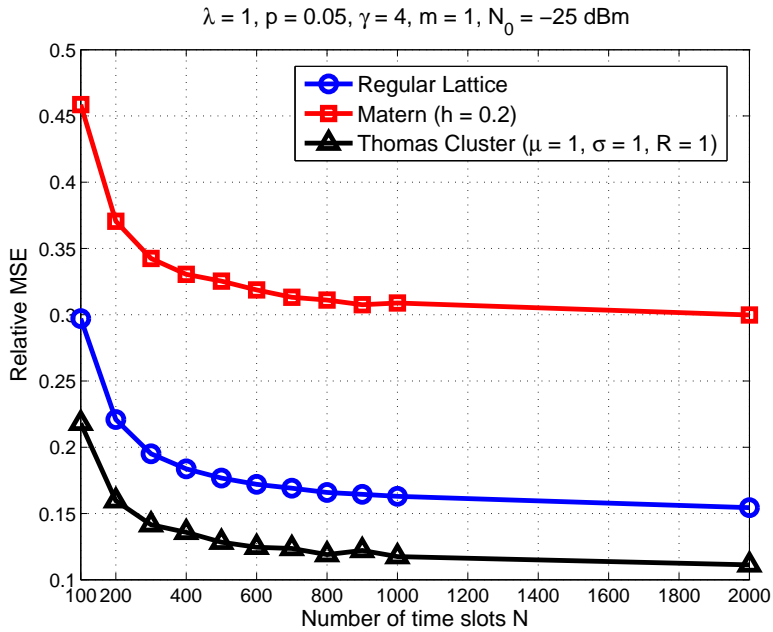


Fig. 12. Relative MSE of $\hat{\gamma}$ versus the number of time slots, for three non-PPP models. In each case, Algorithm 3 is found to estimate accurately.