COVERAGE PERFORMANCE IN CELLULAR NETWORKS: SPATIAL STOCHASTIC MODEL FITTING AND ASYMPTOTIC DEPLOYMENT GAIN

A Thesis

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Master of Science

in

Electrical Engineering

by

Anjin Guo

Martin Haenggi, Director

Graduate Program in Electrical Engineering
Notre Dame, Indiana

April 2014

COVERAGE PERFORMANCE IN CELLULAR NETWORKS: SPATIAL STOCHASTIC MODEL FITTING AND ASYMPTOTIC DEPLOYMENT GAIN

Abstract

by

Anjin Guo

The spatial structure of base stations (BSs) in cellular networks plays a key role in evaluating the downlink performance. The BSs are usually assumed to form a lattice or a Poisson point process (PPP). In reality, however, they are deployed neither fully regularly nor completely randomly.

Accordingly, in this thesis, we first use different spatial stochastic models, including the PPP, the Poisson hard-core process (PHCP), the Strauss process (SP), and the perturbed triangular lattice, to model the spatial structure by fitting them to the locations of BSs in real cellular networks obtained from a public database. We propose the coverage probability (the probability that the signal-to-interference-plus-noise-ratio (SINR) exceeds a threshold) as the criterion for the goodness-of-fit, and provide two general approaches for fitting. One approach is fitting by the method of maximum pseudolikelihood. As for the fitted models, the SP provides a better fit than the PPP and the PHCP. The other approach is fitting by the method of minimum contrast that minimizes the average squared error of the coverage probability. This way, fitted models are obtained whose coverage performance matches that of the given data set very accurately.

Second, we consider the very general class of motion-invariant point processes as

models for the BSs and theoretically analyze the behavior of the outage probability. (complement of the coverage probability). We show that, remarkably, the slope of the outage probability (in dB) as a function of the threshold (also in dB) is the same for essentially all motion-invariant point processes, as the threshold goes to zero. The slope merely depends on the fading statistics. Using this result, we introduce the notion of the asymptotic deployment gain (ADG), which characterizes the horizontal gap between the coverage probability of the PPP and another point process in the high-reliability regime (where the coverage probability is near 1). To demonstrate the usefulness of the ADG for the characterization of the coverage, we investigate the coverage probabilities and the ADGs for different point processes and fading statistics by simulations.

DEDICATION

This thesis is gratefully dedicated to my father and mother. Their encouragement and love are always impetuses to me for exploring new things.

CONTENTS

FIGURES

ACKNOWLEDGMENTS

I would like to acknowledge my advisor Prof. Martin Haenggi. He has given me so much inspiration and help. Without him, this work would not have been possible.

Finally, I would like to thank the NSF, for their generous grants, CNS 1016742 and CCF 1216407, which allowed me to pursue my work.

SYMBOLS

Φ	a point process
λ	intensity of a point process
λ_{p}	intensity of the parent point process
λ_{b}	intensity of the basic point process
P_{o}	outage probability
P_{c}	coverage probability
$P_{\mathrm{c}}^{\mathrm{PPP}}$	coverage probability of the PPP
$P_{\mathrm{o}}^{\mathrm{PPP}}$	outage probability of the PPP
$P_{\mathrm{c}}^{\mathrm{TL}}$	coverage probability of the triangular lattice
$P_{\mathrm{o}}^{\mathrm{TL}}$	outage probability of the triangular lattice
$\mathcal{N}$	the set of counting measures on $\mathbb{R}^2$
$\mathfrak{N}$	the σ-algebra of counting measures
$\Phi(B)$	the number of points in set $B \subset \mathbb{R}^2$ for a point process Φ
φ	a concrete realization of Φ
$\varphi(B)$	deterministic counting measure that denotes the number of points in B
$\tilde{R}$	the interaction radius of the Strauss process
$G(\cdot)$	the nearest-neighbor distance distribution function
$F(\cdot)$	the empty space function
$K(\cdot)$	Ripley's K function
α	the path loss exponent
$\ell(\cdot)$	the path loss function
$\bar{\gamma}$	the average ergodic rate

$G(p_t)$	the deployment gain at p_t
$\hat{S}_g$	the average deployment gain
$\alpha^{(n)}$	the nth-order factorial moment measure
$\rho^{(n)}$	the nth moment density
NP_φ	the nearest-point operator for a point pattern $\varphi \subset \mathbb{R}^2$
$I(\Phi)$	the total interference in Φ
$b(x, r)$	the open disk centered at x with radius r
$\hat{G}$	the asymptotic coding gain
ξ	the contact distance
ξ_max	the supremum of the contact distance
o	the origin
Φ_o^ζ	$(\Phi \mid \mathrm{NP}_\Phi(o) = \zeta)$
Φ^ζ	$(\Phi \mid \zeta \in \Phi)$
$\hat{I}(\Phi^\zeta)$	the interference of Φ^ζ outside a disk of radius $\|\zeta\|/2$ around the origin
$B_{\zeta/2}$	$\mathbb{R}^2 \setminus b(o, \|\zeta\|/2)$
$\nu(\cdot)$	the Lebesgue measure
W	the thermal noise power
δ	$2/\alpha$
M_Φ	the mean SINR for Φ
M_PPP	the mean SINR for the PPP

CHAPTER 1

INTRODUCTION

1.1 Motivation

In cellular networks, as the power of received signals and interferences depends on the distances between the receiver and base stations (BSs), the downlink performance is affected by the spatial structure. System engineers and researchers often use a regular triangular lattice or a square lattice [1–3] to model the structure deterministically. But in reality, the topology of the BSs is not so ideal but depends on many natural or man-made factors, such as the landscape, topography, bodies of water, population densities, and traffic demands. As a consequence, the BSs are more suitably modeled as deployed randomly instead of deterministically, and stochastic geometry is an efficient tool to analyze this kind of geometrical configurations and provide theoretical insights [4, 5]. Recently, it was shown in [8] that a completely irregular point process, the Poisson point process (PPP) [4, 5], may be used without loss in accuracy (compared to the lattice) but significant gain in analytical tractability. Observations of real BS locations in UK, however, show that real deployments fall somewhere in between the two extremes of full regularity (the triangular lattice) and complete randomness (the PPP). They exhibit some degree of repulsion between the BSs, as expected, since the operators do not place the BSs closely together.

A natural question is whether there is a point process model that is better in modeling the BS topology than the PPP. The critical first step to answer the question is to collect data of real BS locations, investigate some regular (or repulsive) point

processes and the PPP by fitting them to the data, and then evaluate the goodness-of-fit. Those works will be discussed in Chapter 2. The answer to the question is positive. The Poisson hard-core process (PHCP), the Strauss process (SP), and the perturbed triangular lattice can provide better fitting results than the PPP, in term of the *coverage probability*, which is defined as the complementary cumulative distribution function of the signal-to-interference-plus-noise ratio (SINR), i.e., $P_{\mathrm{c}}(\theta) \triangleq \mathbb{P}(\mathrm{SINR} > \theta)$.

The next question is how we can deal with those non-Poisson point processes, i.e., whether we can obtain good analytical results from them by applying them to wireless networks. Generally speaking, the analysis of non-Poisson point processes is significantly more difficult than the analysis of the PPP, since dependencies exist between the locations of the BSs. We are not aware of any tractable analytical methods that are applicable in general. In Chapter 3, we will provide an indirect approach to evaluate the coverage probability of cellular networks, where BSs follow a general class of point process models, using the *asymptotic deployment gain* (ADG). The ADG characterizes the horizontal gap between the coverage probability of the PPP and another point process in the high-reliability regime.

Therefore, our work in this thesis can be divided into two separate parts, which are, briefly speaking, the spatial stochastic model fitting and the coverage analysis using the asymptotic deployment gain.

1.2 Spatial Stochastic Model Fitting

Our work on spatial stochastic model fitting is based on the real deployment of BSs; we have several point sets that denote the actual locations of BSs collected from the Ofcom[1] - the independent regulator and competition authority for the UK

[1]Ofcom website: `http://sitefinder.ofcom.org.uk/search`

communications industries, where the data are open to the public. Table 1.1 gives the details of the three point sets used in this thesis, and Figs. 1.1-1.3 visualize these point sets. Note that these point sets all represent the BSs of the operator Vodafone with frequency band 900 MHz (GSM). Although the data sets of certain operators in the Ofcom are almost 10 years old, the data set of the operator Vodafone, is quite up-to-date, since its last update in the Ofcom database occurred in October 2011.

TABLE 1.1

DETAILS OF THE THREE POINT SETS

	Operator	Area ($m \times m$)	Center Location (latitude, longitude)	Number of BSs
Urban region	Vodafone	1500×1050	(51.515° N, -0.132° W)	64
Rural region 1	Vodafone	78200×48200	(52.064° N, -1.381° W)	62
Rural region 2	Vodafone	66700×50000	(52.489° N, 0.704° W)	69

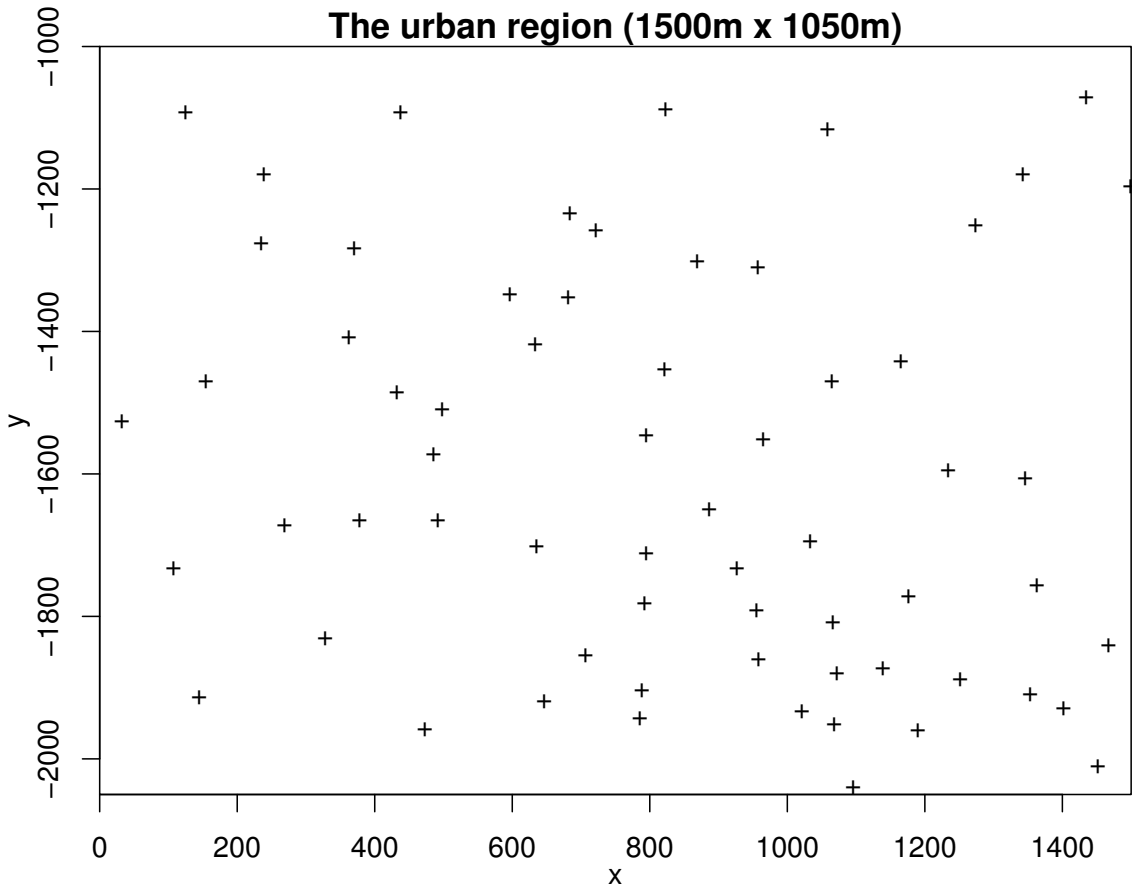

Figure 1.1. The locations of the BSs (the urban region).

The main objective of the model fitting is to find an accurate point process to model the real deployment of BSs. To accomplish it, we have to first define the metrics to evaluate the goodness of different models. Some classical statistics in stochastic geometry, such as the J function and the L function [5], can be used as the metrics. Nevertheless, simulations show they are not sufficient to discriminate between different models. Since we study the point processes in the context of wireless networks, it is natural to instead use a key performance metric of cellular systems, namely the coverage probability [5, Ch. 13], [8, 9].

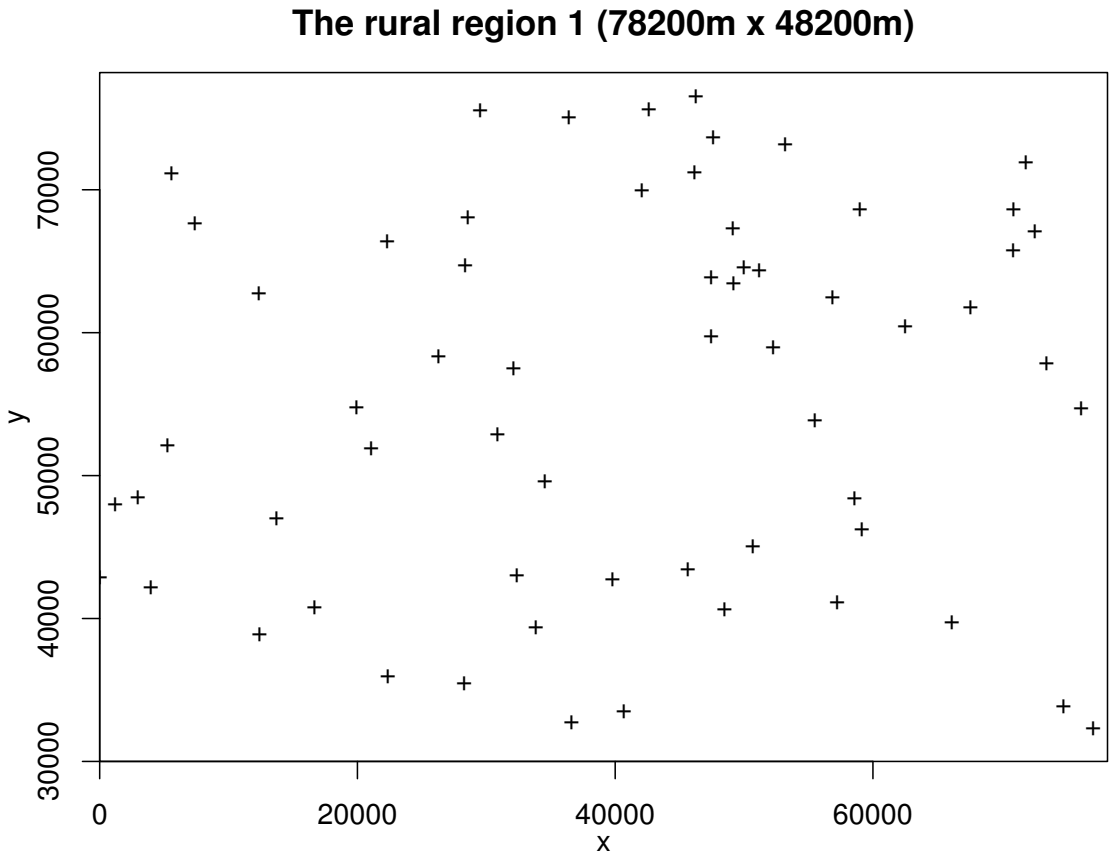

Figure 1.2. The locations of the BSs (the rural region 1).

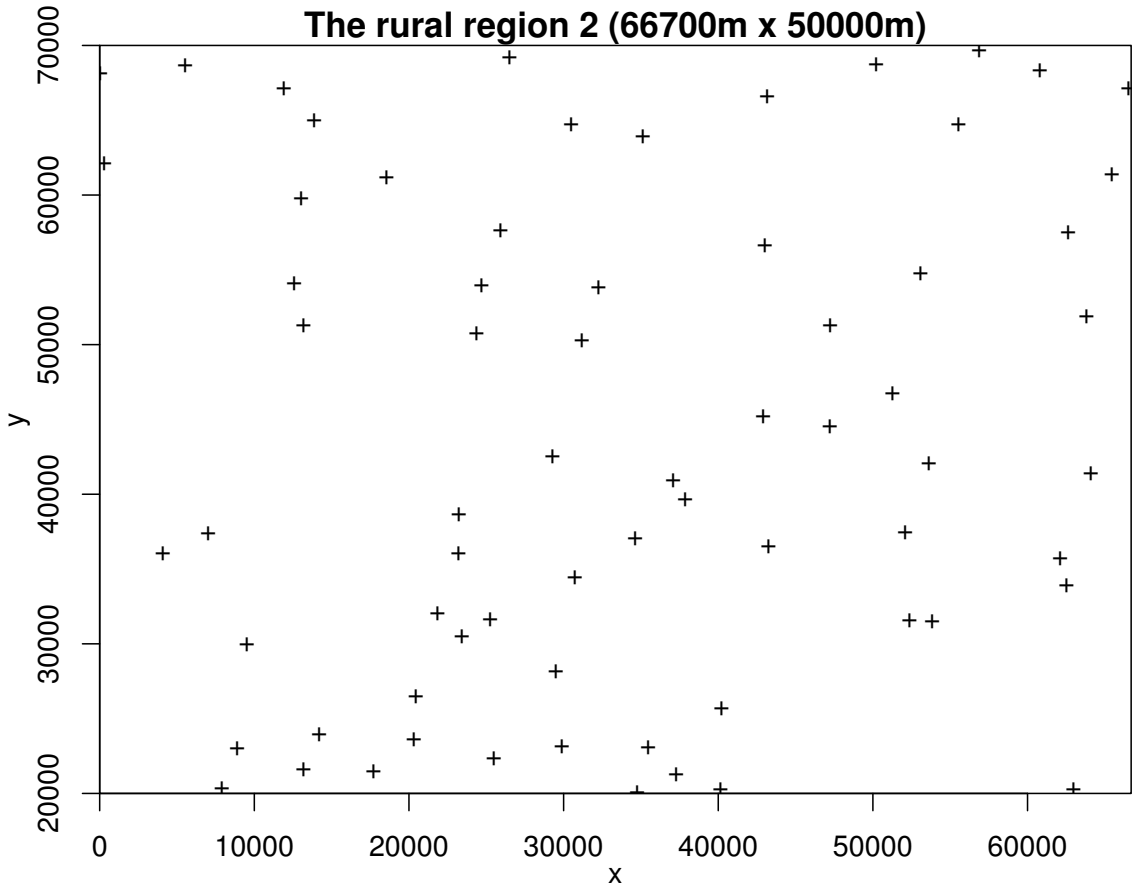

Figure 1.3. The locations of the BSs (the rural region 2).

As [8] indicates, the PPP model and the lattice provide a lower bound and an upper bound on the coverage probability, respectively. Since the point sets appear to be regular and their coverage probabilities lie between the PPP's and the lattice's, we are interested in point process models that lie in between the two in terms of regularity, such as the Poisson hard-core process (PHCP), the Strauss process (SP), and the perturbed triangular lattice. In order to find the desired point process, we use two different fitting methods. The first one is the method of maximum pseudo-likelihood [22], which is the usual method for model fitting in stochastic geometry. The second one is the method of minimum contrast [23], which is used to find the fitted model that minimizes the average squared error of the coverage probability.

Using the first method, we fit the PPP, the PHCP, and the SP to the point sets and determine the best fitted model. Simulations indicate that the SP is the best, followed by the PHCP and then the PPP. But there is still a gap between the coverage probabilities of the SP and the corresponding point set. The perturbed triangular lattice is not considered, since its likelihood and pseudolikelihood are generally unknown.

In the second method, the intensity is assumed to be fixed to the density of the given point sets. The PPP is not considered, since it would result in the same model as with the first method. The fitted models of the SP, the PHCP, and the perturbed triangular lattice for the point sets are obtained. They exhibit quite exactly the same coverage performance as the given point sets. Note that this method is not limited to the average squared error minimization of the coverage probability; it can be applied to many other performance metrics in wireless networks and second-order statistics in stochastic geometry.

Using the two fitting methods, we can find a fitted model that describes the given point set accurately. Although the SP, the PHCP and the perturbed triangular lattice are not as tractable as the PPP, they still have many useful properties. By studying

the fitted model, we can obtain properties for a class of point sets.

For some applications where the chief concern is the coverage evaluation of the point sets rather than their spatial structure, there is a simple way of the evaluation using a novel metric we propose, which is called the *deployment gain*; it measures how close the coverage curve of a point set or a point process model is to that of the PPP. A larger deployment gain means the point set or the model provides better coverage. For example, the deployment gains of the three point sets are: urban region $>$ rural region 1 $>$ rural region 2, which is also the rank of their coverage curves from top to bottom. The deployment gain provides a simple yet highly accurate way of using the analytical results available for the PPP for the analysis of more realistic point process models.

The main contributions on the spatial stochastic model fitting are summarized as follows:

1. We use the coverage probability as a metric to compare different point processes and publicly available point sets, which is shown to be more effective than the classical statistics in stochastic geometry for the data sets we used;
2. Through fitting the PPP, the PHCP, and the SP to the given point set using the method of maximum pseudolikelihood, we discover that the SP has the best coverage performance, while the PPP has the worst;
3. Through fitting the SP, the PHCP, and the perturbed triangular lattice by minimizing the average squared error of the coverage probability, we find that the fitted models have nearly the same coverage probability as the given point set, and thus, in terms of the coverage probability, they are accurate models of the real deployments of the BSs;
4. We propose the deployment gain to analytically compare the coverage probability performances of different point sets or different models and to show how results for the PPP can be applied to more accurate point process models.

1.3 Asymptotic Deployment Gain

In our work on the ADG, we provide an indirect approach to the coverage probability analysis of an arbitrary motion-invariant (isotropic and stationary[2]) point process [5, Ch. 2] by comparing its coverage probability to the coverage probability of the PPP. To validate this approach, we establish that *the outage probability* $1 - P_{\mathrm{c}}(\theta)$ *of essentially all motion-invariant (m.i.) point processes, expressed in dB, as a function of the SINR threshold* θ*, also in dB, has the same slope as* $\theta \to 0$. The slope depends on the fading statistics. This result shows that asymptotically the coverage curves $P_{\mathrm{c}}(\theta)$ of all m.i. models are just (horizontally) shifted versions of each other in a log-log plot, and the shift can be quantified in terms of the horizontal difference $\hat{G}$ along the θ (in dB) axis. Since the coverage probability of the PPP is known analytically, the PPP is a sensible choice as a reference model, which then allows to express the coverage probability of an arbitrary m.i. model as a gain relative to the PPP. This gain is called the *asymptotic deployment gain* (ADG).

We introduced the concept of the *deployment gain* (DG) in our work on the spatial stochastic model fitting. It measures how close a point process or a point set is to the PPP at a given target coverage probability. Here we extend the DG to obtain a quantity that does not depend on a target coverage probability, formally define its asymptotic counterpart—the ADG.

The work on the ADG makes the following contributions:

1. We introduce the asymptotic deployment gain.
2. We formally prove its existence for a large class of m.i. point processes.
3. We show how the asymptotic slope of the outage probability depends on the fading statistics.
4. We demonstrate through simulations how the ADG can be used to quantify the

[2]Stationarity implies that the coverage probability does not depend on the location of the typical user.

coverage probability of several non-Poisson models, even if the SINR threshold θ is not small.

CHAPTER 2

SPATIAL STOCHASTIC MODEL FITTING

In this chapter, we discuss in detail how to find an accurate point process to model the real deployment of BSs, using the method of maximum pseudolikelihood and the method of minimum contrast. Moreover, we introduce the deployment gain.

This chapter is organized as follows. In Section 2.1, we discuss related works. In Section 2.2, basic concepts of point processes are introduced. In Section 2.3, the PPP, the PHCP, and the SP are fitted to the point sets using the method of maximum pseudolikelihood, and some classical statistics, the coverage probability and the average ergodic rate are used to test the goodness of fitted models. In Section 2.4, the SP, the PHCP, and the perturbed triangular lattice are used to model the given point set by the method of minimum contrast. The deployment gain is introduced in Section 2.5. Conclusions are drawn in Section 2.6.

2.1 Related work

Since the Poisson point process (PPP) [4–7] is highly tractable, it is frequently used to model a variety of networks, such as cellular networks [8–12], mobile ad hoc networks [4–6], cognitive radio networks [13] and wireless sensor networks [14]. For cellular networks, in [8], the authors assume the distribution of BSs follows a homogeneous PPP and derive theoretical expressions for the downlink signal-to-interference-plus-noise-ratio (SINR) complementary cumulative distribution function (CCDF) and the average ergodic rate under some assumptions. [9] is an extension of [8], in which the authors model the infrastructure elements in heterogeneous cellular

networks as multi-tier independent PPPs. In [10], the BSs locations are also modeled as a homogenous PPP, and the outage probability and the handover probability are evaluated. Although many useful theoretical results can be derived in closed form for the PPP, the PPP may not be a good model for real BSs' deployments in homogenous networks, as will be shown in this chapter.

Indeed, the BS locations appear to form a more regular point pattern than the PPP, which means there exists repulsion between points, hence the hard-core processes and the Strauss process might be better to describe them. The Matérn hard-core processes [5–7] are often used to model concurrent transmitters in CSMA networks [15–17]. In [17], the author uses them to determine the mean interference in CSMA networks, observed at a node of the process. In [18], a modified Matérn hard-core process is proposed to model the access points in dense IEEE 802.11 networks. But to the best of our knowledge, no prior work has modeled the BSs in cellular networks using hard-core processes.

The Strauss process has not been used in wireless networks, but its generalization, the Geyer saturation process [19], is fitted to the spatial structures of a variety of wireless network types using the method of maximum pseudolikelihood in [20]. The difference between the two processes is that the Strauss process is a regular (or soft-core) process, while the Geyer saturation process can be both clustered and regular depending on its parameters. To evaluate the goodness-of-fit in [20], the authors compare the statistics of the original data and the fitted model, such as the nearest-neighbor distance distribution function, the empty space function, the J function, the L function, and the residuals of the model. Though these statistics verify that the Geyer saturation process is suitable to model the data set, they may not be sufficient to discriminate between different point processes in terms of a metric specific to wireless networks. In this chapter, all the processes mentioned above are studied comprehensively, and we use different statistics to compare their suitability

as models for cellular networks.

The perturbed lattice, which is another soft-core model and thus less regular than the lattice, can also be used to model the BS locations. In [21], the authors consider the BSs as a perturbed lattice network and analyze the fractional frequency reuse technique. The degree of the perturbation is assumed to be a constant. But this constant may not be consistent with real configurations of the BSs. In our work, perturbed lattice networks with different levels of the perturbation are investigated.

2.2 Spatial Point Process Models

2.2.1 Overview

The spatial point processes we considered lie in the Euclidean plane $\mathbb{R}^2$. Informally, a point process is a countable random collection of points in $\mathbb{R}^2$. If it is simple (there is only one point at each location a.s.), it can be represented as a countable random set $\Phi = \{x_1, x_2, \ldots\}$, where $x_i \in \mathbb{R}^2$ are the points. Usually, it is characterized by a random counting measure $N \in \mathcal{N}$, where $\mathcal{N}$ is the set of counting measures on $\mathbb{R}^2$. $(\mathcal{N}, \mathfrak{N})$ is the measurable space, where $\mathfrak{N}$ is the σ-algebra of counting measures. $N(B)$ is a random variable that denotes the number of points in set $B \subset \mathbb{R}^2$ for a point process Φ. Instead of $N(B)$, the notation $\Phi(B)$ is frequently used, since it makes the connection of the point process to the counting measure explicit. A concrete realization of Φ is denoted as φ. Hence $\varphi(B)$ is a deterministic counting measure that denotes the number of points in B. See [5, Ch. 2] for details.

There are many kinds of point processes, such as the PPP, cluster processes, hard-core processes and Gibbs processes [5, Ch. 3]. They can be placed into three categories, the complete spatial randomness (i.e., the PPP), clustered processes, and regular processes. Clustering means there is attraction between points, while regularity means there is repulsion. So the probability of having a nearby neighbor in

regular processes is smaller than in the PPP and clustered processes. Since regularity is good for interference minimization and coverage optimization in wireless networks and the deployment of BSs appears to be regular according to the point sets we collected, some regular point processes, including the PHCP, the SP and the perturbed triangular lattice, are considered. We focus on the motion-invariant case of the PPP, the PHCP, and the SP, and the stationary case of the perturbed triangular lattice. A point process is stationary if its distribution is translation-invariant and isotropic if its distribution is rotationally invariant with respect to rotations about the origin. If a point process is both stationary and isotropic, then it is motion-invariant. A stationary PPP is motion-invariant and also said to be homogeneous [5].

2.2.2 The Poisson Point Process

Definition 1 (Poisson point process)**.** *The PPP with intensity λ is a point process $\Phi \in \mathbb{R}^2$ so that 1) for every bounded closed set B, $\Phi(B)$ follows a Poisson distribution with mean $\lambda|B|$ (where $|\cdot|$ is the Lebesgue measure in two dimensions and λ is the expected number of points per unit area), 2) $\Phi(B_1), \Phi(B_2), \ldots, \Phi(B_m)$ are independent if $B_1, B_2, \ldots, B_m$ are disjoint.*

2.2.3 The Strauss Process

The SP constitutes an important class of Gibbs processes. Loosely speaking, Gibbs processes can be obtained by shaping the distribution of a PPP using a density function $f(\varphi)$ on the space of counting measures $\mathcal{N}$. The density function is also called the *likelihood* function. Suppose $f(\varphi)$ is a function such that $f(\varphi) > 0$ implies $f(\varphi') > 0$ whenever $\varphi' \subseteq \varphi$, and Q is the distribution of a PPP with intensity $\lambda = 1$. Regarding φ as a counting measure, we have $\int_{\mathcal{N}} Q(d\varphi) = 1$. If $\int_{\mathcal{N}} f(\varphi)Q(d\varphi) = 1$, then the probability measure $P(Y)$ on the measurable space $(\mathcal{N}, \mathfrak{N})$ that satisfies $P(Y) = \int_Y f(\varphi)Q(d\varphi)$, $\forall Y \in \mathfrak{N}$, is the distribution of a Gibbs process.

Definition 2 (Strauss process)**.** *The SP is a Gibbs process with a density function* $f : \mathcal{N} \mapsto \mathbb{R}^+$ *with*

$$f(\varphi) = ca^{\varphi(\mathbb{R}^2)} \exp(-bt_{\tilde{R}}(\varphi)), \tag{2.1}$$

where $a, \tilde{R} > 0$*,* $b \in \mathbb{R}^+ \cup \infty$*, c is a normalizing constant, and* $t_{\tilde{R}}(\varphi)$ *is the number of point pairs* $\{x, y\}$ *of* φ *with* $\|x - y\| < \tilde{R}$*.*

$\tilde{R}$ is called the interaction radius. b determines the strength of repulsion between points, which makes the SP suitable for modeling regular point sets. In other words, the SP is a *soft-core process.*

2.2.4 The Poisson Hard-core Process

Just as the name implies, the distance between any two points of the PHCP is larger than a constant R, which is called the hard-core distance.

Definition 3 (Poisson hard-core process)**.** *The PHCP is a special case of the SP. Its density function is obtained by setting* $b = \infty$ *in (2.1), i.e.,*

$$f(\varphi) = \begin{cases} 0 & \text{if } t_R(\varphi) > 0 \\ ca^{\varphi(\mathbb{R}^2)} & \text{if } t_R(\varphi) = 0. \end{cases} \tag{2.2}$$

2.2.5 The Perturbed Triangular Lattice

Definition 4 (Triangular lattice)**.** *The triangular lattice* $\mathbb{L} \subset \mathbb{R}^2$ *is defined as*

$$\mathbb{L} = \{u \in \mathbb{Z}^2 : \mathbf{G}u\}, \tag{2.3}$$

where $\mathbf{G} = \eta \begin{bmatrix} 1 & 1/2 \\ 0 & \sqrt{3}/2 \end{bmatrix}$*,* $\eta \in \mathbb{R}^+$*, is the generator matrix.*

The area of each Voronoi cell is $V = |\det \mathbf{G}| = \eta^2\sqrt{3}/2$, and the density of the triangular lattice is $\lambda_{tri} = V^{-1}$.

The triangular lattice is obviously not stationary. However, we can make it stationary by translating the lattice by a random vector uniformly distributed over the Voronoi cell of the origin. In the rest of the chapter, the triangular lattices considered are all assumed to be stationary.

Definition 5 (Stationary triangular lattice)**.** *Let $V(o)$ be the Voronoi cell of the origin o in $\mathbb{L}$. The stationary triangular lattice is*

$$\Phi = \{u \in \mathbb{Z}^2 : \mathbf{G}u + Y\}, \tag{2.4}$$

where Y is uniformly distributed over $V(o)$.

The perturbed triangular lattice is based on the stationary triangular lattice and is also stationary.

Definition 6 (Perturbed triangular lattice)**.** *Let (X_u), $u \in \mathbb{Z}^2$, be a family of i.i.d. random variables, uniformly distributed on the disk $b(o, R)$. The perturbed triangular lattice, i.e. the triangular lattice with uniform perturbation on the disk $b(o, R)$, is defined as*

$$\Phi = \{u \in \mathbb{Z}^2 : \mathbf{G}u + Y + X_u\}. \tag{2.5}$$

2.3 Fitting By Pseudolikelihood Maximization

In this section, in order to find an accurate model, different point processes (the PPP, the PHCP, and the SP) are fitted to the point sets in Table 1.1 using the method of maximum pseudolikelihood, which is a common fitting method in stochastic geometry. The reason of using this method is that the definitions of the PHCP and the SP are based on their likelihood functions, thus maximizing the likelihood or pseudolikelihood is the most direct way for fitting. Since the likelihood function of the perturbed triangular lattice is generally unknown, it is not considered in this

section. The fitting metric, which is used to compare the models, may be drawn from the classical statistics in stochastic geometry or some statistics relevant in wireless networks.

2.3.1 Fitting Method

For the PPP, the method of maximum pseudolikelihood coincides with maximum likelihood [22, 24]. The likelihood function for the PPP is $f(\varphi) = e^{-(\lambda-1)|W|}\lambda^{\varphi(W)}$, where λ is the intensity and W is the sampling region. The maximum likelihood estimate is $\hat{\lambda} = \varphi(W)/|W|$.

For the PHCP, R is decided by the method of maximum profile pseudolikelihood [22], which means for different values of R, we obtain their corresponding fitted PHCP models by the method of maximum pseudolikelihood and select the value of R whose fitted PHCP model has the largest maximum pseudolikelihood. The other parameters in (2.2) are obtained by fitting using the method of maximum pseudolikelihood given R.

For the SP, $\tilde{R}$ is selected from the range $[R, 4R]$ by the method of maximum profile pseudolikelihood. By fitting, a and b in the SP model (2.1) can then be obtained. We choose the lower bound to be R because the maximum pseudolikelihood of the SP model with $\tilde{R} < R$ would be smaller than that of the model with $\tilde{R} = R$, where R is the minimum distance between each two point of the point set. We choose the upper bound to be $4R$ because: 1) the base stations in the data sets we considered are located in a finite region and the SP model with an $\tilde{R}$ that is too large is inaccurate to model the data; 2) $4R$ is a large enough upper bound, since, for the fitted SP model, the optimum value is always in the range $[R, 4R]$.

The reason why we use the method of maximum pseudolikelihood instead of maximum likelihood is that the likelihood is intractable for the PHCP and the SP, while, except for the computation of an integral over the sampling region, which can be

approximated by a finite sum, the pseudolikelihood is known. As the conditional intensities take an exponential family form, the pseudolikelihood can then be maximized using standard statistical software for generalized linear or additive models. The simulations are all done with the software **R** [25].

We use the function *ppm* of the package "Spatstat" in **R** to perform the fitting, which contains the implementation of the method of maximum pseudolikelihood. It is closely related to the model fitting functions in **R** such as *lm* (for linear models) and *glm* (for generalized linear models). The method is computationally efficient. Details are provided in [22]. The computation times depend on the fitting model type and the number of points. The following are typical times for a standard computer (2.3 GHz processor, 4 GB memory). Consider the point set of the urban region. Fitting the PPP to the 64 points takes about 0.02 seconds; fitting the PHCP (with known R) takes about 0.06 seconds; fitting the SP (with known $\tilde{R}$) takes about 0.05 seconds. In the same region of the point set, fitting the PPP to 640 randomly generated points takes about 0.03 seconds; fitting the PPP to 6400 randomly generated points takes about 0.52 seconds. These numbers indicate that the complexity is $O(n^2)$, perhaps even $O(n \log n)$, as the number of points n grows.

We use the function *profilepl* of the package "Spatstat" in **R**, which uses the method of maximum profile pseudolikelihood, to find the optimal values of the interaction radius $\tilde{R}$ for the SP and the hard-core distance R for the PHCP that give the best fit. In the fitting, we search over a vector whose columns contain values of $\tilde{R}$ or R to find the optimum. Then with the optimal value of $\tilde{R}$ or R, we use the function *ppm* to get the values of other parameters in the equations (1) or (2). The complexity of the fitting is in proportion to that of the function *ppm*.

2.3.2 Classical Statistics

Many statistics can be used to characterize the structure of a point process or a point set, such as the nearest-neighbor distance distribution function $G(r)$ and the empty space function $F(r)$. The J function, $J(r) = (1 - G(r))/(1 - F(r))$, measures how close a process is to a PPP. For the PPP, $J(r) \equiv 1$. $J(r) > 1$ at some r indicates the points are regular at these distances, while $J(r) < 1$ means the points are clustered. Hence, we can easily tell by visual inspection of $J(r)$ whether a point set or a point process is regular or clustered. But it is hard to get more information that can be used to discriminate different regular point processes.

Different from the J function that is related to the inter-point distance, Ripley's K function is related to point correlations. It is a second-order statistic and can be defined as $K(r) = \mathbb{E}[\Phi(b(x, r)) - 1 \mid x \in \Phi]/\lambda$, for $r \geq 0$, where λ is the intensity. $\lambda K(r)$ can be interpreted as the mean number of points $y \in \Phi$ that satisfy $0 < \|y - x\| \leq r$, given $x \in \Phi$. For the PPP, $K(r) = \pi r^2$.

The L function is defined as $L(r) = \sqrt{K(r)/\pi}$. $L(r) < r$ at some r indicates the points are regular at distance r, while $L(r) > r$ means the points are clustered.

Consider the point set of the urban region. The L function of the point set is plotted in Figs. 2.1-2.3 (black solid line). It is seen that the point set is regular for $r < 140$,[1] since $L(r) < r$ for $r < 140$. Clearly, $L(r) = 0$ for $r < 39$, which means no two points are closer than 39. Hence, the point set may be regarded as a realization of a hard-core process with hard-core distance $R = 39$. The grey regions in these figures are the pointwise maximum and minimum of 99 realizations of the fitted PPP, the fitted PHCP and the fitted SP, respectively. The values of the parameters obtained by fitting are $\hat{\lambda} = 4.06 \times 10^{-5}$ for the PPP, $R = 39$ for the PHCP and $\tilde{R} = 63$ for the SP. Note that for the fitted PHCP, the R value coincides with that obtained by

[1]The unit of all distances in this chapter is meter.

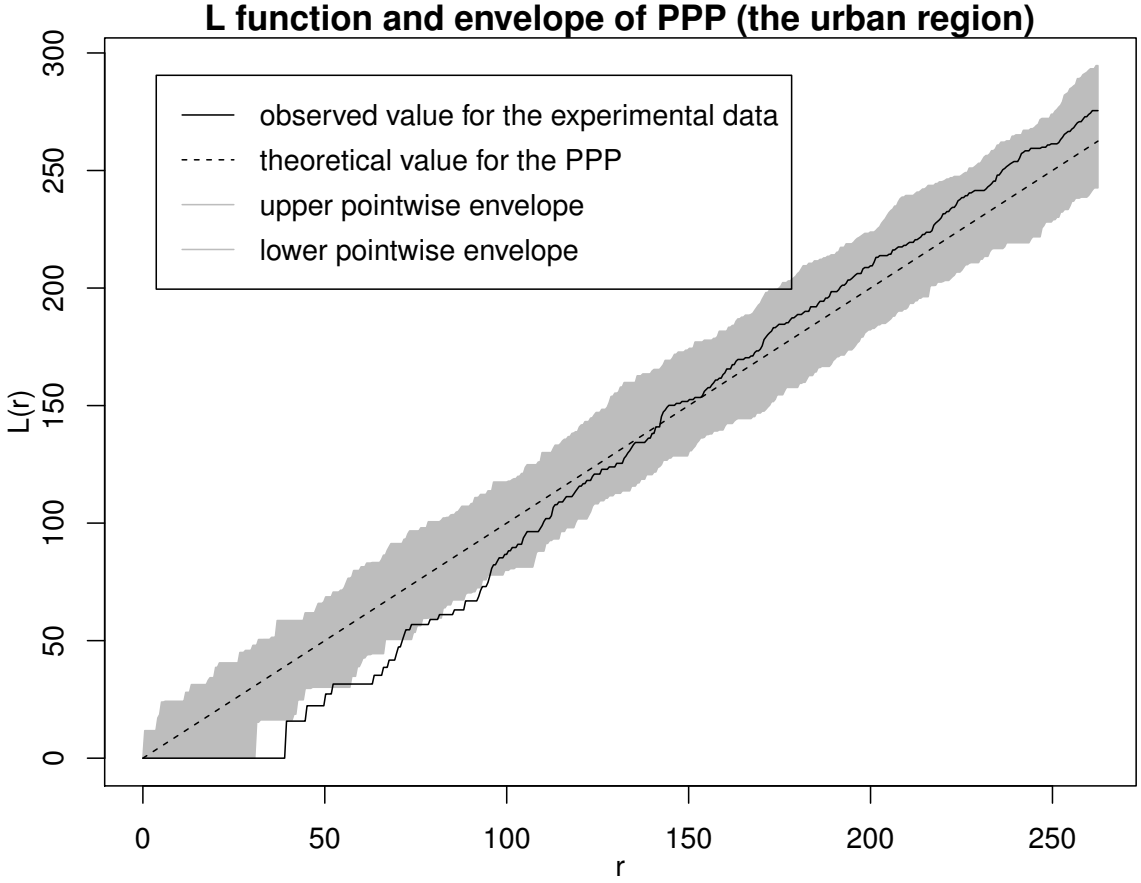

Figure 2.1. L function of BSs of the urban region (the solid line) and the envelope of 99 realizations of the fitted PPP model. The dashed line is the theoretical L function of the PPP.

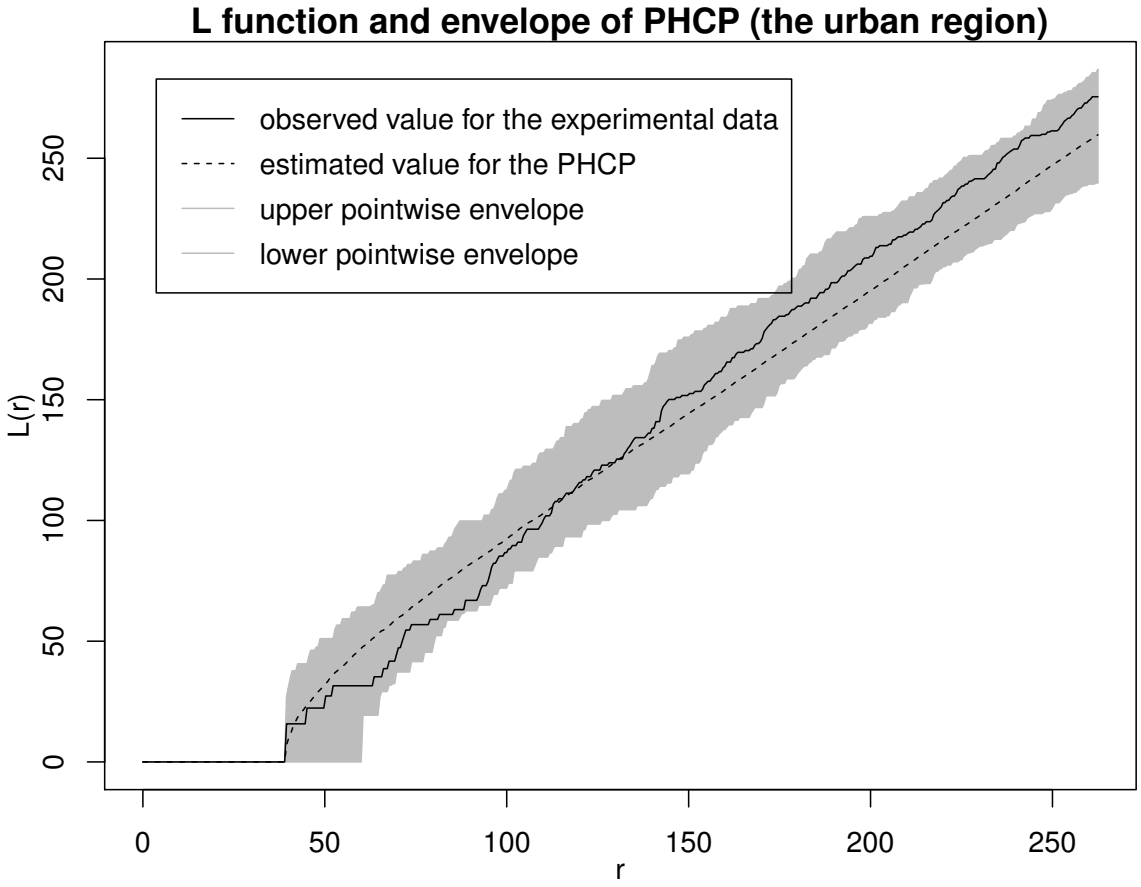

Figure 2.2. L function of BSs of the urban region (the solid line) and the envelope of 99 realizations of the fitted PHCP model. The dashed line is the average value of the L functions of 99 realizations of the fitted PHCP model.

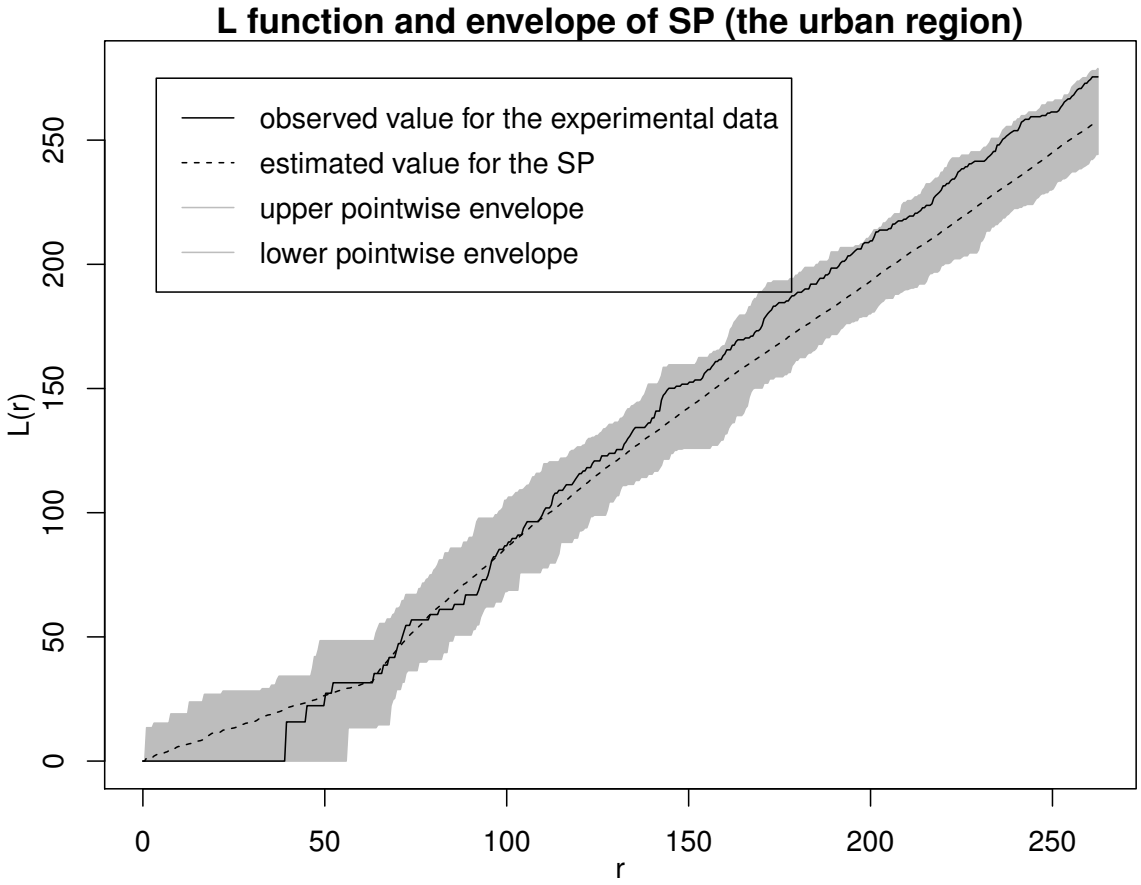

Figure 2.3. L function of BSs of the urban region (the solid line) and the envelope of 99 realizations of the fitted SP model. The dashed line is the average value of the L functions of 99 realizations of the fitted SP model.

visual inspection.

According to the figures, the PPP is not an appropriate model, as the L function of the point set is not within the envelope of the PPP. But the PHCP and the SP fit well. Although the L function is more powerful than the J function when used to compare the three models, it cannot distinguish which of the PHCP and the SP is better. Other statistics are needed.

2.3.3 Definition of Coverage Probability

It is sensible to use a statistic that is related with a standard metric used in wireless networks to decide on the best model. For the three point sets, by evaluating the coverage probabilities of the point sets and fitted models through simulations, we find that the coverage probability is such a statistic that has two desirable properties: it has enough discriminative power to distinguish between different models, and it is relevant to cellular systems. Generally speaking, the coverage probability is the probability that a randomly located user achieves a given SINR threshold with respect to one of the BSs.

A mobile user is assumed to attempt to communicate with the nearest BS, while all other BSs act as interferers (the frequency reuse factor is 1). The received power, the interference, and, in turn, the coverage probability, depend on the transmit power of the BSs, the power loss during propagation, and the random channel effects. We make the following assumptions: (i) the transmit power of all BSs is constant 1; (ii) the path loss exponent $\alpha = 4$; (iii) all signals experience Rayleigh fading with mean 1; (iv) the shadowing effect is neglected; (v) the thermal noise W is ignored, i.e. $\mathrm{SNR} = \infty$, and the SINR reduces to the SIR.

Under these assumptions, the SIR has the form

$$\mathrm{SIR}_z = \frac{h_0 \|x_0\|^{-\alpha}}{\sum_{i:x_i \in \Phi \setminus \{x_0\}} h_i \|x_i - z\|^{-\alpha}}, \tag{2.6}$$

where $\{h_0, h_1, ...\} \sim$ exponential(1) and independent, and $x_0 = \arg\min_{x\in\Phi} \|x - z\|$. We assume that the location z is in coverage if $\mathrm{SIR}_z > \theta$.

Definition 7 (Coverage probability)**.** *For a stationary process,* $\mathbb{P}(SIR_z > \theta)$ *does not depend on* z*, and we call it the coverage probability:*

$$P_{\mathrm{c}}(\theta) = \mathbb{P}(SIR > \theta). \tag{2.7}$$

It is the CCDF of the SIR and can also be interpreted as the average area fraction in coverage.

The theoretical expression of $P_{\mathrm{c}}(\theta)$ for the PPP with intensity λ has been derived in [8]:

$$P_{\mathrm{c}}(\theta) = \frac{1}{1 + \sqrt{\theta}\arctan(\sqrt{\theta})}. \tag{2.8}$$

Since the coverage probability of the PPP does not depend on the intensity, no fitting method based on adjusting the intensity is possible. On the other hand, the intensity is easily matched to the intensity of a given point set.

2.3.4 Results for Coverage Probability

The regions where the BSs reside are not infinite. Thus, for the fitted point process, which is stationary, we only consider a finite region that has the same area and shape as the point set under consideration.

In the finite region, $P_{\mathrm{c}}(\theta)$ can be estimated by determining the average fraction of the whole area where $\mathrm{SIR} > \theta$. In the following simulations, $P_{\mathrm{c}}(\theta)$ is obtained by evaluating 3,000,000 SIR values. In order to mitigate the boundary effect, we only use the central $[\frac{2}{3}\mathrm{length} \times \frac{2}{3}\mathrm{width}]$ area of the entire region to compute $P_{\mathrm{c}}(\theta)$. For the point sets, the SIRs of 3,000,000 randomly chosen locations (uniformly distributed) are computed. For point processes, 3,000 realizations are generated and for each

realization, 1,000 randomly chosen locations are generated. The SIR is evaluated at all chosen locations.

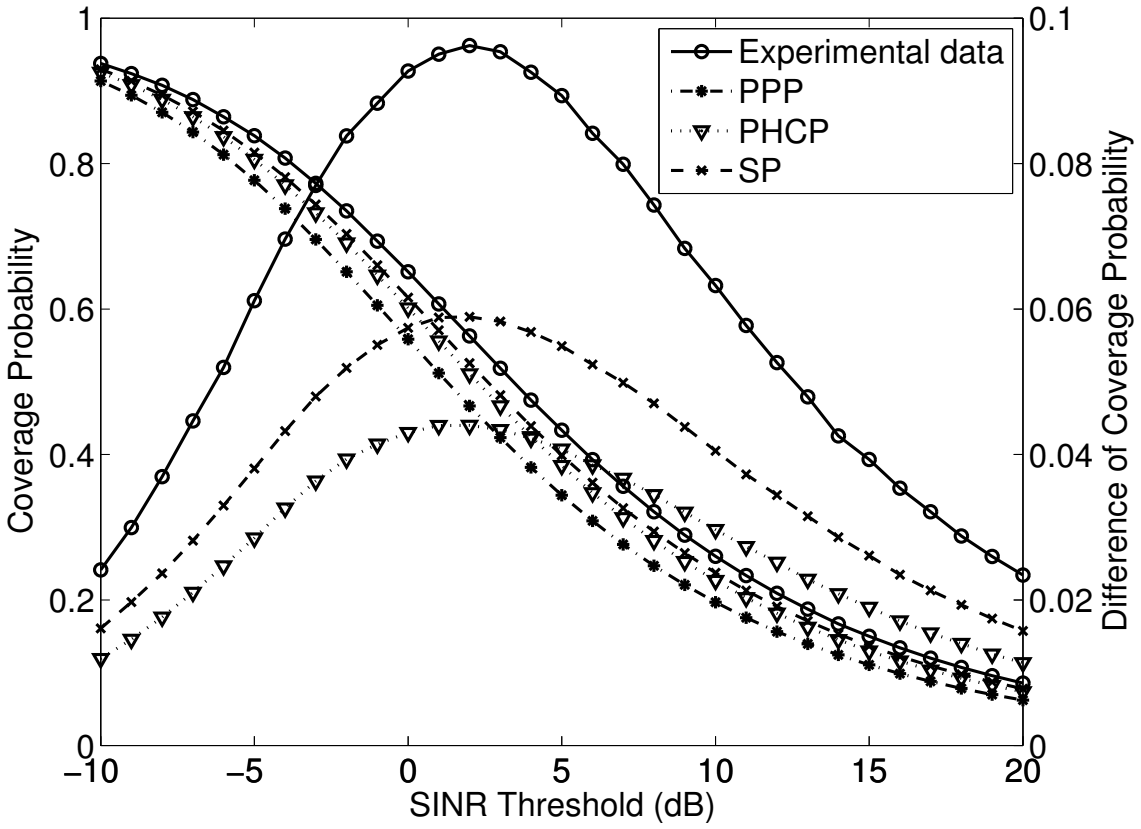

Figure 2.4. Left axis: the coverage curves of the experimental data of the urban region and different fitted point process models. Right axis: the difference between the coverage curve of the PPP and the other curves.

Consider the point set of the urban region. The coverage curves of the experimental data and the fitted models of the PPP, the PHCP, and the SP are shown in Fig. 2.4. The left axis shows the coverage curves, while the right axis shows the difference between the coverage curve of the PPP and the other curves. Clearly, the

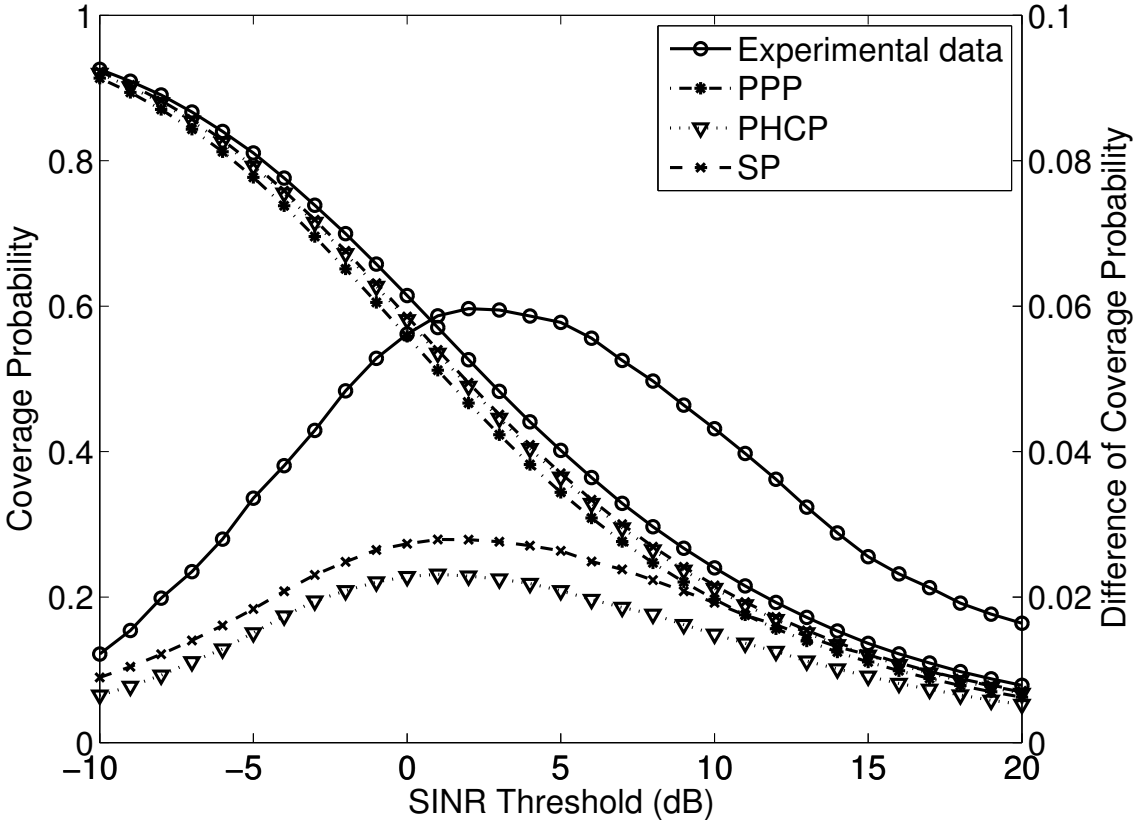

Figure 2.5. Left axis: the coverage curves of the experimental data of the rural region 1 and different fitted point process models. Right axis: the difference between the coverage curve of the PPP and the other curves.

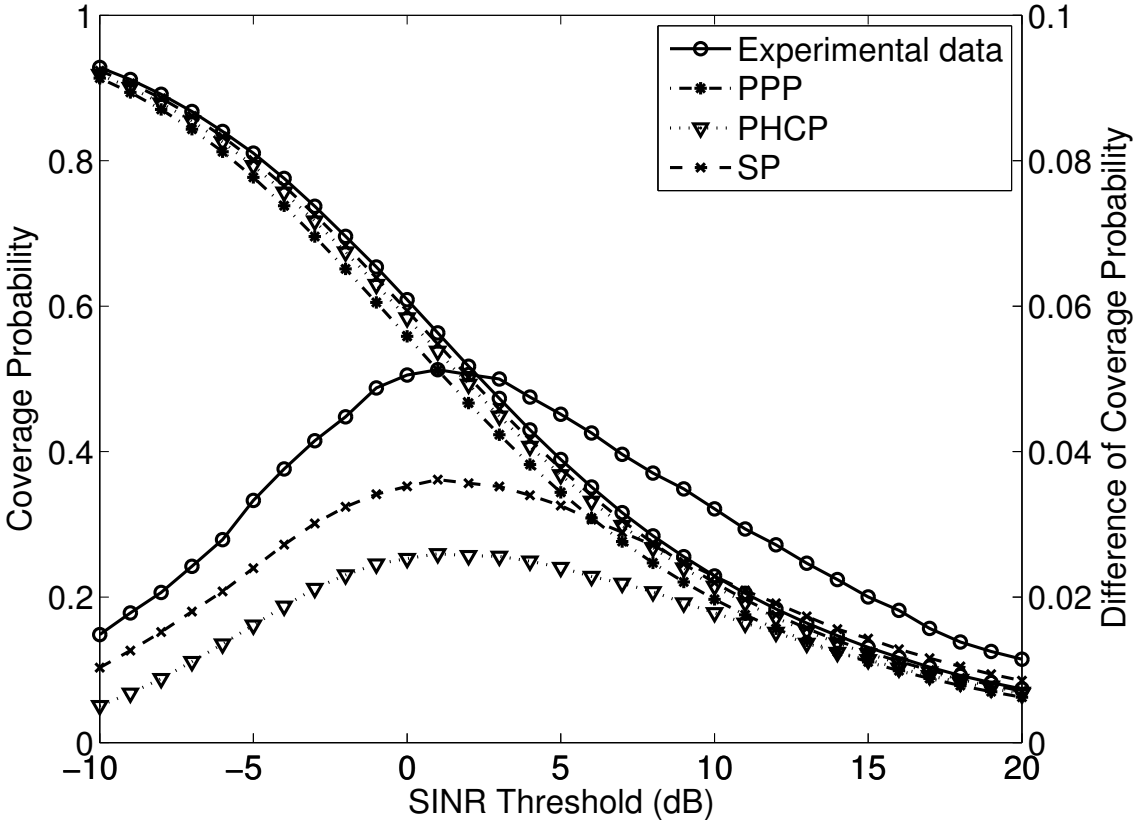

Figure 2.6. Left axis: the coverage curves of the experimental data of the rural region 2 and different fitted point process models. Right axis: the difference between the coverage curve of the PPP and the other curves.

curves of three models are all below the curve of the experimental data. Among the three point processes, the SP provides the best fit, followed by the PHCP and then the PPP.

We use the other two point sets in Table 1.1 to test the statistic. For the fitted models, the hard-core distances in the two rural regions are $R_1 = 1194$ and $R_2 = 1474$ and the interaction radii are $\tilde{R}_1 = 2120$ and $\tilde{R}_2 = 5490$. Figs. 2.5 and 2.6 show the coverage curves of the two point sets. The PPP still performs the worst. In Fig. 2.6, the SP is better than the PHCP, while in Fig. 2.5, the curves of the SP and the PHCP are quite close, thus, the two processes can be considered equivalent when fitted to that point set. Generally, it depends on the given point set. The SP is often better. Note that this is not because the PHCP is a special case of the SP. The method of maximum pseudolikelihood is used to do the fittings, but a larger pseudolikelihood does not imply a better matching coverage probability.

2.3.5 Average Rate

We can also distinguish the best fitted model in terms of the average ergodic rate. Similar results are obtained. The average ergodic rate (or Shannon throughput) is defined as $\bar{\gamma} = \mathbb{E}[\ln(1 + \mathrm{SIR})]$, which is measured in nats/s/Hz. Denote $\bar{\gamma}_e$, $\bar{\gamma}_p$, $\bar{\gamma}_h$, $\bar{\gamma}_s$ as the average ergodic rates of the experimental data, the PPP, the PHCP, and the SP respectively. Let the simulation parameters remain the same. For the point set of the urban region, $\bar{\gamma}_e \approx 1.786$, $\bar{\gamma}_p \approx 1.513$, $\bar{\gamma}_h \approx 1.635$, $\bar{\gamma}_s \approx 1.682$. For the point set of the rural region 1, $\bar{\gamma}_e \approx 1.679$, $\bar{\gamma}_p \approx 1.506$, $\bar{\gamma}_h \approx 1.566$, $\bar{\gamma}_s \approx 1.572$. For the point set of the rural region 2, $\bar{\gamma}_e \approx 1.634$, $\bar{\gamma}_p \approx 1.515$, $\bar{\gamma}_h \approx 1.581$, $\bar{\gamma}_s \approx 1.605$. So we have $\bar{\gamma}_p < \bar{\gamma}_h < \bar{\gamma}_s < \bar{\gamma}_e$.

The theoretical average ergodic rate of the PPP is $\bar{\gamma}'_p \approx 1.49$, which is smaller than the values of simulations of the PPP. The reason is that the theoretical average ergodic rate of the PPP considers all the points on the whole plane $\mathbb{R}^2$, while in the

simulations, we only consider the points of the PPP in the same region of the given point set, which leads to the fact that the SIR value at any location is larger than the theoretical value at the same location and thus the theoretical average ergodic rate is smaller than the simulation values.

2.4 Fitting Using the Coverage Probability

We have fitted the PPP, the PHCP, and the SP to the given point sets by the method of maximum pseudolikelihood, but none of these models precisely describes the coverage probability of the data, and all their coverage curves are *below* the actual curve of the point set. If we want to find a point process that has a similar performance as the given point set, we cannot just use the three fitted models, because they are all not regular enough due to the limitation of the fitting methods. In this section, we adopt the method of minimum contrast as a fitting method and fit the SP, the PHCP, and the perturbed triangular lattice to the point sets in Table 1.1.

2.4.1 Fitting Method

In the method of minimum contrast, there is a suitable summary statistic S and a point process model with some adjustable parameters $\{\theta_i\}$. Ideally the chosen point process model has analytically tractable expressions for the summary statistic S as a function of $\{\theta_i\}$.

Here, the summary statistic is the coverage probability, and the chosen models are the SP, the PHCP and the perturbed triangular lattice. But there are no analytically tractable expressions for the coverage probability of the SP and the PHCP. For different values of the adjustable parameters, the coverage probabilities are estimated through simulations.

We assume the intensity of the fitted model is the same as the given point set. By this method, the coverage curve of the fitted model should have the minimum

difference from that of the given point set.

Definition 8 (Average squared error of the coverage probability)**.** *The average squared error of the coverage probability, denoted as E, measures the difference between two coverage curves. It is defined as:*

$$E(a,b) = \frac{1}{b-a}\int_a^b \Big(P_{c1}(t) - P_{c2}(t)\Big)^2 dt, \tag{2.9}$$

where $a, b \in \mathbb{R}$, t is the SIR threshold in dB, and $P_{c1}(t)$, $P_{c2}(t)$ denote two coverage curves.

The average squared error of the coverage probability is used as the contrast criterion of the method of minimum contrast. Under the condition of the fixed intensity, the relevant parameters in the model are adjusted to find the model that has the minimum average squared error between its coverage curve and the given point set's. Here, we set $a = -9.38$ dB and $b = 16.07$ dB (for the PPP, $P_{\mathrm{c}}(a) = 0.9$ and $P_{\mathrm{c}}(b) = 0.1$), because $[0.1, 0.9]$ is the coverage probability range where the curves differ the most and $[-10, 16]$ dB is a reasonable SIR interval for practical systems.

This fitting method is not restricted to the contrast criterion defined by the coverage probability. The criteria defined by other performance metrics in wireless networks and second-order statistics in stochastic geometry can also be used. Similarly, the method is not limited to Rayleigh fading either, when we simulate the network. Other fading types can also be applied depending on the propagation environment.

2.4.2 The SP and the PHCP

In the fitting method, the intensity of the fitted model is fixed. Thus, the PPP is not considered. As the accurate intensity values of the SP and the PHCP are unknown for given values of the parameters in (2.1) and (2.2), it is not quite suitable to use the method for the two processes. But there are some approximations of the

intensity for the SP [26], e.g.,

$$\lambda \approx W(a\Gamma)/\Gamma, \tag{2.10}$$

where $W(x)$ is the principle branch of the *Lambert W function* [27] and $\Gamma = \big(1 - \exp(-b)\big)\pi\tilde{R}^2$. This is the Poisson-saddlepoint approximation [26], which is more accurate than the mean field approximation.

If we use the approximated intensity (2.10) in the fitting method for the SP, we have to adjust the three parameters a, b, and $\tilde{R}$ in (2.1) to minimize the average squared error of the coverage probability. Note that, as b increases, the strength of the repulsion between the points in the SP increases, and as $\tilde{R}$ increases, the repulsion range increases. Both adjustments increase the regularity of the process. From (2.10), we have $a \approx \lambda \exp(\lambda\Gamma)$. a increases as b and $\tilde{R}$ increase with λ fixed. So in order to increase the regularity of the SP with fixed intensity, we can fix b, increase $\tilde{R}$ and a, or fix $\tilde{R}$, increase b and a according to (2.10). We can also first increase a, and then adjust b and $\tilde{R}$. But in this way, the regularity may not increase, or even decrease for some b and $\tilde{R}$. To get a more regular model, we can compare models with different settings of b and $\tilde{R}$ in simulations. The above three methods are used to obtain the fitting results of the SP in simulations.

TABLE 2.1

FITTING RESULTS OF THE STRAUSS PROCESS

Parameters	a	b	$\tilde{R}$	Actual intensity $\hat{\lambda}$	Desired intensity λ	$\hat{\lambda}/\lambda - 1$
Urban region	1×10^{-4}	3.745	85	3.737×10^{-5}	4.063×10^{-5}	-8.02%
Rural region 1	2.44×10^{-8}	1.892	3000	1.622×10^{-8}	1.645×10^{-8}	-1.40%
Rural region 2	5.00×10^{-8}	0.599	5490	2.086×10^{-8}	2.069×10^{-8}	0.82%

Given a point set, to obtain a fitted SP, we can first fit the SP to the point set using the method of the maximum pseudolikelihood, and then, based on the parameters we get, increase the regularity to minimize the average squared error of the coverage probability. Table 2.1 shows the fitting results of the SP for the three point sets in Table 1.1. As shown in Fig. 2.7, for each fitted model, the coverage curve matches the one of the corresponding point set very closely. Note that the simulation is not perfectly accurate, since the number of realizations of the point process used to calculate the coverage probability is limited to 3,000; also, when calculating the average squared error of the coverage probability, we only compute the average over a finite number of sample points on the coverage curve; and when we increase b and $\tilde{R}$, the step width is not infinitesimal. We say an SP model has the "minimum" average squared error of the coverage probability, if $E < 10^{-5}$.

The fitted SP is not unique. For some different values of a, we can find different fitted models that satisfy $E < 10^{-5}$, by adjusting b and $\tilde{R}$. For instance, the SP with $a = 1.1 \times 10^{-4}$, $b = 2.547$, $\tilde{R} = 92$ is also a fitted model for the urban region, which is shown as the curve of another fitted SP model in Fig. 2.7.

Since the PHCP is a special case of the SP, its approximated intensity can be obtained by setting $b = \infty$ in (2.10), $\lambda \approx W(a\pi R^2)/(\pi R^2)$. To increase the regularity of the PHCP with fixed intensity, we can increase R. Table 2.2 shows the fitting results of the PHCP for the three point sets. The coverage curves of the fitted models and their corresponding point sets are visually indistinguishable, as shown in Fig. 2.8.

Although the models are fitted well to the point sets, there are two main shortcomings of the fitting for the SP and the PHCP. One is that the actual intensity is not the same as the density of the given point set as shown in Tables 2.1 and 2.2, and the difference can be as large as 10%. Note that each value of the actual intensity is obtained by averaging over 10,000 independent realizations of the model.

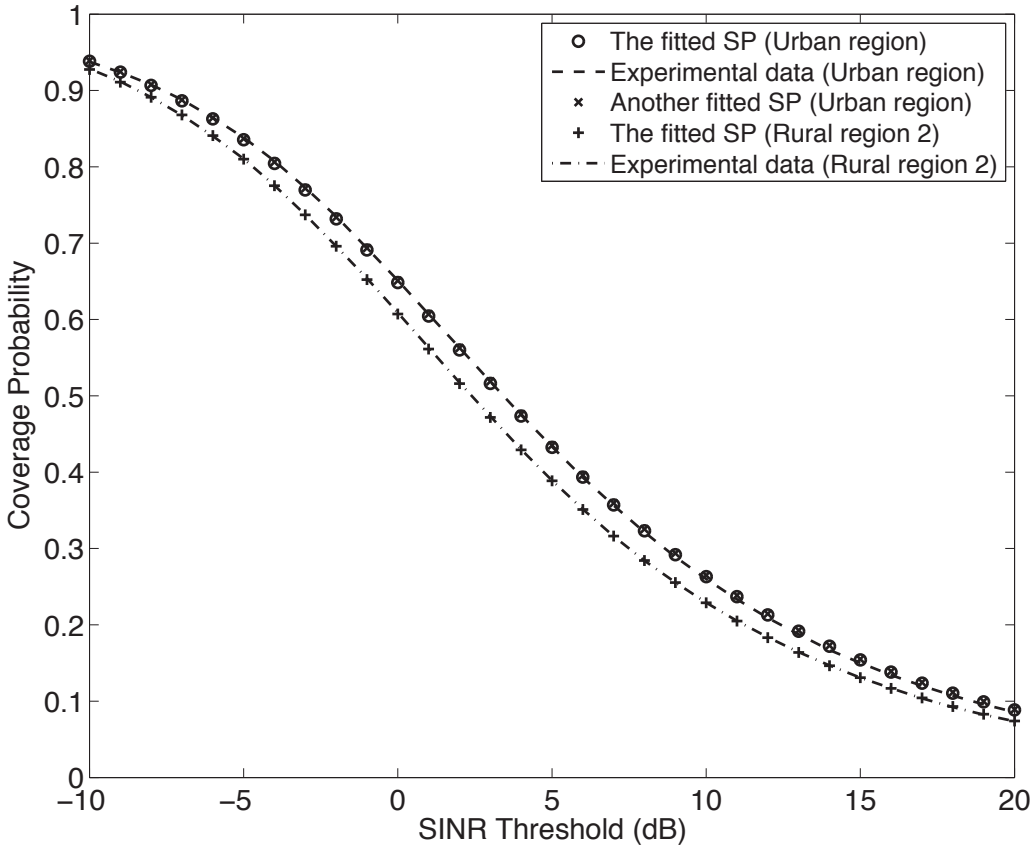

Figure 2.7. The coverage curves of the experimental data and the fitted SP models. The curves of the rural region 1, not shown in this figure, are very similar to those of the rural region 2.

TABLE 2.2

FITTING RESULTS OF THE POISSON HARD-CORE PROCESS

Parameters	a	R	Actual intensity $\hat{\lambda}$	Desired intensity λ	$\hat{\lambda}/\lambda - 1$
Urban region	9.38×10^{-5}	78	3.885×10^{-5}	4.063×10^{-5}	-4.38%
Rural region 1	2.28×10^{-8}	2500	1.626×10^{-8}	1.645×10^{-8}	-1.16%
Rural region 2	2.37×10^{-8}	2000	1.864×10^{-8}	2.069×10^{-8}	-9.91%

The other drawback is that we may not get a well fitted model for some point sets. In simulations, we use the function *rStrauss* in the **R** package "Spatstat" [28] to generate realizations for the SP and the function *rHardcore* for the PHCP. In *rStrauss* and *rHardcore*, the coupling-from-the-past (CFTP) algorithm [29] is used, but it is not practicable for all parameter values. Its computation time and storage increase rapidly with a, $\tilde{R}$ and R. For example, for a point set that has a coverage curve close to that of the triangular lattice, we cannot get the fitted SP or PHCP, due to the limited storage and time. It turns out, though, that the three point sets in Table 1.1 are not too regular to use *rStrauss* and *rHardcore*.

2.4.3 The Perturbed Triangular Lattice

There are no such shortcomings described in the previous subsection when the perturbed triangular lattice is fitted by the method of minimum contrast. The reasons are 1) the intensity is fixed once η is fixed; 2) as R increases from 0 to ∞, the coverage curve of the perturbed triangular lattice degrades from that of the triangular lattice to that of the PPP, and we can easily get the realizations of the perturbed triangular lattice for all values of η and R. To do the fitting, we first compute η and then

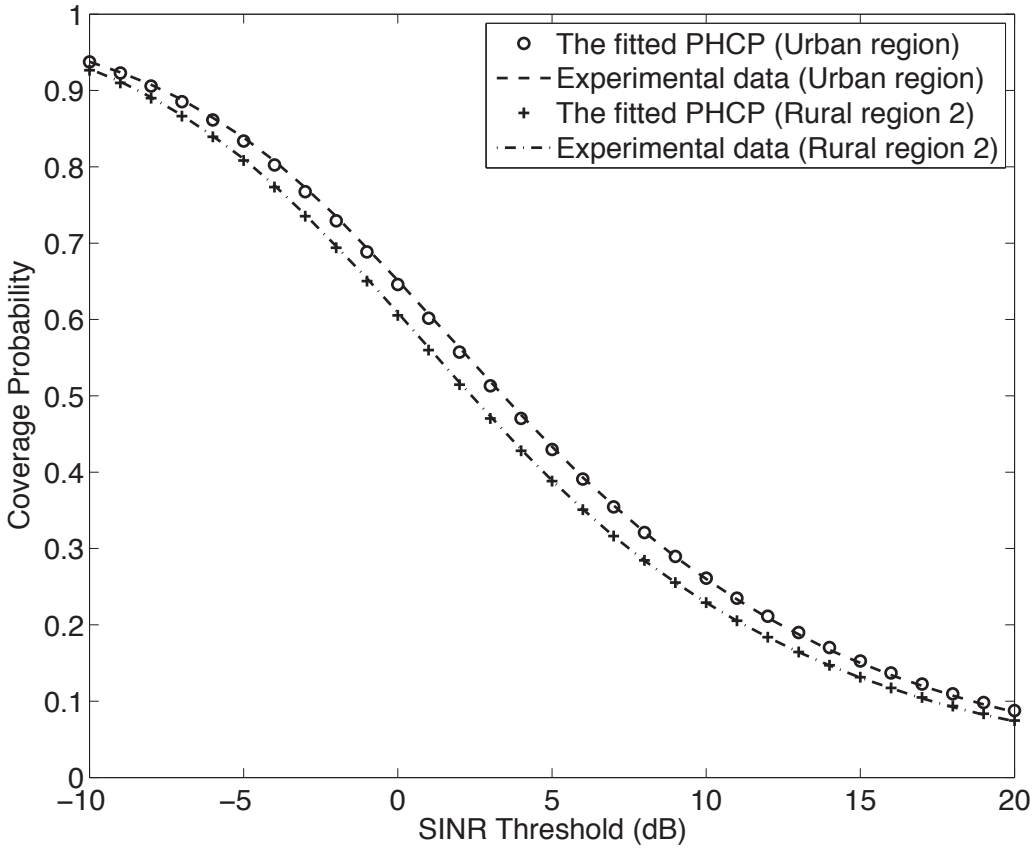

Figure 2.8. The coverage curves of the experimental data and the fitted PHCP models. The curves of the rural region 1, not shown in this figure, are very similar to those of the rural region 2.

increase R from 0 to find the fitted model.

Consider the point set of the urban region. The intensity of the point set is $\hat{\lambda} = 4.06 \times 10^{-5}$. Equating $\lambda_{tri} = \hat{\lambda}$, we get $\eta = 168.57$. Fig. 1.1 shows the locations of the BSs in the urban region. Figs. 2.9-2.11 give the realizations of the fitted PPP, the triangular lattice, and the triangular lattice with uniform perturbation on the disk $b(o, 0.52\eta)$, respectively. Note that the fitted PPP means that the intensity of the PPP is estimated by the method of maximum likelihood. To compute the coverage probability of the triangular lattice with $\eta = 168.57$, the lattice is generated on the same region as the point set. Under the same simulation conditions as those in Section III, the coverage probability is obtained, which is shown in Fig. 2.12. As expected, the coverage probability of the lattice is larger than that of the given point set. The lattice provides an upper bound on the coverage probability.

To compare the coverage performances of the perturbed triangular lattices with the PPP and the triangular lattice, we simulate the cases with $R = 0.2\eta, 0.5\eta$ and 0.8η. Fig. 2.13 shows the coverage curves. As expected and observed in the figure, the coverage probability degrades as R increases. As $R \to \infty$, the perturbed triangular lattice approaches the PPP with intensity $\lambda = 4.06 \times 10^{-5}$. Therefore, the coverage curves of the perturbed triangular lattices with different R span the region between the PPP and the triangular lattice. It is thus guaranteed that we can obtain the desired perturbed triangular lattice that is fitted tightly to a point set.

For the point set of the urban region, the fitting value of R is $R = 0.52\eta$. Fig. 2.10 indicates that the disks centered at the triangular lattice points with radii 0.52η overlap slightly, as the distance between each two triangular lattice points is η. In Fig. 2.11, a realization of this perturbed triangular lattice is shown. The coverage curves of this perturbed triangular lattice and the point set closely overlap, as shown in Fig. 2.12. For the point sets of the rural region 1 and the rural region 2, the fitting values are $R_1 = 0.70\eta$ and $R_2 = 0.74\eta$, respectively. So the point set of the urban

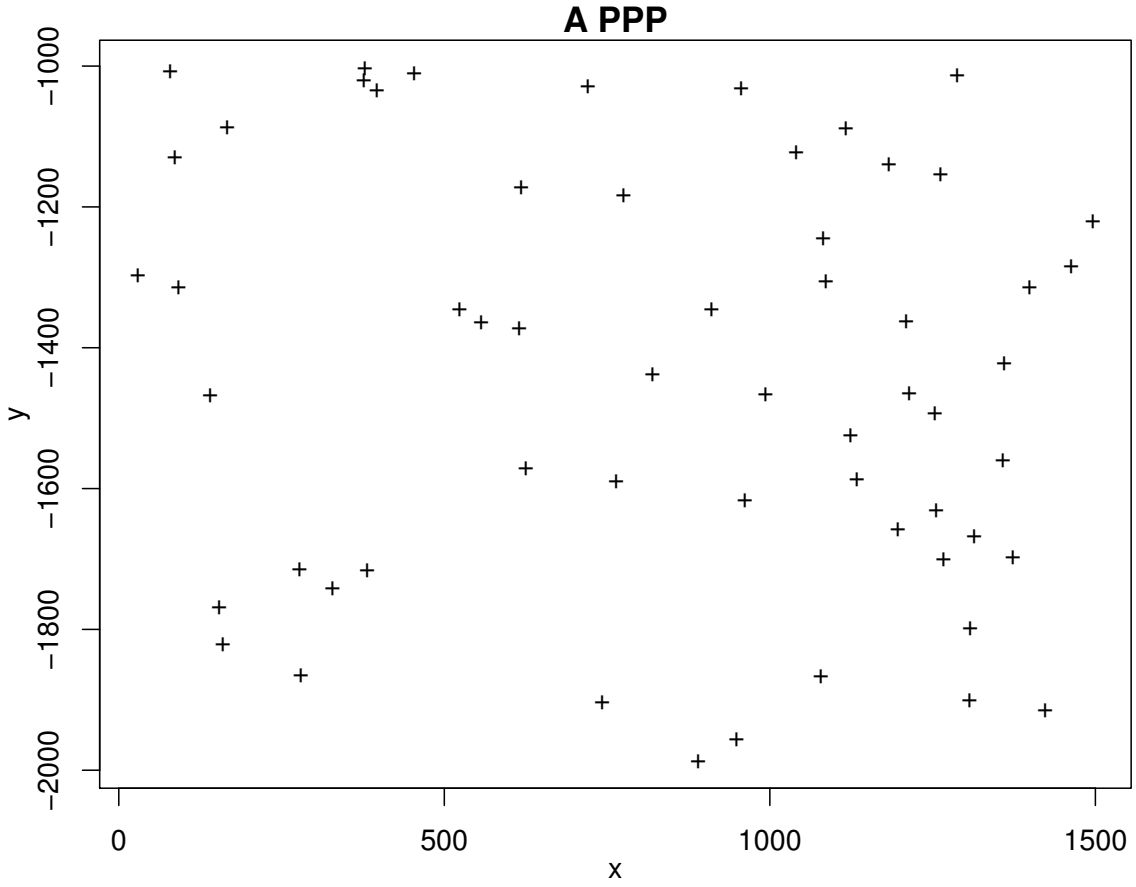

Figure 2.9. A realization of the PPP fitted to the urban data set.

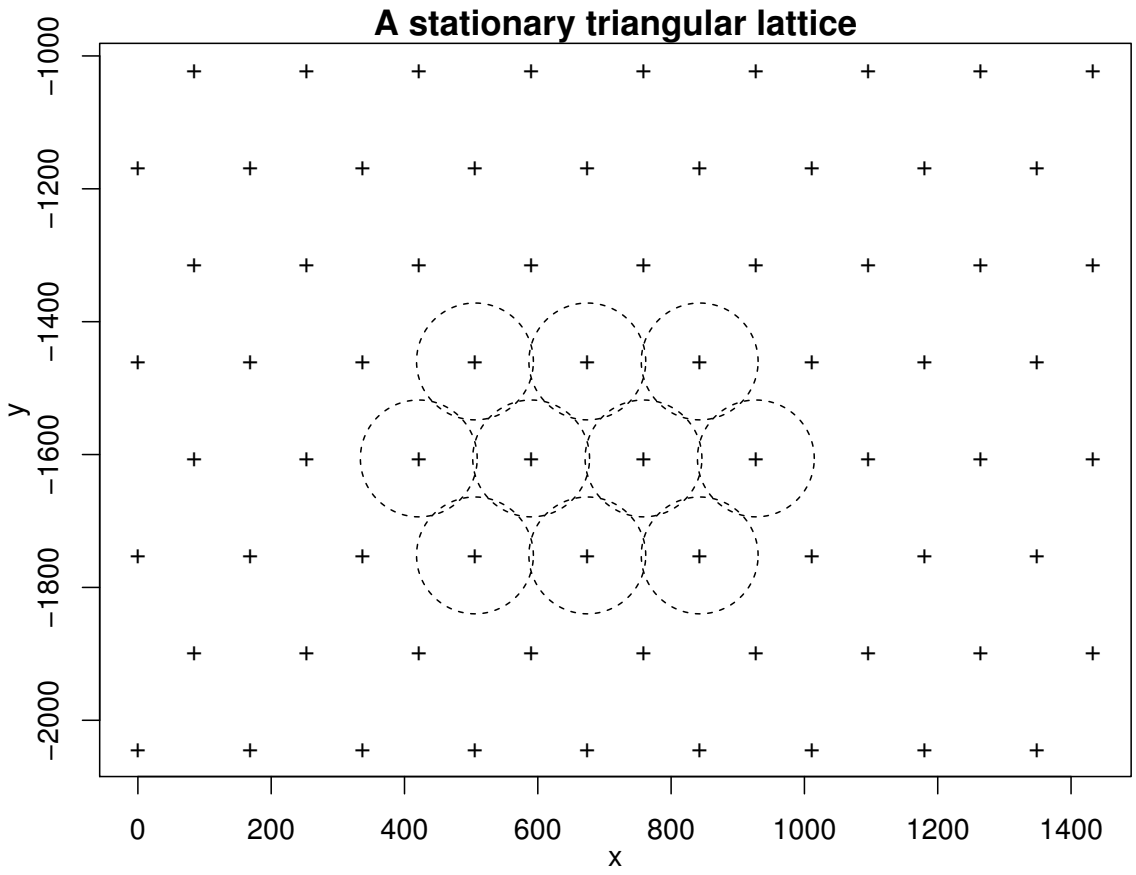

Figure 2.10. A realization of the triangular lattice on the urban region. The dashed disks have centers at the lattice points and their radii are 0.52η.

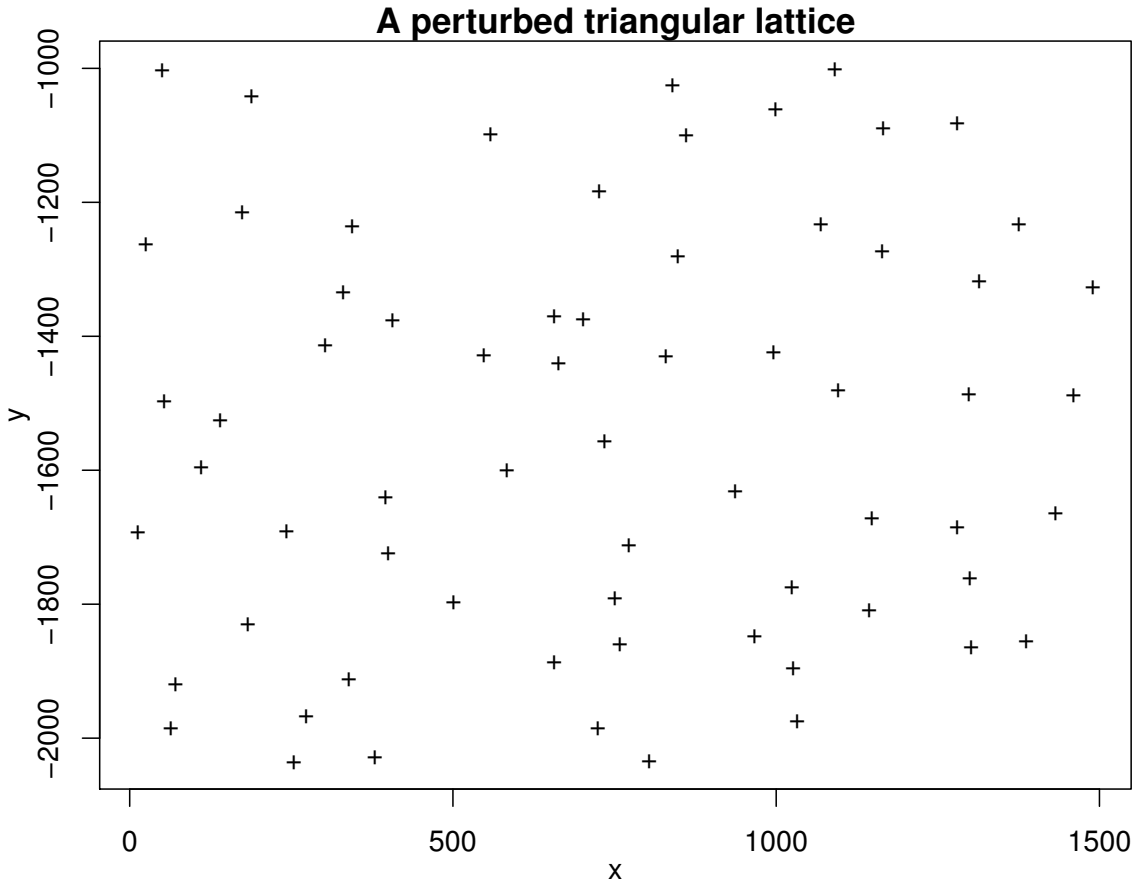

Figure 2.11. A realization of the triangular lattice with uniform perturbation on the disk $b(o, 0.52\eta)$ fitted to the urban data set.

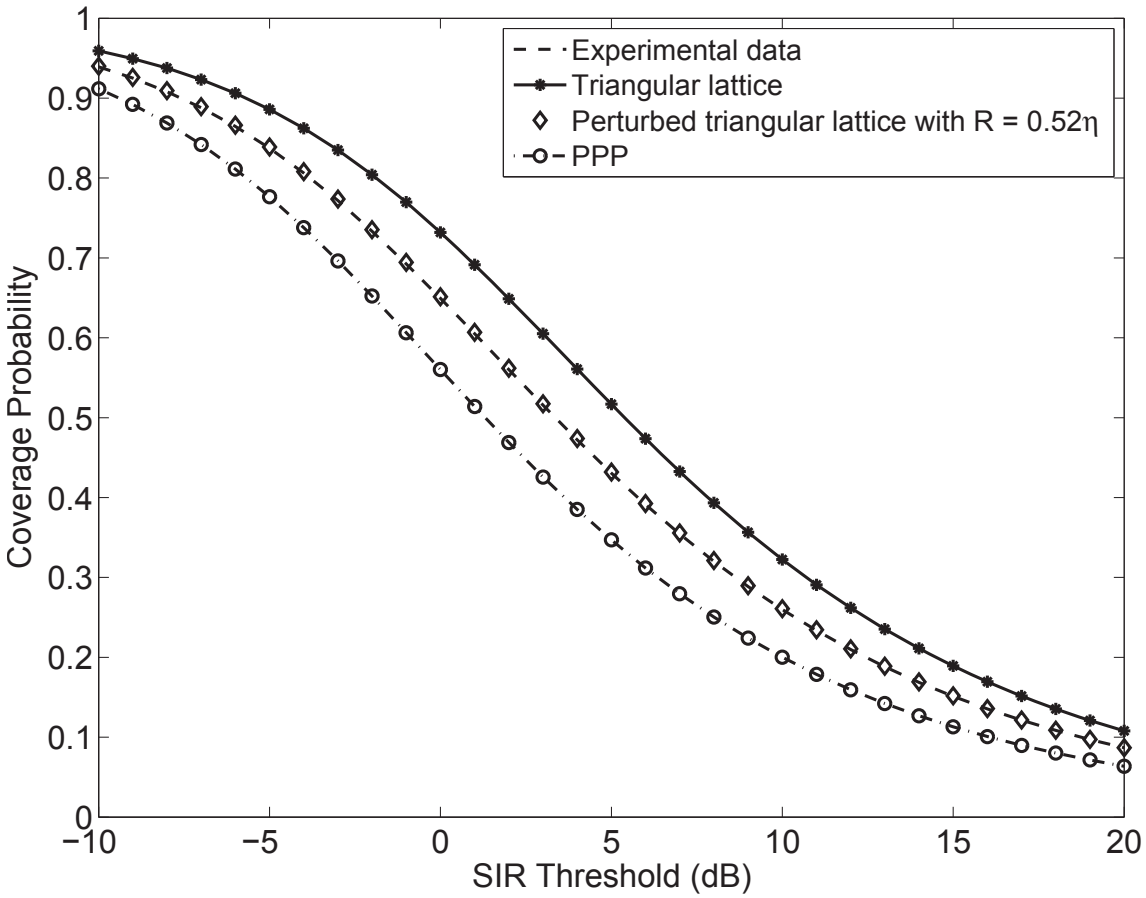

Figure 2.12. The coverage curves of the experimental data (the urban region), the triangular lattice, the triangular lattice with uniform perturbation on the disk $b(o, 0.52\eta)$ and the PPP.

region is the most regular of the three, followed by the point set of the rural region 1 and then the point set of the rural region 2.

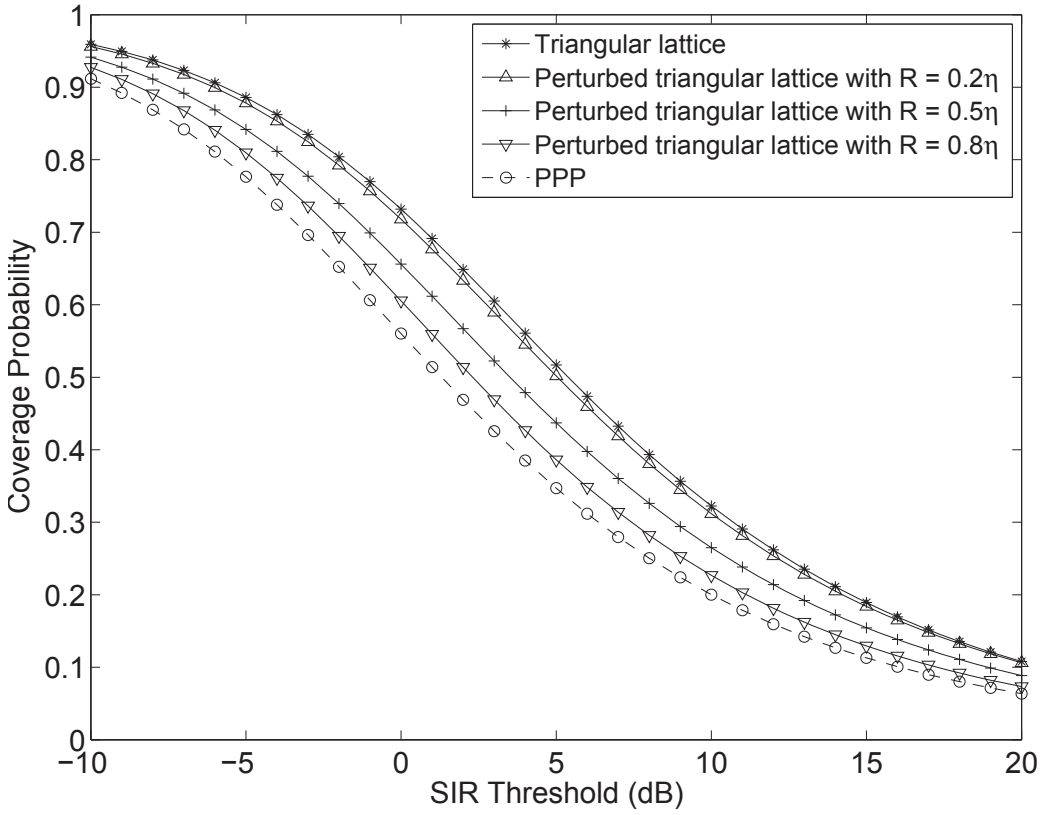

Figure 2.13. The coverage curves of the triangular lattice, the perturbed triangular lattices and the PPP.

To obtain a point set from the model that has approximately the same performance of the coverage probability as the given point set, we can generate a realization of the triangular lattice with uniform perturbation on the disk $b(o, R)$. Although the coverage curve of the realization may have some deviations, its average, the coverage probability, is quite exactly that of the point set.

Thus, we can model the given point set as a realization of the triangular lattice with uniform perturbation on the disk $b(o, R)$, where R can be determined by minimizing the average squared error of the coverage probability, which is of great significance in practice. When analyzing performance metrics that are related with the distribution of the BSs in real cellular networks, we can use the perturbed triangular lattice instead of the lattice or the PPP to model the BSs. Although the perturbed triangular lattice is not as tractable as the PPP, it still has some desirable properties. For the PPP, the distribution of the area of the Voronoi cell is usually approximated by a generalized gamma function [11, 30, 31]. The area is unbounded for the PPP, while for the perturbed triangular lattice, the area is bounded and depends on R.

2.5 Deployment Gain

Here we define a metric that measures how close the point set is to the PPP. This metric can be considered as a "distance" between the point set and the PPP whose coverage curve only depends on the SIR threshold θ. We call this metric the *deployment gain*. It is a function of the coverage probability and is a gain in SIR, relative to the PPP, provided by the deployment.

Definition 9 (Deployment gain)**.** *The deployment gain, denoted by $G(p_t)$, is the SIR difference between the coverage curves of the given point set and the PPP at a given target coverage probability p_t.*

As such, it mimics the notion of the coding gain[2] commonly used in coding theory. We can evaluate different deployment gains at different p_t, for different considerations. In the rest of the chapter, we choose $p_t = 0.5$. At this target probability, the coverage curves are steep, and the gap between curves is easy to observe. More importantly,

[2]Coding gain [32, Ch. 1], always a function of the target bit-error-rate (BER), is a measure to quantify the performance of a given code, and is defined by the difference in minimum signal-to-noise-ratio (SNR) required to achieve the same BER with and without the code.

$G(0.5)$ gives a good approximation of the *average deployment gain*, which is the value by which the coverage curve of the PPP is right shifted such that the difference between the new curve and the curve of the point set is minimized.

Definition 10 (Average deployment gain)**.** *Let the difference between two curves be the average squared error defined in (2.9). The average deployment gain, denoted by $\hat{S}_g$, is then defined as:*

$$\hat{S}_g = \arg\min_x \int_a^b \left(P_{\mathrm{c}}^{\mathrm{th}}(t-x) - P_{\mathrm{c}}^{\mathrm{ed}}(t) \right)^2 dt, \tag{2.11}$$

where $a = -9.38$ dB and $b = 16.07$ dB, $P_{\mathrm{c}}^{\mathrm{th}}(t)$ is the theoretical value of the coverage probability for the PPP, and $P_{\mathrm{c}}^{\mathrm{ed}}(t)$ is the experimental value of the coverage probability for the data.

For fixed α, the theoretical expression of the coverage probability of the PPP [8] is

$$P_{\mathrm{c}}^{\mathrm{th}}(\theta) = \frac{1}{1+\rho(\theta,\alpha)}, \tag{2.12}$$

where $\rho(\theta,\alpha) = \theta^{2/\alpha} \int_{\theta^{-2/\alpha}}^{\infty} 1/(1+u^{\alpha/2}) du$. For $\alpha = 4$, $P_{\mathrm{c}}^{\mathrm{th}}(\theta)$ is equal to $P_{\mathrm{c}}(\theta)$ in (2.8).

The average deployment gain $\hat{S}_g$ is a measure of regularity. The point set with a larger average deployment gain has a better performance than the one with a smaller value. For the triangular lattice, when $\alpha = 4$, $\hat{S}_g^l = 4.38$ dB, which is the maximal value of the average deployment gain. Similar to $\hat{S}_g$, $G(0.5)$ is also a measure of regularity and satisfies that $|G(0.5) - \hat{S}_g|/\hat{S}_g < 5\%$, which is verified in simulations. Hence, we can evaluate $G(0.5)$ instead of $\hat{S}_g$ in practice, since $G(0.5)$ is much easier to obtain.

Fig. 2.14 shows the coverage curves of the experimental data and the PPP and the right shifted curves of the PPP by the average deployment gains, when $\alpha = 4$. As the figure shows, the right shifted curve of the PPP and the curve of the point set are

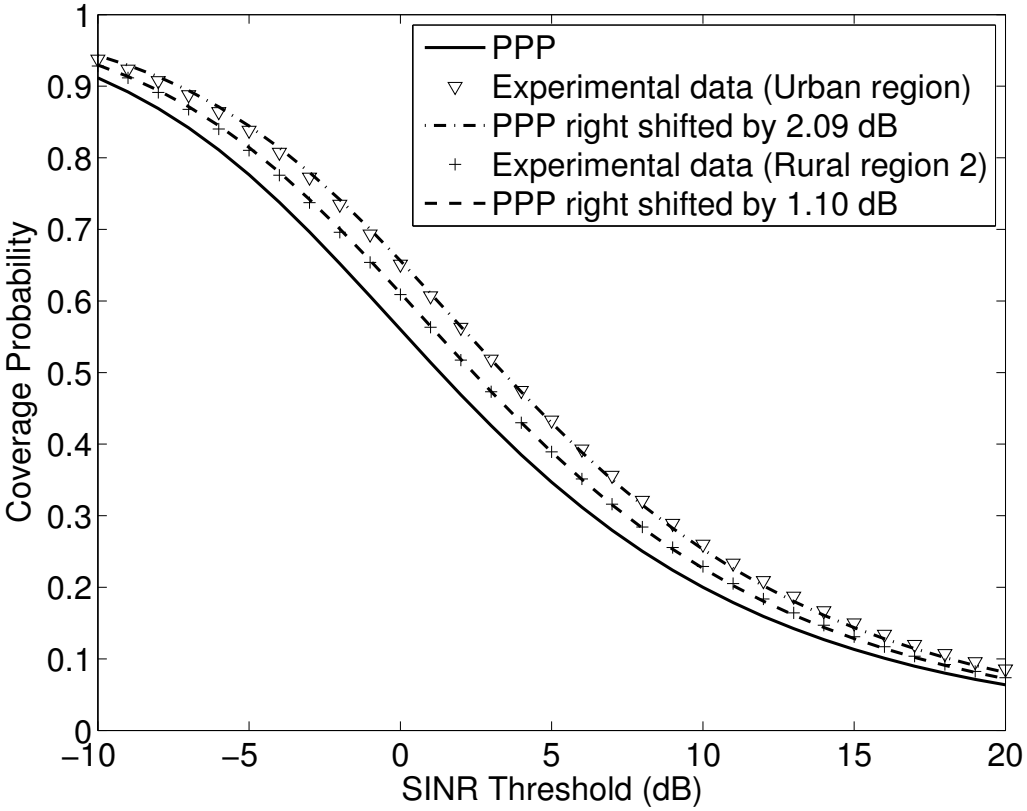

Figure 2.14. The coverage curves of the experimental data and the PPP and the curves of the PPP right shifted by 2.09 dB and 1.10 dB, which are the average deployment gains ($\alpha = 4$). The coverage curve of the experimental data (Rural region 1) and the curve of the PPP right shifted by 1.28 dB are not shown in this figure, but they are well matched, similar to the cases of the other two regions.

well matched. For the point sets of the urban region, the rural region 1 and the rural region 2, the average deployment gains are, respectively, $\hat{S}_{g0} = 2.09$ dB, $\hat{S}_{g1} = 1.28$ dB and $\hat{S}_{g2} = 1.10$ dB. While, the deployment gains at $p_t = 0.5$ are, respectively, $G_0(0.5) = 2.07$ dB, $G_1(0.5) = 1.26$ dB and $G_2(0.5) = 1.08$ dB, which are very close to the average deployment gains. Because $G_0(0.5) > G_1(0.5) > G_2(0.5)$, in terms of the deployment gain, the deployment of the point set of the urban region is the best, followed by the point set of the rural region 1 and then the point set of the rural region 2.

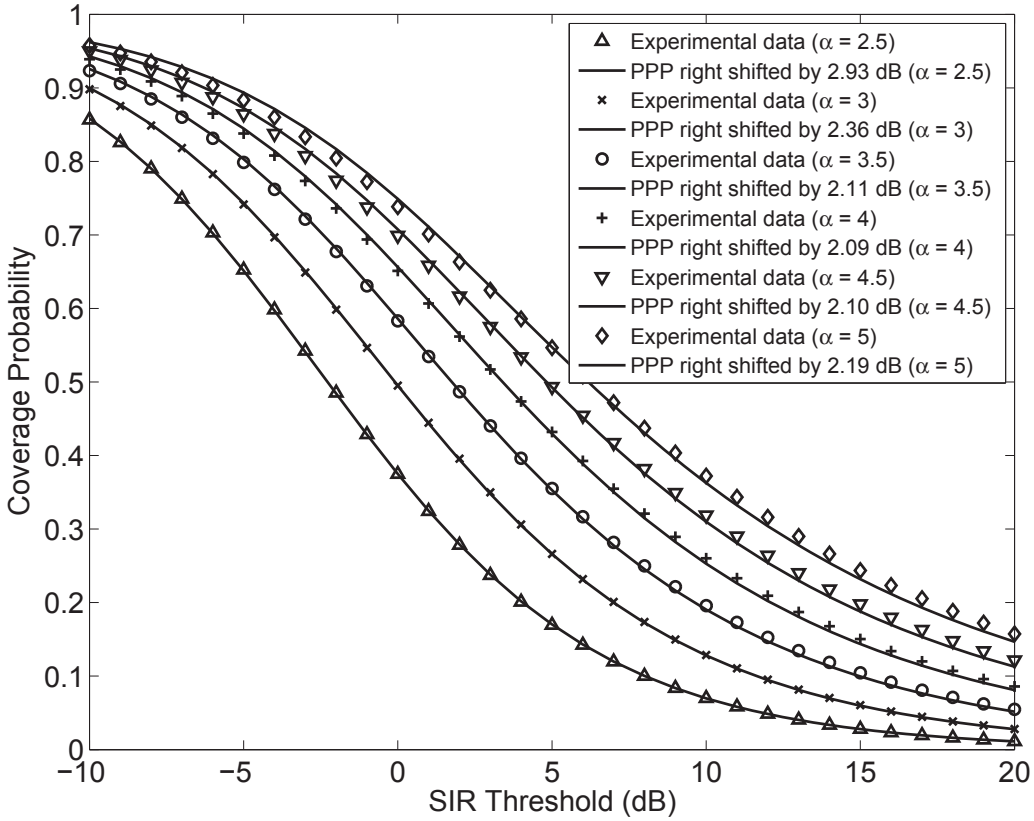

Figure 2.15. The coverage curves of the experimental data (the urban region) and the curves of the PPP right shifted by the corresponding average deployment gains $\hat{S}_g = 2.93, 2.36, 2.11, 2.09, 2.10, 2.19$ (dB) under different values of $\alpha = 2.5, 3, 3.5, 4, 4.5, 5$.

In the above case, the path loss exponent $\alpha = 4$ is fixed. If the value of α varies, $G(0.5)$ and $\hat{S}_g$ will also change. Fig. 2.15 shows the coverage curves of the experimental data (the urban region) and the curves of the PPP right shifted by the corresponding $\hat{S}_g$ under different values of α. For the triangular lattice, as the parameter η of the triangular lattice in the SIR can be eliminated, the coverage probability and the average deployment gain do not depend on η. Fig. 2.16 shows the deployment gains $G(0.5)$ and the average deployment gains $\hat{S}_g$ of all point sets and the triangular lattice when α takes different values, which indicates that $\hat{S}_g$ and $G(0.5)$ are not monotonic as a function of α, but first decrease and then increase as α increases from 2.5 to 5. In this figure, the lines or dashed lines indicate the average deployment gains, and the marks indicate the deployment gains. The inequality $|G(0.5) - \hat{S}_g|/\hat{S}_g < 5\%$ is also satisfied here. The figure also reveals that $G_0(0.5) > G_1(0.5) > G_2(0.5)$ for all $\alpha \in \{2.5, 3, 3.5, 4, 4.5, 5\}$, and the deployment gain of the triangular lattice gives an upper bound.

We have demonstrated that in all cases considered, the coverage probability is very closely approximated by the coverage curve of the PPP, right shifted along the SIR axis by the deployment gain. This general behavior has important implications for the analysis of point process models that are more accurate than the PPP. For the coverage performance evaluation of an arbitrary cellular model, we may take the value analytically obtained for the PPP, and adjust the SIR threshold θ by the deployment gain. Since the coverage probability (or the SIR distribution) affects most first-order metrics, the deployment gain can be used to estimate other metrics (e.g. the average ergodic rate), according to their definitions. Of course, the deployment gain of a model or point set first needs to be established.

One important implication is the estimation of the average ergodic rate of the network using the deployment gain. Since the coverage probability is the CCDF of the SIR and the average ergodic rate only depends on the PDF of the SIR, we can

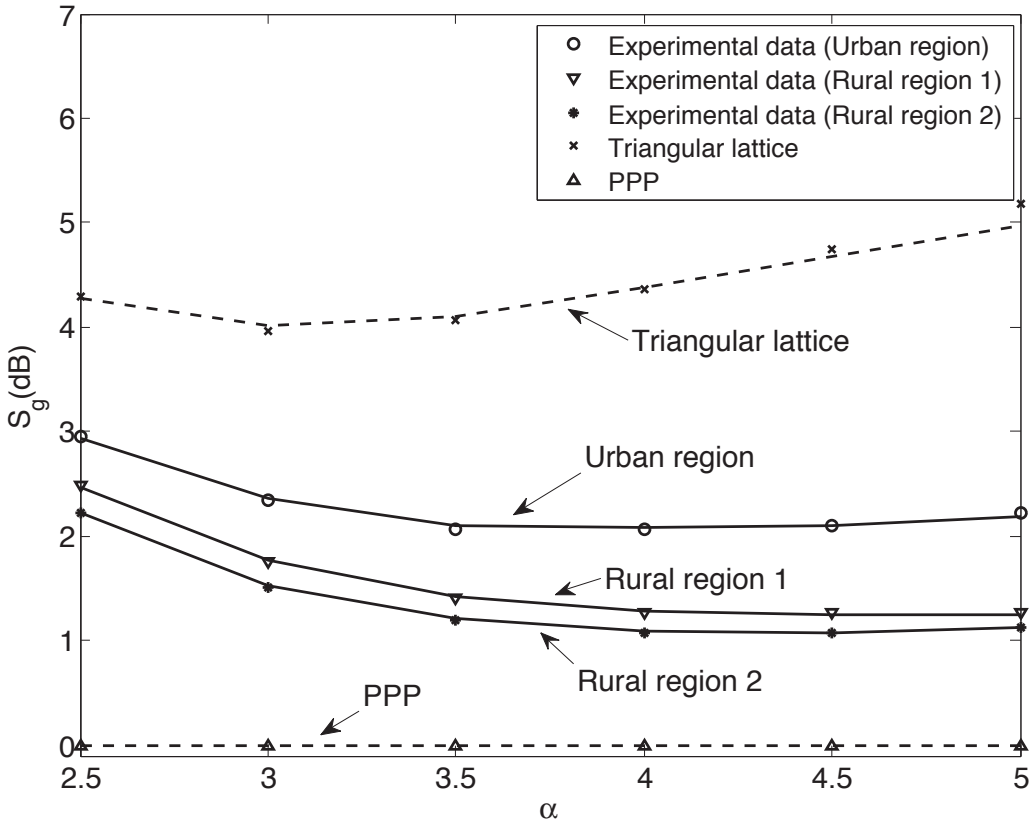

Figure 2.16. The deployment gains $G(0.5)$ and the average deployment gains $\hat{S}_g$ of the experimental data and the triangular lattice when α takes different values. ($G(0.5)$: the marks, $\hat{S}_g$: the lines or dashed lines.)

first obtain the approximated CDF of the SIR by the deployment gain, and then compute the average ergodic rate. Using the definition of the average ergodic rate $\bar{\gamma}$ and the theoretical expression of $P_{\mathrm{c}}(\theta)$ for the PPP in (2.8), $\bar{\gamma}$ can be expressed as

$$\begin{aligned}\bar{\gamma} &= \mathbb{E}[\ln(1+\mathrm{SIR})] \\ &= -\int_0^\infty \ln(1+Gx) dP_{\mathrm{c}}(x) \\ &\overset{(a)}{=} \int_0^\infty P_{\mathrm{c}}\left(\frac{e^x-1}{G}\right) dx, \end{aligned} \tag{2.13}$$

where the unit of G is 1, not dB. (a) follows since the CCDF of the random variable $X = \ln(1+G\cdot\mathrm{SIR})$ is $\mathbb{P}(X > x) = \mathbb{P}\big(\mathrm{SIR} > (\exp(x)-1)/G\big) = P_{\mathrm{c}}\big((\exp(x)-1)/G\big)$ and the expectation of a positive random variable can be expressed as the integral over the CCDF. Numerically evaluating the above integral, we can obtain the approximated average ergodic rate of the three point sets. For the point set of the urban region, $\bar{\gamma} \approx 1.770$. For the point set of the rural region 1, $\bar{\gamma} \approx 1.660$. For the point set of the rural region 2, $\bar{\gamma} \approx 1.6311$. Compared with the values obtained in Section III-E, the difference is smaller than 2%.

2.6 Summary

We proposed a general procedure for point process fitting and applied it to publicly available base station data. To the best of our knowledge, this is the first time public data was used for model fitting in cellular systems. We also introduced the deployment gain, which is a metric on the regularity of a point set or a point process and greatly simplifies the analysis of general point process models.

Two methods are used to fit different point processes to real deployments of BSs in wireless networks in the UK. One is the method of maximum pseudolikelihood, the other is the method of minimum contrast, which minimizes the average squared error of the coverage probability between the point process model and the point set.

Using the former method for fitting, we can decide which model fits best in terms of the coverage probability. But the fitted model cannot perfectly fit the coverage probability—there is still a significant gap between the fit and the data. Conversely, using the latter method, the fitted model fits the data perfectly with respect to the metric.

The deployment gain can be used to evaluate the coverage probability and compare the coverage performances of different point sets analytically. It has considerable practical significance in system design. For example, it can help guide the placement of additional BSs and judge the goodness of a concrete deployment of BSs, which includes recognizing how much better the deployment is than the PPP and how much the deployment could be improved theoretically.

Our work sheds light on real BSs modeling in cellular networks in terms of coverage. For a specified BS data set, we can use the methodology in this chapter to model it. The SP, the PHCP and the perturbed triangular lattice are shown to be accurate models. However, for detailed theoretical analyses, these models may not be suitable. They do not have the tractability of the PPP, since their probability generating functionals are unknown. We can carry out the analysis for the PPP instead and then add the deployment gain to the coverage curve to evaluate the performance of the real deployments.

CHAPTER 3

ASYMPTOTIC DEPLOYMENT GAIN

In this chapter, we propose a novel approach to evaluate the coverage probability of cellular networks, where BSs follow a general class of point processes, using the ADG.

This chapter is organized as follows. In Section 3.1, related works are discussed. In Section 3.2, we introduce the system model and the ADG. We investigate the existence of the ADG and study the asymptotic properties of the outage probability in Section 3.3. Some potential applications of the ADG are provided in Section 3.4. In Section 3.5, we show simulation results for some specific network models. Conclusions are drawn in Section 3.6.

3.1 Related Work

The spatial configuration of the BSs (or transmitters) plays a critical role in the performance evaluation of cellular networks (or general wireless networks), since the SINR critically relies on the distances between BSs and users (or transmitters and receivers). Network performance analysis using stochastic geometry have drawn considerable attention [4, 8, 17, 20, 33–47]. Recent related works can be roughly divided into two categories. One is based on the assumption of modeling the BSs or access points as Poisson-based point processes (e.g., the PPP and the Poisson cluster process) in cellular networks, e.g. [8, 34–37]. The other one is dealing with general point processes in non-cellular networks, especially in wireless ad hoc networks, e.g. [42–46]. Of course, there are some other types of works, such as the type of using the

Poisson-based point processes in non-cellular networks, e.g. [38–40], and the type of using non-Poisson point processes in non-cellular networks, e.g. [17, 41], but they are not related with our concern in this chapter. Our focus is applying general point processes to cellular networks, which has seldom been studied.

Regarding the first category, in cellular networks, the PPP is advantageous for modeling the BSs configuration [8, 34–36], due to its analytical tractability. Poisson-based processes, especially Poisson cluster processes, e.g. in [37], have been used to model the small cell tier in heterogeneous cellular networks, where the BS tier is still modeled as the PPP. Non-Poisson processes, such as hard-core processes, are less studied in cellular networks, due to the absence of an analytical form for the probability generating functional and the Palm characterization of the point process distribution. But recently, there are some works dealing with that topic [20, 47]. In [47], the authors applied the β-Ginibre point process, where points exhibit repulsion, in cellular networks. In [20], the Geyer saturation process was used to model the real cellular service site locations.

As for the second category, general point processes have been used to model the transmitting nodes in non-cellular networks, e.g. [43–46]. In [43], the authors analyzed the success probability in an asymptotic regime where the density of interferers goes to 0 in wireless networks with general fading and node distribution. The paper [44] provided an in-depth study of the outage probability of general ad hoc networks, where the nodes form an arbitrary motion-invariant point process, under Rayleigh fading as the density of interferers goes to 0. In [45], the tail properties of interference for any motion-invariant spatial distribution of transmitting nodes were derived. In [46], dealing with a wide range of point processes, the authors provided accurate approximations of the transmission capacity in the low-outage regime based on the second-order product density of the node distribution in wireless ad hoc networks.

In this chapter, we consider a general class of point processes for modeling possible

BS configurations. In homogeneous cellular networks, each user is usually serviced by its nearest BS, though not necessarily. When general point processes are applied in such networks, one of the main emerging difficulties is that the point process distribution conditioned on an empty ball around the user is unknown. Moreover, the empty space function has to be considered, resulting in the growth of the complexity. Tackling those difficulties directly is seldom seen in the literature.

3.2 System Model and Asymptotic Deployment Gain

3.2.1 System Model

We consider a cellular network that consists of BSs and mobile users. The BSs are modeled as a general m.i. point process Φ of intensity λ on the plane. We assume that Φ is mixing [5, Def. 2.31], which implies that the second moment density $\rho^{(2)}(x_1, x_2) \to \lambda^2$ as $\|x_1 - x_2\| \to \infty$. Intuitively, $\rho^{(2)}(x_1, x_2)$ is the probability that there are two points of Φ at x_1 and x_2 in the infinitesimal volumes dx_1 and dx_2. Rigorously, it is the density (with respect to the Lebesgue measure) pertaining to the second factorial moment measure [5, Def. 6.4], which is given by

$$\alpha^{(2)}(A \times B) = \mathbb{E}\Big(\sum_{x,y \in \Phi}^{\neq} \mathbf{1}_A(x)\mathbf{1}_B(y) \Big) = \int_{A \times B} \rho^{(2)}(y - x) dx dy,$$

where A, B are two compact subsets of $\mathbb{R}^2$, and the $\neq$ symbol indicates that the sum is taken only over distinct point pairs. Since the point processes considered are m.i., $\rho^{(2)}(x_1, x_2)$ only depends on $\|x_1 - x_2\|$. Without ambiguity, we let $\rho^{(2)}(x_2 - x_1) \triangleq \rho^{(2)}(x_1, x_2)$. Similarly, the nth moment density $\rho^{(n)}(x_1, x_2, \ldots, x_n)$ is the the density pertaining to the nth-order factorial moment measure $\alpha^{(n)}$, and we let $\rho^{(n)}(x_2 - x_1, \ldots, x_n - x_1) \triangleq \rho^{(n)}(x_1, \ldots, x_n)$.

We assume all BSs are always transmitting and the transmit power is fixed to

1. Each mobile user receives signals from its nearest BS, and all other BSs act as interferers (the frequency reuse factor is 1). Every signal is assumed to experience path loss and fading. We consider both non-singular and singular path loss models, which are, respectively, $\ell(x) = (1 + \|x\|^{\alpha})^{-1}$ and $\ell(x) = \|x\|^{-\alpha}$, where $\alpha > 2$. (Since $\ell(x)$ only depends on $\|x\|$, in this chapter, $\ell(x)$ and $\ell(\|x\|)$ are equivalent.) We assume that the fading is independent and identically distributed (i.i.d.) for signals from all BSs. The fading can be small-scale fading, shadowing or a combination of the two. We mainly focus on Nakagami-m fading, which includes Rayleigh fading as a special case, and the combination of Nakagami-m fading and log-normal shadowing. The thermal noise is assumed to be additive and constant with power W. We define the mean SNR as the received SNR at a distance of $r = 1$, where its value is $1/(2W)$ for the non-singular path loss model and $1/W$ for the singular path loss model.

To formulate the SINR and the coverage probability, we first define the nearest-point operator NP_{φ} for a point pattern $\varphi \subset \mathbb{R}^2$ as

$$\mathrm{NP}_{\varphi}(x) \triangleq \arg\min_{y \in \varphi} \{\|y - x\|\}, \quad x \in \mathbb{R}^2. \tag{3.1}$$

If the nearest point is not unique, the operator picks one of the nearest points uniformly at random. The SINR at location $z \in \mathbb{R}^2$ has the form

$$\mathrm{SINR}_z = \frac{h_u \ell(u - z)}{W + \sum_{x \in \Phi \setminus \{u\}} h_x \ell(x - z)}, \tag{3.2}$$

where $u = \mathrm{NP}_{\Phi}(z)$ and h_x denotes the i.i.d. fading variable for $x \in \Phi$ with CDF F_h and PDF f_h. For a m.i. point process, the coverage probability $\mathbb{P}(\mathrm{SINR}_z > \theta)$ does not depend on z, and we define

$$P_{\mathrm{c}}(\theta) = \mathbb{P}(\mathrm{SINR} > \theta). \tag{3.3}$$

Hence, without loss of generality, we focus on the coverage probability at the origin o. Since each user communicates with its nearest BS, the interference at o only comes from the BSs outside the open disk $b(o,r) \triangleq \{x \in \mathbb{R}^2 : \|x\| < r\}$, where $r = \|\mathrm{NP}_\Phi(o)\|$. The total interference, denoted by $I(\Phi)$, is

$$I(\Phi) = \sum_{x \in \Phi \setminus \mathrm{NP}_\Phi(o)} h_x \ell(x). \tag{3.4}$$

3.2.2 Asymptotic Deployment Gain

In Section 2.5, we introduced the deployment gain (DG). Here we redefine the DG rigorously.

Definition 11 (Deployment gain)**.** *The deployment gain, denoted by $G(p_\mathrm{t})$, is the ratio of the θ values between the coverage curves of the given point process (or point set) and the PPP at a given target coverage probability p_t, i.e.,*

$$G(p_\mathrm{t}) = \frac{P_\mathrm{c}^{-1}(p_\mathrm{t})}{(P_\mathrm{c}^{\mathrm{PPP}})^{-1}(p_\mathrm{t})} \tag{3.5}$$

where $P_\mathrm{c}^{\mathrm{PPP}}(\theta)$ and $P_\mathrm{c}(\theta)$ are, respectively, the coverage probabilities of the PPP and the given point process Φ.

This definition is analogous to the notion of the coding gain commonly used in coding theory [32, Ch. 1].

Fig. 3.1 shows the coverage probability of the PPP, the Matérn cluster process (MCP) [5, Ch. 3], and the randomly translated triangular lattice. The intensities of all the three point processes are the same. We observe that for $p_\mathrm{t} > 0.3$, the DG is approximately constant, e.g. the DG of the MCP is about -3 dB. In Fig. 3.1, the coverage probability curves of the PPP that are shifted by $G(0.6)$ (in dB) of the MCP and the triangular lattice are also drawn. We see that the shifted curves overlap quite exactly with the curves of the MCP and the triangular lattice, respectively, for all

$p_\mathrm{t} > 0.3$. It is thus sensible to study the DG as $p_\mathrm{t} \to 1$ and find out whether the DG approaches a constant. To do so, analogous to the notion of the asymptotic coding gain, we define the asymptotic deployment gain (ADG).

Definition 12 (Asymptotic deployment gain)**.** *The ADG, denoted by $\hat{G}$, is the deployment gain $G(p_\mathrm{t})$ when $\theta \to 0$, or, equivalently, when $p_\mathrm{t} \to 1$:*

$$\hat{G} = \lim_{p_\mathrm{t} \to 1} G(p_\mathrm{t}). \tag{3.6}$$

Note that, the ADG may not exist for some point processes and fading types. In the following section, we will provide some sufficient conditions for the existence of the ADG. For Rayleigh fading, the ADG of the MCP exists.

Similar to the DG, the ADG measures the coverage probability but characterizes the difference between the coverage probability of the PPP and a given point process as the coverage probability approaches 1 instead of for a target coverage probability, and by observation from Fig. 3.1, the ADG closely approximates the DG for all practical values of the coverage probability. Hence, given the ADG of a point process, we can evaluate its coverage probability by shifting (in dB) the corresponding PPP results, that is to say, $P_\mathrm{c}(\theta) \approx P_\mathrm{c}^\mathrm{PPP}(\theta/\hat{G})$ and $P_\mathrm{c}(\theta) \sim P_\mathrm{c}^\mathrm{PPP}(\theta/\hat{G})$, $\theta \to 0$. In Fig. 3.1, we observe that $\hat{G} \approx 2.4$ for the triangular lattice and $\hat{G} \approx 0.5$ for the MCP. Note that the ADG relative to the PPP permits an immediate calculation of the ADG between two arbitrary point processes.

3.3 Existence of the Asymptotic Deployment Gain

In this section, we derive several important asymptotic properties of the SINR distribution, given some general assumptions about the point process and the CDF of the fading variables. These asymptotic properties, in turn, prove the existence of the ADG.

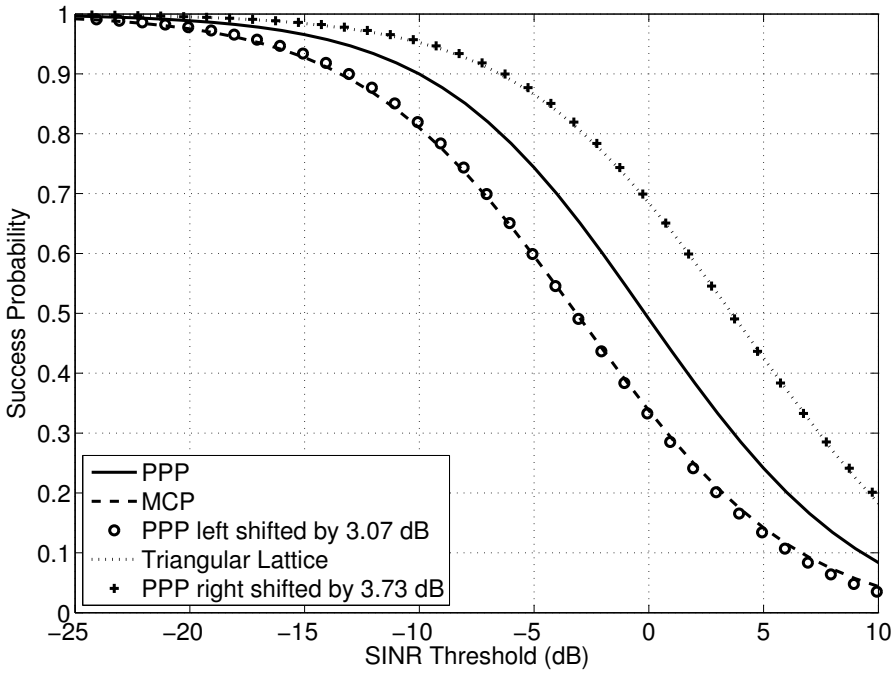

Figure 3.1. The coverage probability of the PPP with intensity $\lambda = 0.1$, the MCP with $\lambda_{\mathrm{p}} = 0.01$, $\bar{c} = 10$ and $r_{\mathrm{c}} = 5$, and the triangular lattice with density $\lambda = 0.1$ (see Section III-B for an explanation of these parameters) for Rayleigh fading, path loss model $\ell(x) = (1 + \|x\|^4)^{-1}$ and no noise, which are simulated on a 100×100 square. The lines are the coverage probability curves of the three point processes, while the markers indicate the coverage probability curves of the PPP shifted by the deployment gains of the MCP and the triangular lattice at $p_{\mathrm{t}} = 0.6$.

First we give several notations, based on which we introduce the precise class of point processes we focus on. We define the contact distance $\xi \triangleq \|\mathrm{NP}_\Phi(o)\|$, and define the supremum of ξ as

$$\xi_{\max} \triangleq \sup_{x\in\mathbb{R}^2} \min_{y\in\Phi}\{\|x-y\|\}. \tag{3.7}$$

Due to the ergodicity of the point process (which follows from the mixing property) [5, Ch. 2], $\xi_{\max}$ does not depend on the realization of Φ. $\xi_{\max} = \infty$ in many mixing point processes.

We define $\Phi_o^\zeta \triangleq (\Phi \mid \mathrm{NP}_\Phi(o) = \zeta)$, where $\zeta \in \mathbb{R}^2 \setminus \{o\}$, as the conditional point process that satisfies $\mathrm{NP}_\Phi(o) = \zeta$, which implies $\zeta \in \Phi_o^\zeta$ and $\Phi_o^\zeta(b(o,\|\zeta\|)) = 0$.[1] So given that ζ is the closest point of Φ to o, the total interference is $I(\Phi_o^\zeta)$. However, it is tricky to directly handle the conditional point process conditioned on that there is an empty disk, if the point process is not the PPP. Thus, we compare the interference in Φ_o^ζ with the interference from a point process where the desired BS ζ is not necessarily the closest one. To this end, we define $\Phi^\zeta \triangleq (\Phi \mid \zeta \in \Phi)$ and consider its interference outside a disk of radius $\|\zeta\|/2$ around the origin:

$$\hat{I}(\Phi^\zeta) = \sum_{x\in\Phi^\zeta \bigcap B_{\zeta/2}\setminus\{\zeta\}} h_x \ell(x), \tag{3.8}$$

where $B_{\zeta/2} \triangleq \mathbb{R}^2 \setminus b(o,\|\zeta\|/2)$. Note that it is not necessary to let the radius of the disk be $\|\zeta\|/2$ and, in fact, the radius could be any quantity that is smaller than $\|\zeta\|$. Since we can use Palm theory [5] for Φ^ζ, it is easier to deal with Φ^ζ than Φ_o^ζ.

To better understand the above notations, we give an illustration of them in Fig. 3.2. Both Φ_o^ζ and Φ^ζ have a point at ζ, and we let $\|\zeta\| = y$. All points of Φ_o^ζ are located in the striped region (outside $b(o,y)$), and $I(\Phi_o^\zeta)$ is the interference from all

[1] For a point process Φ, $\Phi(B)$ is a random variable that denotes the number of points in set $B \subset \mathbb{R}^2$.

these points except ζ. In contrast, Φ^{ζ} may have points throughout the whole plane, but $\hat{I}(\Phi^{\zeta})$ is the interference only from the points of Φ^{ζ} in the shaded region (outside $b(o, y/2)$) except ζ.

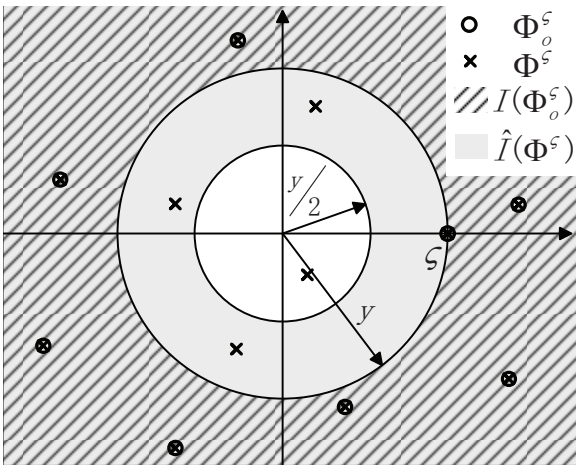

Figure 3.2. An illustration of Φ_o^{ζ}, Φ^{ζ}, $I(\Phi_o^{\zeta})$ and $\hat{I}(\Phi^{\zeta})$, where $\|\zeta\| = y$.

Using the above notations, we define a general class of point process distributions that we use to rigorously state our main result on the SINR distribution.

Definition 13 (Set $\mathcal{A}$)**.** *The set $\mathcal{A} = \{P_{\Phi}\}$ is the set of all m.i. point process distributions P_{Φ} that are mixing and that satisfy the following four conditions. If a point process Φ is distributed as $P_{\Phi} \in \mathcal{A}$,*

1. *for all $n \geq 2$, the nth moment density of Φ is bounded, i.e., $\exists q_n < \infty$, such that $\rho^{(n)}(x_1, \ldots, x_n) < q_n$, for $x_1, \ldots, x_n \in \mathbb{R}^2$;*
2. *for all $y > 0$, $\exists \zeta \in \mathbb{R}^2$ with $\|\zeta\| = y$, such that $\mathbb{P}(\Phi^{\zeta}(b(o, y)) = 0) \neq 0$;*
3. *$\exists y_0 > 0$, such that for all $y > y_0$ and $\zeta \in \mathbb{R}^2$ with $\|\zeta\| = y$, $\hat{I}(\Phi^{\zeta})$ stochastically dominates $I(\Phi_o^{\zeta})$, i.e., $\mathbb{P}(I(\Phi_o^{\zeta}) > z) \leq \mathbb{P}(\hat{I}(\Phi^{\zeta}) > z)$, for all $z \geq 0$;*

4. *$\forall n \in \mathbb{N}$, the n-th moment of the contact distance ξ is bounded, i.e., $\exists b_n < \infty$, s.t. $\mathbb{E}(\xi^n) < b_n$.*

To have a better understanding of the definition, we have the following remarks with respect to each condition:

1. By the mixing property, we have that $\rho^{(2)}(x_1, x_2) \to \lambda^2$, as $\|x_1 - x_2\| \to \infty$, which indicates that when $\|x_1 - x_2\|$ is large enough, $\rho^{(2)}(x_1, x_2)$ is bounded. The first condition means more than that. It guarantees that the nth moment measure of Φ is absolutely continuous with respect to the Lebesgue measure, which, in turn, implies that Φ is *locally finite* [5, Ch. 2.2]. A point process is locally finite if and only if $\Phi(B) < \infty$ a.s., for any $B \subset \mathbb{R}^2$ with $\nu(B) < \infty$, where $\nu(\cdot)$ is the Lebesgue measure. The locally finiteness is a standard assumption in point process theory. Note that, the first condition is stronger than the locally finiteness. This condition excludes some extreme cases, such as the Gauss-Poisson point process as described in [5, Sec. 3.4].

2. Since Φ is a m.i. point process, the second condition is equivalent to requiring that for all $y > 0$, $\forall \zeta \in \mathbb{R}^2$ with $\|\zeta\| = y$, such that $\mathbb{P}(\Phi^{\zeta}(b(o, y)) = 0) \neq 0$. That is to say, if $\zeta \in \Phi$ ($\|\zeta\| \neq 0$), the probability of no points of Φ being located in $b(o, \|\zeta\|)$ is larger than zero. The condition also implies that $\xi_{\max} = \infty$. Because if $\xi_{\max} < \infty$, there surely is at least one point of Φ in $B(o, \xi_{\max})$, which leads to a contradiction that $\forall \zeta \in \mathbb{R}^2$ with $\|\zeta\| > \xi_{\max}$, $\mathbb{P}(\Phi^{\zeta}(b(o, \xi_{\max})) > 0) = 1$. Therefore, the condition excludes the m.i. and mixing point processes where there exists $r_0 > 0$, such that for all $x \in \mathbb{R}^2$, there is at least one point in the region $b(x, r_0)$. Those point processes may be constructed, but are rarely considered in the context of wireless networks.

3. The third condition is based on the two random variables $I(\Phi_o^{\zeta})$ and $\hat{I}(\Phi^{\zeta})$, whose expressions contain the fading variables. But, in fact, the condition is independent of the fading type, since the fading variables are i.i.d. and their expectation is bounded. The condition means that there exists $y_0 > 0$, such that for all $y > y_0$ and $\zeta \in \mathbb{R}^2$ with $\|\zeta\| = y$, the CCDF of the interference from $\Phi^{\zeta} \bigcap B_{\zeta/2} \setminus \{\zeta\}$ is always no smaller than the CCDF of the interference from $\Phi_o^{\zeta} \setminus \{\zeta\}$. Most point processes meet the condition, since an extra region $b(o, y) \setminus b(o, y/2)$ is included in $\hat{I}(\Phi^{\zeta})$, but not in $I(\Phi_o^{\zeta})$. Some point processes, which are seldom considered, may violate the condition. For example, albeit somewhat artificial, in the point process Φ, for small $\epsilon > 0$, the expectation of $\Phi_o^{\zeta}(b(o, \|\zeta\| + \epsilon))$ is much greater than that of $\Phi^{\zeta}(b(o, \|\zeta\| + \epsilon))$, which, at last, leads to the violence of the third condition. Such kind of point processes are beyond our consideration.

4. The fourth condition is satisfied by most point processes that are considered. One sufficient condition of the fourth condition is that $F_{\xi}^{\mathrm{c}}(x) < \exp(-c_0 x)$, as $x \to \infty$, where F_{ξ}^{c} is the CCDF of ξ and $c_0 \in \mathbb{R}^+$. One simple example is the PPP with intensity λ, whose CCDF of ξ is $F_{\xi}^{\mathrm{c}}(x) = \exp(-\lambda \pi x^2)$.

In summary, the four conditions in Def. 13 are quite mild; they are satisfied by most point processes that are usually considered in wireless networks and in stochastic geometry, such as the PPP, the MCP and the Matérn hard-core process (MHP) [5, Ch. 3]. The triangular lattice is not included, since it is not mixing and $\xi_{\max} < \infty$. We will prove that the laws of the PPP, the MCP and the MHP belong to $\mathcal{A}$ in Section 3.3.3.

3.3.1 Main Results

Before presenting the main theorems of this section, we state a property of the distribution of $I(\Phi_o^\zeta)$.

Lemma 1. *Assume the fading variable h satisfies that $\forall n \in \mathbb{N}$, $\mathbb{E}(h^n) < +\infty$. For a point process Φ with $P_\Phi \in \mathcal{A}$, the following statements hold:*

1. *for $\ell(x) = (1 + \|x\|^\alpha)^{-1}$, all moments of the interference $I(\Phi_o^\zeta)$ are bounded, i.e., $\forall n \in \mathbb{N}$, $\exists c_n \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^\zeta)^n) < c_n$, where c_n does not depend on ζ;*
2. *for $\ell(x) = \|x\|^{-\alpha}$, all moments of the interference are bounded, and $\forall n \in \mathbb{N}$, $\exists c_n \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^\zeta)^n) < c_n \max\{1, \|\zeta\|^{2-\alpha n}\}$.*

Proof. See Appendix A.1. □

Since $I(\Phi_o^\zeta)$ can be interpreted as the total interference at o if the nearest base station to o is at ζ, Lemma 1 shows that all moments of the total interference are bounded. If the path loss model is non-singular, the bound can be chosen to be independent of $\|\zeta\|$, the location of the nearest base station. However, if the path loss model is singular, the bound depends on $\|\zeta\|$, and if $\|\zeta\|$ goes to 0, it can be proved that the mean interference becomes arbitrarily large for some BS processes, e.g., the PPP.

Now we are equipped to state our main result: if the CDF of the fading variable h decays polynomially around 0 and all moments of h are bounded, then as a result of

the boundedness of the moments of the interference, the outage probability $1 - P_{\mathrm{c}}(\theta)$ expressed in dB, as a function of the SINR threshold θ, also in dB, has the same slope as $\theta \to 0$, for all Φ with $P_\Phi \in \mathcal{A}$.

Theorem 1. *For a point process Φ with $P_\Phi \in \mathcal{A}$, if the fading variable satisfies*

1. $\exists m \in (0, +\infty)$, s.t. $F_h(t) \sim at^m$, as $t \to 0$, where $a > 0$ is constant,

2. $\forall n \in \mathbb{N}$, $\mathbb{E}(h^n) < +\infty$,

then we have

$$\frac{1 - P_{\mathrm{c}}(\theta)}{\theta^m} \to \kappa, \qquad \text{as } \theta \to 0, \tag{3.9}$$

where $0 < \kappa < \infty$ does not depend on θ and is given by

$$\kappa = \int_0^\infty \mathbb{E}_{I(\Phi_o^\zeta)} \Big[a\ell(y)^{-m} \big(I(\Phi_o^\zeta) + W \big)^m \Big] f_\xi(y) dy \tag{3.10}$$

($\|\zeta\| = y$) and f_ξ is the PDF of ξ.

Proof. See Appendix A.2. □

Theorem 1 shows that the ADG exists and how it depends on the other network parameters. The following theorem quantifies the ADG.

Corollary 1. *Under the same condition as in Theorem 1, the ADG of Φ exists and is given by*

$$\hat{G} = \Big(\frac{\kappa^{\mathrm{PPP}}}{\kappa} \Big)^{\frac{1}{m}}, \tag{3.11}$$

where κ^{PPP} is the value for the PPP and κ is the value for Φ. For the PPP with intensity λ,

$$\kappa^{\mathrm{PPP}} = 2\lambda\pi \int_0^\infty \mathbb{E}_{I_r} \Big[\frac{m^{m-1}}{\Gamma(m)} \ell(r)^{-m} \big(I_r + W \big)^m \Big] r \exp(-\lambda\pi r^2) dr, \tag{3.12}$$

where $I_r = \sum_{x \in \Phi \bigcap b(o,r)^{\mathrm{c}}} h_x \ell(x)$.

Proof. Given a target coverage probability p_{t}, define $\theta_1 \triangleq P_{\mathrm{c}}^{-1}(p_{\mathrm{t}})$ and $\theta_2 \triangleq (P_{\mathrm{c}}^{\mathrm{PPP}})^{-1}(p_{\mathrm{t}})$. As $p_{\mathrm{t}} \to 1$, we have $\theta_1 \to 0$ and $\theta_2 \to 0$. By Theorem 1, $1 - P_{\mathrm{c}}(\theta_1) \sim \kappa\theta_1^m$ and $1 - P_{\mathrm{c}}^{\mathrm{PPP}}(\theta_2) \sim \kappa^{\mathrm{PPP}}\theta_2^m$. Since $p_{\mathrm{t}} = P_{\mathrm{c}}(\theta_1) = P_{\mathrm{c}}^{\mathrm{PPP}}(\theta_2)$, as $p_{\mathrm{t}} \to 1$, $\kappa\theta_1^m = \kappa^{\mathrm{PPP}}\theta_2^m$. Thus, $\hat{G} = \lim_{p_{\mathrm{t}}\to 1} \theta_1/\theta_2 = (\kappa^{\mathrm{PPP}}/\kappa)^{1/m}$. □

Note that Rayleigh fading meets the requirements in Theorem 1 with $m = 1$. For the special case of the PPP with intensity λ, no noise and Rayleigh fading, it has been obtained in [8] that

$$P_{\mathrm{c}}(\theta) = \left(1 + \theta^{\delta} \int_{\theta^{-\delta}}^{\infty} \frac{1}{1 + u^{1/\delta}} du\right)^{-1}, \tag{3.13}$$

where $\delta \triangleq 2/\alpha$. It follows that $\kappa^{\mathrm{PPP}} = \lim_{\theta\to 0} \frac{1-P_{\mathrm{c}}(\theta)}{\theta} = \frac{2}{\alpha-2}$. For $\alpha = 4$, $P_{\mathrm{c}}(\theta) = 1/(1 + \sqrt{\theta}\arctan\sqrt{\theta})$, and $\kappa^{\mathrm{PPP}} = 1$.

A point process has different ADGs depending on the value of m. So it is sensible to compare the ADGs of different point process models only under the same fading assumption.

We have proved that the ADG exists with certain constrains on the fading and point processes. In the rest of this section, we consider some special cases for the point processes and fading distributions.

3.3.2 Special Cases - Fading Types

Regarding the fading, we mainly consider Nakagami-m fading and *composite fading*, which is a combination of Nakagami-m fading and log-normal shadowing.

3.3.2.1 Nakagami-m Fading

The fading variable $h \sim \mathrm{gamma}(m, 1/m)$. On the one hand, we have

$$\lim_{t\to 0} \frac{F_h(t)}{t^m} = \lim_{t\to 0} \frac{(mt)^{m-1}\exp(-mt)}{\Gamma(m)t^{m-1}} = \frac{m^{m-1}}{\Gamma(m)} < +\infty. \tag{3.14}$$

On the other hand, since $F_h^{\mathrm{c}}(x)$ has an exponential tail, all moments of h are finite. Thus, Nakagami-m fading meets the requirements in Theorem 1.

In addition, we find an interesting phenomenon that for a point process Φ with $P_\Phi \in \mathcal{A}$, the behavior of the CCDF of the fading at the tail determines the behavior of the CCDF of the interference $I(\Phi_o^\zeta)$ at the tail. The following corollary presents the property. As usual, $f(x) = \Omega(g(x))$, as $x \to \infty$ means $\limsup_{x\to\infty} \left|\frac{f(x)}{g(x)}\right| > 0$.

Corollary 2. *For a point process Φ with $P_\Phi \in \mathcal{A}$, if the fading has at most an exponential tail, i.e., $-\log F_h^{\mathrm{c}}(x) = \Omega(x)$, $x \to \infty$, where $F_h^{\mathrm{c}}(x)$ is the CCDF of the fading variable h, then the interference tail is bounded by an exponential, i.e., $-\log F_{I(\Phi_o^\zeta)}^{\mathrm{c}}(x) = \Omega(x)$, $x \to \infty$, where $F_{I(\Phi_o^\zeta)}^{\mathrm{c}}(x)$ is the CCDF of $I(\Phi_o^\zeta)$.*

Proof. See Appendix A.3. □

A similar property has been derived in [45], namely, that in ad hoc networks modeled by m.i. point processes, an exponential tail in the fading distribution implies an exponential tail in the interference distribution. The result cannot be directly applied to cellular networks, because in the cellular network that we consider, each user communicates with its nearest BS u and thus no interferers can be closer than u, while the authors in [45] assume the receiver communicates with a transmitter with a fixed location and there can be some interferers closer to the receiver than the transmitter.

3.3.2.2 Composite Fading

The signals from all BSs experience both Nakagami-m fading and log-normal shadowing. A similar kind of fading has been investigated in [48, 49], where the fading was composed of Rayleigh fading and log-normal shadowing. Denote the fading variable with respect to Nakagami-m fading by $\tilde{h}$ and the fading variable

with respect to log-normal shadowing by $\hat{h}$. The composite fading variable can be represented as $h = \tilde{h}\hat{h}$, where $\tilde{h}$ and $\hat{h}$ are independent.

For the log-normal shadowing, we use the definition from [53]. Without loss of generality, we assume $\hat{h} = 10^{X/10}$, where $X \sim N(0, \sigma^2)$. The CDF of $\hat{h}$, denoted by $F_{\hat{h}}(t)$, is

$$F_{\hat{h}}(t) = \frac{1}{2}\mathrm{erfc}\bigg(- \frac{10\log t}{\sigma\sqrt{2}\log 10}\bigg) = \frac{1}{\sqrt{\pi}}\int_{-\frac{10\log t}{\sigma\sqrt{2}\log 10}}^{\infty} \exp(-v^2)dv, \tag{3.15}$$

where erfc is the complementary error function. It is straightforward to obtain that $\mathbb{E}[\hat{h}] = \exp((\frac{\log 10}{10})^2\frac{\sigma^2}{2})$ and $\mathbb{E}[\hat{h}^2] = \exp((\frac{\log 10}{10})^2 2\sigma^2)$.[2] We can easily show that as $t \to \infty$, $F^{\mathrm{c}}_{\hat{h}}(t)$ decays faster than t^{-n} for any $n \in \mathbb{N}$, but slower than $\exp(-at)$ for any $a > 0$.

For the composite fading, we have the following lemma about the distribution of h.

Lemma 2. *If $\tilde{h} \sim$ gamma$(m, \frac{1}{m})$, $10\log\hat{h}/\log 10 \sim N(0, \sigma^2)$, and $\tilde{h}$ is independent of $\hat{h}$, the distribution of $h = \tilde{h}\hat{h}$ has the following properties:*

1. *F_h decays polynomially around 0 and*

$$\lim_{t\to 0}\frac{F_h(t)}{t^m} = \int_0^{\infty} \frac{10m^{m-1}}{\sigma\log 10\sqrt{2\pi}\Gamma(m)u^{m+1}} \exp\bigg(-\Big(\frac{10\log u}{\sigma\sqrt{2}\log 10}\Big)^2\bigg)du < \infty; \tag{3.16}$$

2. *$F^{\mathrm{c}}_h(t) = o(t^{-n})$, as $t \to \infty$, for any $n \in \mathbb{N}$, and $-\log F^{\mathrm{c}}_h(t) = o(t)$, $t \to \infty$.*

Proof. See Appendix A.4. □

The two properties in Lemma 2 indicate that the composite fading retains the asymptotic property of Nakagami-m fading for $t \to 0$ and that of log-normal shad-

[2]Note that the mean of $\hat{h}$ is not 1. Actually, we can normalize it to 1 and replace it with the normalized variable in our results, but since it does not affect our results, for convenience, we just leave it as it is.

owing for $t \to \infty$, respectively. They also imply that the composite fading meets the requirements in Theorem 1.

Regarding the distribution of the interference at the tail, we also have the following corollary.

Corollary 3. *For a point process Φ with $P_\Phi \in \mathcal{A}$ and composite fading, the interference tail is upper bounded by a power law with arbitrary parameter β, i.e., $F^{\mathrm{c}}_{I(\Phi^\zeta_\delta)}(y) = o(y^{-\beta})$, $\forall \beta \in \mathbb{N}$, as $y \to +\infty$.*

Proof. We can simply apply the Markov inequality and have that $\forall \beta \in \mathbb{N}$,

$$\mathbb{P}(I(\Phi^\zeta_o) > y) \leq \frac{\mathbb{E}(I(\Phi^\zeta_o)^\beta)}{y^\beta}. \tag{3.17}$$

Hence, using Lemma 1, we have $F^{\mathrm{c}}_{I(\Phi^\zeta_o)}(y) = o(y^{-\beta})$, $\forall \beta \in \mathbb{N}$, as $y \to +\infty$. □

3.3.3 Special Cases - Point Processes

As for the point processes, we specifically concentrate on the PPP, the MCP and the MHP. We first give an introduction to the three point processes, then we will present the asymptotic properties with respect to the fading types.

Poisson Point Process: The PPP is the simplest model of point processes, which exhibits complete spatial randomness. The points in the PPP are stochastically independent, which makes the PPP the most tractable point process.

Matérn Cluster Process: As a class of clustered point processes on the plane built on a PPP, the MCPs are doubly Poisson cluster processes, where the parent points form a uniform PPP Φ_{p} of intensity λ_{p} and the daughter points are uniformly scattered on the ball of radius r_{c} centered at each parent point x_{p} with intensity

$$\lambda_0(x) = \frac{\bar{c}}{\pi r_{\mathrm{c}}^2} \mathbf{1}_{B(x_{\mathrm{p}}, r_{\mathrm{c}})}(x), \tag{3.18}$$

where $B(x_{\rm p}, r_{\rm c}) \triangleq \{x \in \mathbb{R}^2 : \|x - x_{\rm p}\| \leq r_{\rm c}\}$ is the closed disk of radius $r_{\rm c}$ centered at $x_{\rm p}$. The mean number of daughter points in one cluster is $\bar{c}$. So the intensity of the process is $\lambda = \lambda_{\rm p}\bar{c}$.

Matérn Hard-core Process: The MHPs are a class of repulsive point processes, where points are forbidden to be closer than a certain minimum distance. There are several types of MHPs. Here we only consider the MHP of type II [5, Ch. 3], which is generated by starting with a basic uniform PPP $\Phi_{\rm b}$ of intensity $\lambda_{\rm b}$, adding to each point x an independent random variable $m(x)$, called a mark, uniformly distributed on $[0, 1]$, then flagging for removal all points that have a neighbor within distance $r_{\rm h}$ that has a smaller mark and finally removing all flagged points. The intensity of the MHP is $\lambda = \frac{1-\exp(-\lambda_{\rm b}\pi r_{\rm h}^2)}{\pi r_{\rm h}^2}$. The highest density $\lambda_{\max} = 1/(\pi r_{\rm h}^2)$ is achieved as $\lambda_{\rm b} \to \infty$.

Lemma 3. *The distributions of the PPP, the MCP and the MHP belong to the set* $\mathcal{A}$.

Proof. See Appendix A.5. □

By Lemma 3, regarding Nakagami-m fading and composite fading, we have the following corollary directly from Theorem 1.

Corollary 4. *If the fading is Nakagami-m or the composite fading, then for the PPP, the MCP and the MHP,*

$$\frac{1 - P_{\rm c}(\theta)}{\theta^m} \to \kappa, \qquad as\ \theta \to 0, \tag{3.19}$$

where $\kappa > 0$ *is given by*

$$\kappa = \int_0^\infty \mathbb{E}_{I(\Phi_o^\zeta)}\Big[a\ell(y)^{-m}\big(I(\Phi_o^\zeta) + W\big)^m\Big] f_\xi(y)dy \tag{3.20}$$

($\|\zeta\| = y$) *and* f_ξ *is the PDF of* ξ. *For Nakagami-m fading,* $a = \frac{m^{m-1}}{\Gamma(m)}$; *for the composite fading,* $a = \int_0^\infty \frac{10m^{m-1}}{\sigma \log 10 \sqrt{2\pi}\Gamma(m)u^{m+1}} \exp(-(\frac{10 \log u}{\sigma\sqrt{2}\log 10})^2)du$.

3.4 Applications of the Asymptotic Deployment Gain

Since the ADG characterizes the gap of the coverage probability between a point process and the PPP, any statistic that depends on the distribution of the SINR (e.g., the average ergodic rate and the mean SINR) can be approximated using the ADG. In this section, we focus on the average ergodic rate and the mean SINR.

3.4.1 Average Ergodic Rate

We assume each mobile user adopts adaptive modulation/coding to achieve the Shannon bound of the rate for the instantaneous SINR. That is to say, each user adjusts its rate of transmission to $\gamma = \ln(1 + \mathrm{SINR})$. Using the same definition as in Section 2.5, the average ergodic rate (expressed in nats) is $\bar{\gamma} \triangleq \mathbb{E}[\ln(1 + \mathrm{SINR})]$.

Provided the ADG of Φ as $\hat{G}$ and the coverage probability of the corresponding PPP as $P_{\mathrm{c}}^{\mathrm{PPP}}(\theta)$, the coverage probability of Φ is approximated as $P_{\mathrm{c}}^{\mathrm{PPP}}(\theta/\hat{G})$. Similar to (2.13), the average ergodic rate can be expressed as

$$\begin{aligned}
\bar{\gamma} &\approx -\int_0^\infty \ln(1+\theta) dP_{\mathrm{c}}^{\mathrm{PPP}}\left(\frac{\theta}{\hat{G}}\right) \\
&= -\int_0^\infty \ln(1+\hat{G}\theta) dP_{\mathrm{c}}^{\mathrm{PPP}}(\theta) \qquad (3.21) \\
&= \int_0^\infty P_{\mathrm{c}}^{\mathrm{PPP}}\left(\frac{e^x - 1}{\hat{G}}\right) dx. \qquad (3.22)
\end{aligned}$$

In general, if each user adopts a certain type of modulation or coding instead of adaptive modulation/coding, the rate can be approximated by adding a gap G_{g}, i.e., $\gamma \approx \ln(1 + \mathrm{SINR}/G_{\mathrm{g}})$, where $G_{\mathrm{g}} \geq 1$. From (3.21), we observe that in fact, the distribution of base stations has the same effect on the average ergodic rate as the

modulation and coding. The modulation and coding degrade the level of the average ergodic rate, but the distribution of base stations can either increase or decrease the average ergodic rate, depending on whether the base stations are regular ($\hat{G} > 1$) or clustered ($\hat{G} < 1$). If $\hat{G} = 1$, the average ergodic rate approximates that of the PPP, and as $\hat{G}$ grows, the average ergodic rate increases.

3.4.2 Mean SINR

Just as the coverage probability and the average ergodic rate, the mean SINR is also an important criterion that has been discussed in wireless networks, e.g. in [50]. Denote M_Φ as the mean SINR for Φ, and M_{PPP} the mean SINR for the PPP with the same intensity as that of Φ. It can be proved that the mean SINR for the PPP is infinite if the path loss model is singular. Briefly, for $\zeta = \mathrm{NP}_\Phi(o)$, letting $y = \|\zeta\|$, we have

$$\begin{aligned}
\mathbb{E}(\mathrm{SINR}) &= \mathbb{E}\bigg(\frac{\ell(\zeta)}{W + I(\Phi_o^\zeta)}\bigg) \overset{(a)}{\geq} \mathbb{E}_y\bigg(\frac{\ell(\zeta)}{W + \mathbb{E}[I(\Phi_o^\zeta)]}\bigg) \overset{(b)}{\geq} \mathbb{E}_y\bigg(\frac{y^{-\alpha}}{W + c_1 \max\{1, y^{2-\alpha}\}}\bigg) \\
&= \int_0^1 \frac{x^{-\alpha}}{W + c_1 x^{2-\alpha}} f_{\|\zeta\|}(x)dx + \int_1^\infty \frac{x^{-\alpha}}{W + c_1} f_{\|\zeta\|}(x)dx \\
&\geq \int_0^1 \frac{x^{-1}}{W + c_1} 2\pi\lambda e^{-\lambda\pi x^2} dx + \int_1^\infty \frac{x^{-\alpha}}{W + c_1} f_{\|\zeta\|}(x)dx = \infty, \qquad (3.23)
\end{aligned}$$

where $f_{\|\zeta\|}(x) = 2\pi\lambda x e^{-\lambda\pi x^2}$ is the contact distance distribution for the PPP, (a) follows from Jensen's inequality, and (b) follows from Lemma 1.

So, we only consider the non-singular path loss model. We have $\mathbb{E}(\mathrm{SINR}) = \mathbb{E}(h)\mathbb{E}\big(\frac{\ell(\zeta)}{W+I(\Phi_o^\zeta)}\big) \leq \frac{\mathbb{E}(h)}{W}\mathbb{E}(\ell(\zeta)) < \infty$, where $\zeta = \mathrm{NP}_\Phi(o)$. Given the ADG $\hat{G}$ of Φ, we have a simple approximation for M_Φ:

$$M_\Phi \approx \hat{G} M_{\mathrm{PPP}}.$$

Therefore, the ADG can also be interpreted as the approximate gain in the mean

SINR.

Conditioned on a small SINR, the relationship becomes exact asymptotically: Define $M_\Phi^t = \mathbb{E}(\text{SINR} \mid \text{SINR} < t)$. Since $M_\Phi^t = \int_0^t P_\text{c}(\theta)d\theta - tP_\text{c}(t)$, we have $M_\Phi^t \sim \hat{G} M_\text{PPP}^t$, as $t \to 0$, where M_PPP^t is M_Φ^t when Φ is the PPP.

3.5 Simulations

In this section, we perform simulations on a 100×100 square, where we consider the non-singular path loss model and fix the path loss exponent to $\alpha = 4$ and the intensity of the point processes to $\lambda = 0.1$. For the MCP, we let $\lambda_\text{p} = 0.01$, $\bar{c} = 10$ and $r_\text{c} = 5$; for the MHP, we let $\lambda_\text{b} = 0.263$ and $r_\text{h} = 1.7$. We present our results in two subsections corresponding to the SINR distribution and the applications of the ADG.

3.5.1 SINR Distribution

3.5.1.1 Nakagami-m Fading

In this part, we present simulation results of the outage probability for the PPP, the MCP, and the MHP under Nakagami-m fading.

Fig. 3.3 shows the outage curves $1 - P_\text{c}(\theta)$ of the PPP for $m \in \{1, 2\}$ and different mean SNR values. Note that the SNR value here is the mean SNR measured at a unit distance from the base station, and is $1/(2W)$. As θ approaches 0, the slopes of the curves for $m = 1$ are all 10 dB/decade, and the slopes for $m = 2$ are all 20 dB/decade, in agreement with Corollary 4. We also observe that there is only a rather small gap between the cases of SNR = 20 dB and SNR = ∞, thus the thermal noise does not significantly affect the asymptotic performance of the coverage probability. We will neglect noise in the rest of this section.

In Fig. 3.4, we find that for the same point process, a different m implies a

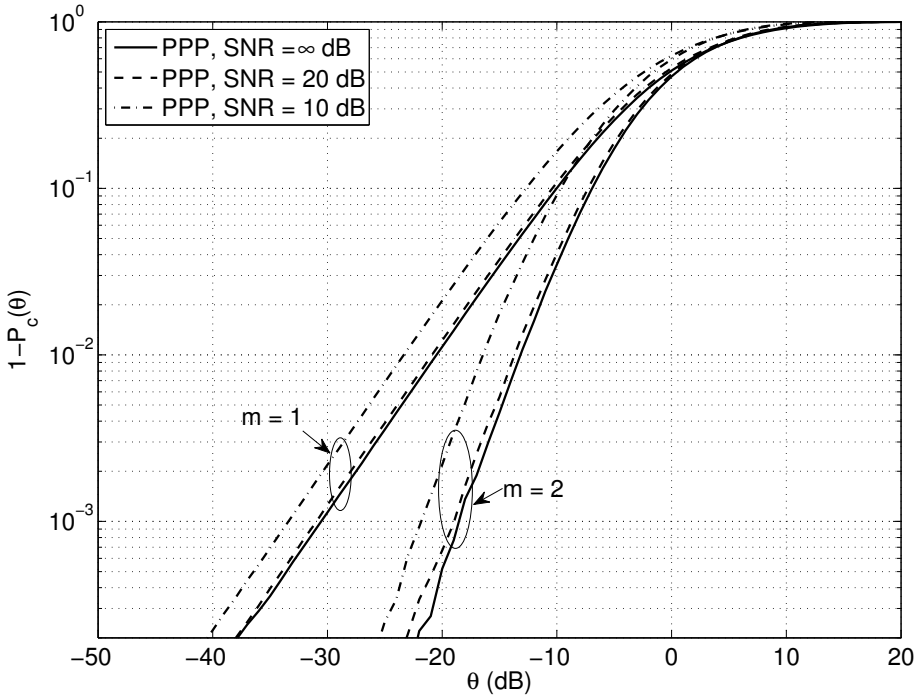

Figure 3.3. Nakagami-m fading: the outage probability $1 - P_{\rm c}(\theta)$ vs. θ for the PPP when $m \in \{1, 2\}$ under different SNR settings.

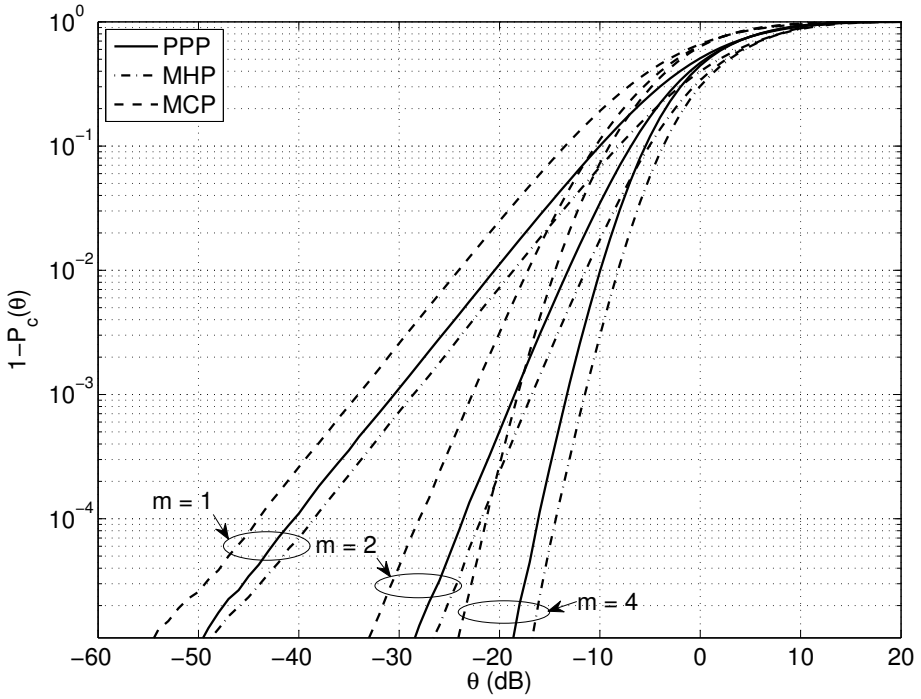

Figure 3.4. Nakagami-m fading: the outage probability $1 - P_{\rm c}(\theta)$ vs. θ for the PPP, the MCP and the MHP when $m \in \{1, 2, 4\}$ (no noise).

different asymptotic slope. In fact, the slope is $10m$ dB/decade, just as Corollary 4 indicates. For the same m, different point processes have the same asymptotic slope, thus in the high-reliability regime, the coverage probability of a non-Poisson process can be obtained accurately simply by shifting the coverage probability curve of the PPP with the same intensity by the ADG. Besides, we observe that for any m, the coverage probability of the MHP is the largest of the three processes, followed by the PPP and then the MCP. Intuitively, the MHP has a better coverage probability because it is more regular than the PPP. Similarly, the MCP has a poorer coverage probability because it is more clustered than the PPP. In addition, since the value of κ for the MCP and the MHP can be approximated through the simulation, by Corollary 1, we can approximate their ADGs. Denote by $\hat{G}_m^{\mathrm{MCP}}$ the ADG for the MCP with respect to m, and by $\hat{G}_m^{\mathrm{MHP}}$ that of the MHP. We obtain that for the MCP, $\hat{G}_1^{\mathrm{MCP}} \approx 0.49$, $\hat{G}_2^{\mathrm{MCP}} \approx 0.37$ and $\hat{G}_4^{\mathrm{MCP}} \approx 0.29$; for the MHP, $\hat{G}_1^{\mathrm{MHP}} \approx 1.58$, $\hat{G}_2^{\mathrm{MHP}} \approx 1.48$ and $\hat{G}_4^{\mathrm{MHP}} \approx 1.41$. Note that $\hat{G}_1^{\mathrm{MCP}}$ is consistent with the approximated value 0.49 obtained from Fig. 3.1.

3.5.1.2 Composite Fading

We consider the combination of Nakagami-m fading and log-normal shadowing in this part. Fig. 3.5 shows the outage probability for the PPP when $\sigma = 2$ and $m \in \{1, 2, 4\}$. As Corollary 4 indicates, the case of the composite fading has the same asymptotic property as the case of Nakagami-m fading. As $\theta \to 0$, the slope of the curve is $10m$ dB/decade.

In Fig. 3.6, the outage probabilities for the PPP, the MCP and the MHP are exhibited. The MHP still has the best outage probability, followed by the PPP and the MCP. We also observe that the value of σ does not affect the slope of the outage curve as $\theta \to 0$, which is $10m$ dB/decade. The ADGs of the MCP and MHP can also be estimated: for $m = 1$ and $\sigma = 2$, $\hat{G}_1^{\mathrm{MCP}} \approx 0.51$ and $\hat{G}_1^{\mathrm{MHP}} \approx 1.55$; for $m = 2$ and

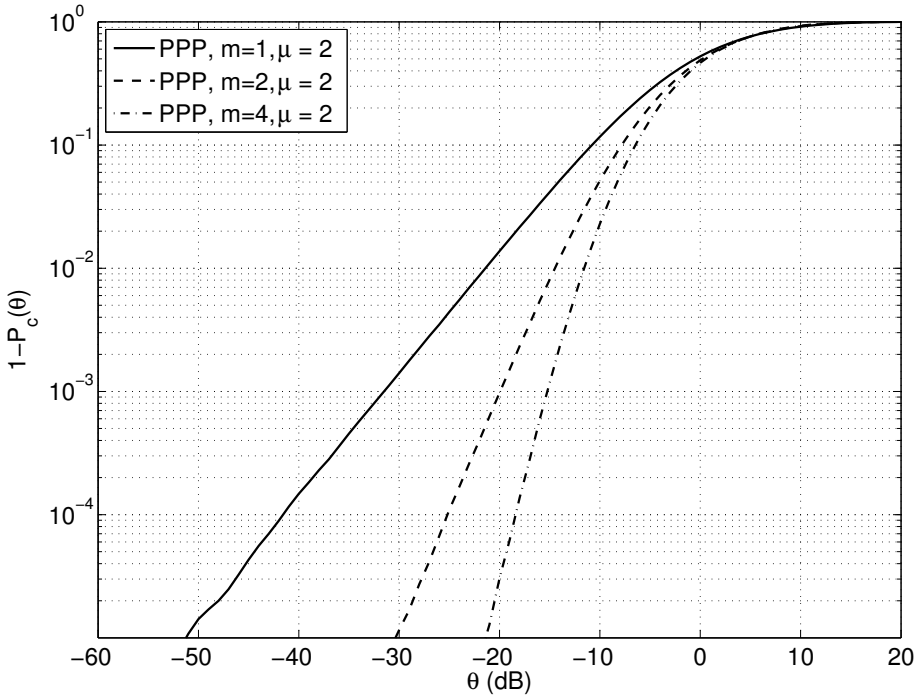

Figure 3.5. Compound fading: the outage probability $1 - P_{\rm c}(\theta)$ vs. θ for the PPP when $m \in \{1, 2, 4\}$ and $\sigma = 2$ (no noise, $\alpha = 4$).

$\sigma = 4$, $\hat{G}_2^{\mathrm{MCP}} \approx 0.40$ and $\hat{G}_2^{\mathrm{MHP}} \approx 1.37$.

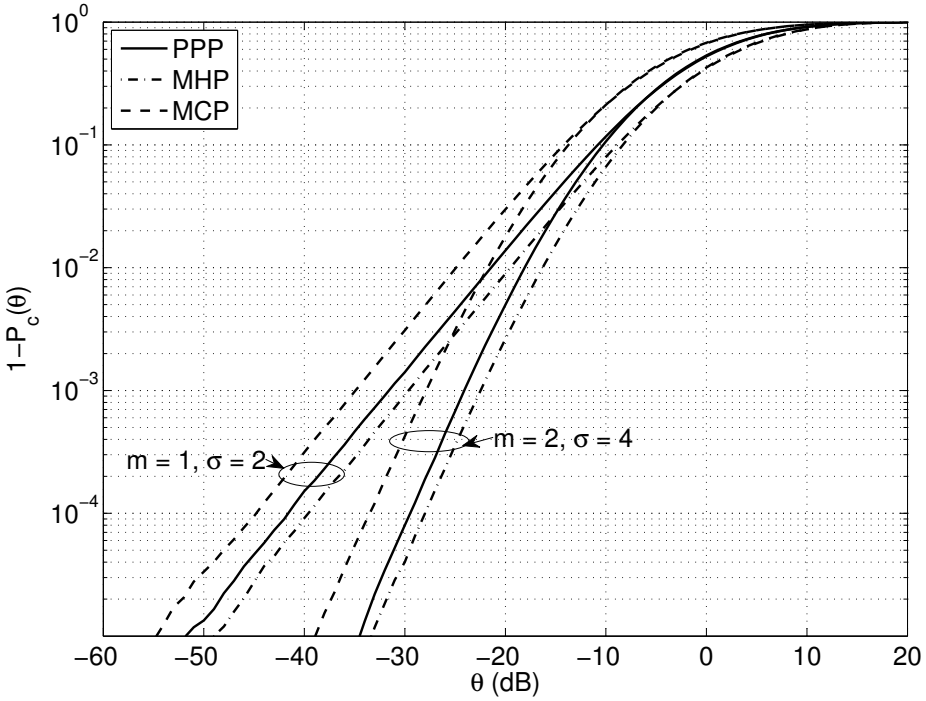

Figure 3.6. Compound fading: the outage probability $1 - P_{\mathrm{c}}(\theta)$ vs. θ for the PPP, the MCP and the MHP when $m = 1, \sigma = 2$ and $m = 2, \sigma = 4$ (no noise, $\alpha = 4$).

3.5.2 Applications of the ADG

In this subsection, we evaluate the average ergodic rate and the mean SINR for the PPP, the MCP and the MHP through simulations, and also estimate them using the ADGs. The ADG values approximated by the DG values at $p_{\mathrm{t}} = 1 - 10^{-4}$ for the three point processes are presented in Table 3.1. We see that as a function of α,

the ADG does not monotonically increase as α increases.

TABLE 3.1

THE ADGS FOR DIFFERENT α (RAYLEIGH FADING, NO NOISE)

ADG	$\alpha = 2.5$	$\alpha = 3.0$	$\alpha = 3.5$	$\alpha = 4$	$\alpha = 4.5$	$\alpha = 5$
MCP	0.46	0.40	0.41	0.49	0.42	0.46
MHP	1.27	1.37	1.37	1.58	1.40	1.65

3.5.2.1 Average Ergodic Rate

In Fig. 3.7, the average ergodic rates $\bar{\gamma}$ for the three point processes as a function of α are shown as the lines. We also use the simulation results of the PPP and the ADGs in Table 3.1 to estimate the average ergodic rates for the MCP and the MHP. The estimated values are shown as the markers in Fig. 3.7. From the figure, we verify that the average ergodic rates estimated using the ADGs provide fairly good approximations to the empirical values. We also observe that $\bar{\gamma}$ increases as α grows, which is obvious since the interference decays much faster than the desired signal power.

3.5.2.2 Mean SINR

In Fig. 3.8, the lines are the mean SINRs for the three point processes as a function of α. The markers indicate the mean SINRs for the MCP and the MHP estimated

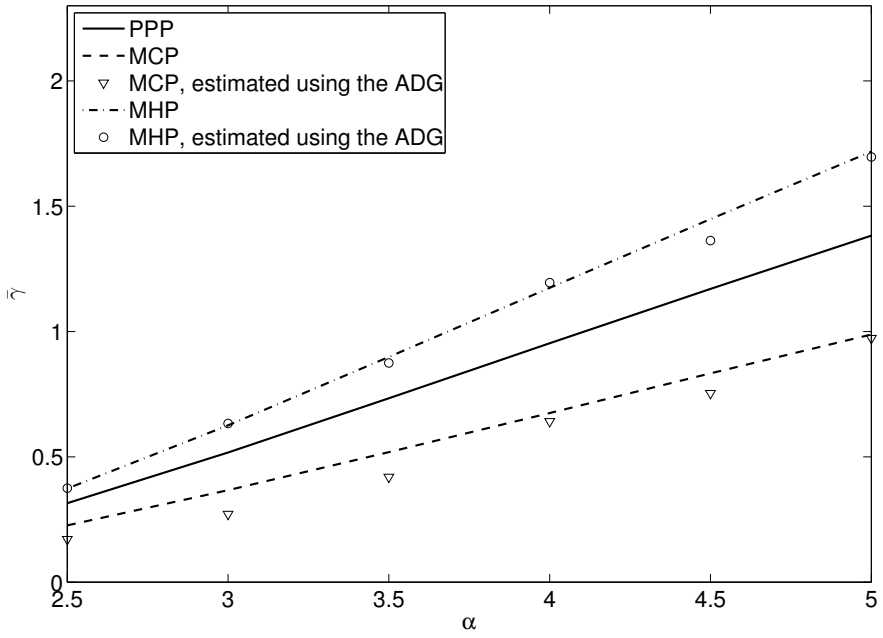

Figure 3.7. The average ergodic rate $\bar{\gamma}$ vs. α for the PPP, the MCP and the MHP. The lines are the average ergodic rates obtained directly from simulations, while the markers are the average ergodic rates estimated using the ADGs.

using the simulation results of the PPP and the ADGs. The approximations using the ADGs are acceptable, although not quite good. The gaps between the values estimated using the ADG and the empirical value are mainly due to the fact that the ADG can provide accurate approximations of the CCDF of the SINR when the SINR threshold θ is small enough, but when θ becomes large, the accuracy of the approximation degrades.

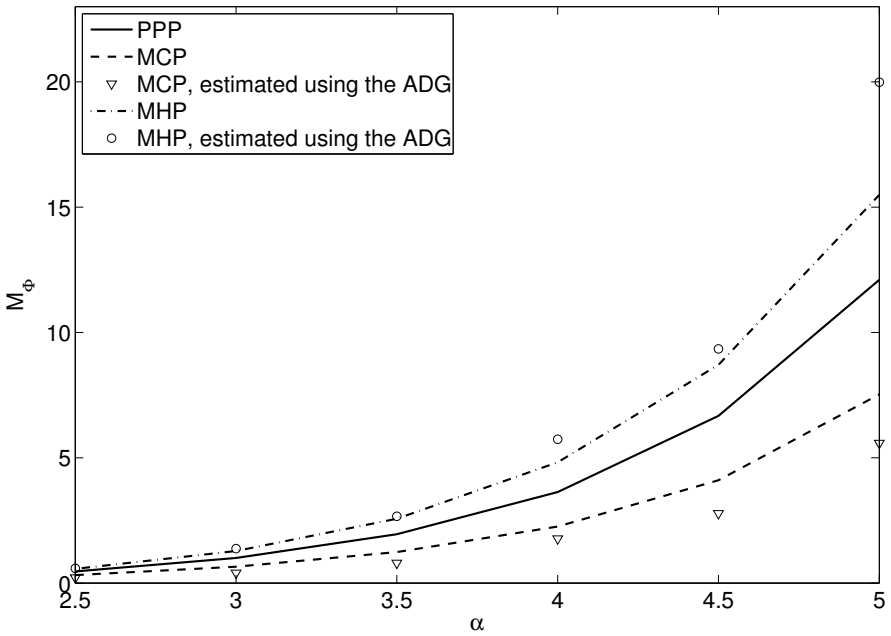

Figure 3.8. The mean SINR M_Φ vs. α for the PPP, the MCP and the MHP. The lines are the mean SINRs obtained directly from simulations, while the markers are the mean SINRs estimated using the ADGs.

3.6 Summary

We examined the asymptotic properties of the SINR distribution for a variety of motion-invariant point processes, given some general assumptions on the point process and general fading assumptions. The assumptions on the point process are satisfied by many commonly used point processes, e.g. the PPP, the MHP and the MCP. Similarly, the fading assumptions are satisfied by Nakagami-m fading and composite fading. We proved that $1 - P_{\mathrm{c}}(\theta) \sim \kappa\theta^m$, as $\theta \to 0$, which shows that the ADG exists.

Under the same system configurations on the fading and path loss, different point processes with the same intensity have different ADGs. Thus, the ADG can be used as a simple metric to characterize the coverage probability. Given the ADG of a point process, we can obtain the precise CCDF of the SINR near 1 by shifting the coverage probability curve of the PPP with the same intensity by the ADG (in dB), and numerical studies show that the shifted coverage probability curve is highly accurate for all practical coverage probabilities.

CHAPTER 4

CONCLUSIONS AND FUTURE WORK

4.1 Conclusions

In this thesis, we first conducted spatial stochastic model fitting in cellular networks using real data sets of BS locations in the UK. It turned out that, the regular point processes outperform the PPP in modeling the network topology, in term of the coverage probability. But those models do not have the tractability of the PPP. As a consequence, they are not widely considered in the literature on wireless networks. By observations from the simulations, we proposed the deployment gain, which can provide somewhat analytical power to those models by utilizing the analytical results for the PPP.

Furthermore, based on the concept of the deployment gain, we studied the asymptotic deployment gain, and theoretically proved its existence for a general class of point process models under general fading assumptions. We proposed it as a simple metric to characterize the coverage probability, and used it to estimate other statistics usually considered in the context of wireless networks, such as the average ergodic rate and the mean SINR.

4.2 Future Work

Possible extensions of this thesis could be as follows:

- The ADG defined in this thesis is the deployment gain as the SINR threshold θ goes to 0. We can explore whether the asymptotic gain exists as θ goes to

∞. If it exists, we can use it together with the ADG we defined to improve the estimations of the average ergodic rate and the mean SINR.

- In the cellular industry, the coverage (i.e., the CCDF of the SINR) is usually evaluated without consideration of small-scale fading. We can investigate whether the ADG exists when only log-normal shadowing is taken into consideration.

- We can study the coverage without fading by re-defining the SINR. In that case, the coverage depends on the BS distribution, the path loss model and the thermal noise. Verifying the existence of the ADG is the key problem. By the study, we can give some insights on how the fading affects the coverage for various BS distributions.

- For some point process models that are usually considered in wireless networks, such as the Poisson cluster process and the Matérn hard-core process, we can derive some closed-form bounds of the ADG.

APPENDIX A

PROOFS

A.1 Proof of Lemma 1

Proof. We first prove that $\forall n \in \mathbb{N}$, there exists a positive $K_0 < \infty$, s.t. $\mathbb{E}(I(\Phi_o^\zeta)^n) \leq K_0 \mathbb{E}(\hat{I}(\Phi^\zeta)^n)$. Let $\zeta \in \mathbb{R}^2$ and $\|\zeta\| = y$. According to Def. 13, for $y > y_0$, $\mathbb{P}(I(\Phi_o^\zeta) > z) \leq \mathbb{P}(\hat{I}(\Phi^\zeta) > z)$, $\forall z \geq 0$, hence $\mathbb{E}(I(\Phi_o^\zeta)^n) \leq \mathbb{E}(\hat{I}(\Phi^\zeta)^n)$. For $y \leq y_0$, we have

$$\begin{aligned} \mathbb{E}(\hat{I}(\Phi^\zeta)^n) &\geq \mathbb{E}(\hat{I}(\Phi^\zeta)^n \mid \Phi^\zeta(b(o,y)) = 0)) \mathbb{P}(\Phi^\zeta(b(o,y)) = 0) \\ &\overset{(a)}{\geq} \mathbb{E}(I(\Phi_o^\zeta)^n) \mathbb{P}(\Phi^{\zeta'}(b(o,y_0)) = 0), \end{aligned} \tag{A.1}$$

where $\zeta' \in \mathbb{R}^2$, $\|\zeta'\| = y_0$, and (a) holds since Φ is motion-invariant, $y_0 \geq y$ and thus $\mathbb{P}(\Phi^{\zeta'}(b(o,y_0)) = 0) \leq \mathbb{P}(\Phi^\zeta(b(o,y)) = 0)$. The second condition in Def. 13 implies that for all $y > 0$, $\forall \zeta \in \mathbb{R}^2$ with $\|\zeta\| = y$, $\mathbb{P}(\Phi^\zeta(b(o,y)) = 0) \neq 0$. So, we have $\mathbb{P}(\Phi^{\zeta'}(b(o,y_0)) = 0) \neq 0$, letting $K_0 = \max\{1, 1/\mathbb{P}(\Phi^{\zeta'}(b(o,y_0)) = 0)\}$, we have

$$\mathbb{E}(I(\Phi_o^\zeta)^n) \leq K_0 \mathbb{E}(\hat{I}(\Phi^\zeta)^n). \tag{A.2}$$

Second, we prove that all moments of $I(\Phi_o^\zeta)$ are bounded. For $n = 1$, by the third

condition in Def. 13, we have

$$\begin{aligned}\mathbb{E}(I(\Phi_o^\zeta)) \leq K_0\mathbb{E}(\hat{I}(\Phi^\zeta)) &= K_0\mathbb{E}_h\mathbb{E}^{!\zeta}\Big(\sum_{x\in\Phi\bigcap B_{\zeta/2}} h_x\ell(x)\Big)\\ &= K_0\mathbb{E}^{!\zeta}\Big(\sum_{x\in\Phi\bigcap B_{\zeta/2}} \mathbb{E}(h_x)\ell(x)\Big)\\ &\overset{(a)}{=} K_0\frac{\mathbb{E}(h)}{\lambda}\int_{B_{\zeta/2}} \ell(x)\rho^{(2)}(x-\zeta)dx, \end{aligned} \tag{A.3}$$

where $\mathbb{E}^{!\zeta}(\cdot)$ is the expectation with respect to the reduced Palm distribution $P^{!\zeta}$, which is the conditional expectation conditioned on $\zeta \in \Phi$ but excluding ζ. (a) follows from the Campbell-Mecke theorem.

For $n \geq 2$, we have

$$\begin{aligned}\mathbb{E}(I(\Phi_o^\zeta)^n) &\leq K_0\mathbb{E}_h\mathbb{E}^{!\zeta}\Big(\sum_{x\in\Phi\bigcap B_{\zeta/2}} h_x\ell(x)\Big)^n\\ &\overset{(a)}{=} K_0\mathbb{E}_h\mathbb{E}^{!\zeta}\Big[\sum_{x\in\Phi\bigcap B_{\zeta/2}} \big(h_x\ell(x)\big)^n\Big] + K_0\sum_{k_1+k_2=n,k_1\geq k_2>0}\binom{n}{k_1,k_2}\\ &\quad\cdot\mathbb{E}_h\mathbb{E}^{!\zeta}\Big[\sum_{x_1,x_2\in\Phi\bigcap B_{\zeta/2}}^{\neq} \big(h_{x_1}\ell(x_1)\big)^{k_1}\big(h_{x_2}\ell(x_2)\big)^{k_2}\Big] + \cdots\\ &\quad + K_0\sum_{\sum_{j=1}^n k_j=n,k_n\geq\cdots\geq k_1>0}\binom{n}{k_1,...,k_n}\mathbb{E}_h\mathbb{E}^{!\zeta}\Big[\sum_{x_1,...,x_n\in\Phi\bigcap B_{\zeta/2}}^{\neq}\prod_{j=1}^n\big(h_{x_j}\ell(x_j)\big)^{k_j}\Big]\\ &\overset{(b)}{=} K_0\frac{\mathbb{E}(h^n)}{\lambda}\int_{B_{\zeta/2}}(\ell(x))^n\rho^{(2)}(x-\zeta)dx + \frac{K_0}{\lambda}\sum_{J=2}^n\sum_{\sum_{j=1}^J k_j=n,k_J\geq\cdots\geq k_1>0}\binom{n}{k_1,...,k_J}\\ &\quad\cdot\Big(\prod_{j=1}^J\mathbb{E}(h^{k_j})\Big)\int_{B_{\zeta/2}}\cdots\int_{B_{\zeta/2}}\prod_{j=1}^J(\ell(x_j))^{k_j}\rho^{(J+1)}(x_1-\zeta,...,x_J-\zeta)dx_1...dx_J, \end{aligned} \tag{A.4}$$

where (a) follows by the multinomial theorem and (b) follows from the Campbell-Mecke theorem.

We discuss the cases of the non-singular and singular path loss models, separately.

For $\ell(x) = (1+\|x\|^{\alpha})^{-1}$, when $n = 1$, since by Def. 13, there exists $q_2 < \infty$, such that $\rho^{(2)}(x) < q_2$ for $x \in \mathbb{R}^2$, it yields that $\int_{B_{\zeta/2}} \ell(x)\rho^{(2)}(x)dx \leq \int_{\mathbb{R}^2} \ell(x)\rho^{(2)}(x)dx < \infty$ and thus by (A.3), there exists $c_1 \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^{\zeta})) < c_1$. Similarly, when $n > 1$, by (A.4), there exists $c_n \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^{\zeta})^n) < c_n$, where c_n does not depend on ζ.

For $\ell(x) = \|x\|^{-\alpha}$, when $n = 1$, we have that $\int_{B_{\zeta/2}} \ell(x)\rho^{(2)}(x)dx \leq q_2 \int_{B_{\zeta/2}} \|x\|^{-\alpha}dx = \frac{2\pi q_2}{(\alpha-2)2^{2-\alpha}}\|\zeta\|^{2-\alpha} \leq \frac{2\pi q_2}{(\alpha-2)2^{2-\alpha}} \max\{1, \|\zeta\|^{2-\alpha}\}$, and hence by (A.3), there exists $c_1 \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^{\zeta})) < c_1 \max\{1, \|\zeta\|^{2-\alpha}\}$. When $n > 1$, for $k_j \in \{1, 2, ..., n\}$, $\int_{B_{\zeta/2}} (\ell(x))^{k_j}dx = \int_{B_{\zeta/2}} \|x\|^{-\alpha k_j}dx = \frac{2\pi}{(\alpha k_j - 2)2^{2-\alpha k_j}}\|\zeta\|^{2-\alpha k_j}$, and therefore, we have $\int_{B_{\zeta/2}} \cdots \int_{B_{\zeta/2}} \prod_{j=1}^{J} (\ell(x_j))^{k_j} dx_1 ... dx_J = (\prod_{j=1}^{J} (\frac{2\pi}{(\alpha k_j - 2)2^{2-\alpha k_j}}))\|\zeta\|^{2J-\alpha n}$. Further, we have $\|\zeta\|^{2J-\alpha n} \leq \max\{1, \|\zeta\|^{2-\alpha n}\}$. Hence, by (A.4), there exists $c_n \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^{\zeta})^n) < c_n \max\{1, \|\zeta\|^{2-\alpha n}\}$.

□

A.2 Proof of Theorem 1

Proof. We first consider the case when the noise power $W = 0$. Since Φ is m.i., we assume $\zeta = (y, 0)$. Let $\hat{\ell}(x) = 1/\ell(x)$. The coverage probability is

$$\begin{aligned} P_{\mathrm{c}}(\theta) &= \mathbb{E}_{\xi}[\mathbb{P}(\mathrm{SINR} > \theta \mid \xi)] \\ &= \int_0^{\infty} \mathbb{P}(h_{\zeta} > \theta\hat{\ell}(\zeta) I(\Phi_o^{\zeta})) f_{\xi}(y)dy \\ &= \int_0^{\infty} \mathbb{E}_{I(\Phi_o^{\zeta})}[F_h^{\mathrm{c}}(\theta\hat{\ell}(\zeta) I(\Phi_o^{\zeta}))] f_{\xi}(y)dy, \end{aligned} \tag{A.5}$$

Thus,

$$\lim_{\theta\to 0} \frac{1 - P_{\mathrm{c}}(\theta)}{\theta^m} = \lim_{\theta\to 0} \int_0^{\infty} \mathbb{E}_{I(\Phi_o^{\zeta})} \left[\frac{F_h(\theta\hat{\ell}(\zeta) I(\Phi_o^{\zeta}))}{\theta^m} \right] f_{\xi}(y)dy. \tag{A.6}$$

Assume $G(t) \triangleq F_h(t)/t^m$, for $t > 0$, and $G(0) = \lim_{t\to 0} F_h(t)/t^m = a$. $\forall \epsilon > 0$,

there exists $\tau > 0$, such that for all $t \in (0, \tau)$, $|G(t) - a| < \epsilon$. So, $G(t) < a + \epsilon$ for $t \in (0, \tau)$. For $t \geq \tau$, $G(t) = F_h(t)/t^m < \tau^{-m}$. Letting $A = \max\{a+\epsilon, \tau^{-m}\}$, we have $G(t) < A$, for all $t \geq 0$.

In the following, we discuss the cases of $\ell(x) = (1 + \|x\|^\alpha)^{-1}$ and $\ell(x) = \|x\|^{-\alpha}$, separately.

For $\ell(x) = (1 + \|x\|^\alpha)^{-1}$, by Lemma 1, we have that $\forall n \in \mathbb{N}$, $\exists c_n \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^\zeta)^n) < c_n$. It follows that

$$\begin{aligned} H(y) &\triangleq \mathbb{E}_{I(\Phi_o^\zeta)}\left[\frac{F_h(\theta\hat{\ell}(\zeta)I(\Phi_o^\zeta))}{\theta^m}\right] \\ &< \mathbb{E}_{I(\Phi_o^\zeta)}\left[A(\hat{\ell}(\zeta)I(\Phi_o^\zeta))^m\right] \\ &< Ac_m\hat{\ell}(y)^m < +\infty, \end{aligned} \tag{A.7}$$

and thus, by the fourth condition in Def. 13,

$$\int_0^\infty H(y)f_\xi(y)dy < Ac_m\mathbb{E}_\xi\big(\hat{\ell}(\xi)^m\big) < +\infty. \tag{A.8}$$

For $\ell(x) = \|x\|^{-\alpha}$, by Lemma 1, we have that $\forall n \in \mathbb{N}$, $\exists d_n \in \mathbb{R}^+$, such that $\mathbb{E}(I(\Phi_o^\zeta)^n) < d_n \max\{1, \|\zeta\|^{2-\alpha n}\}$. Therefore, $H(y) < Ay^{\alpha m} d_m \max\{1, y^{2-\alpha m}\} < +\infty$, and $\int_0^\infty H(y)f_\xi(y)dy < Ad_m\mathbb{E}_\xi\big(\xi^{\alpha m}\max\{1, \xi^{2-\alpha m}\}\big) \leq Ad_m\big(\mathbb{E}_\xi(\xi^{\alpha m}) + \mathbb{E}_\xi(\xi^2)\big) < +\infty$.

Assume $\{\theta_n\}$ is any sequence that converges to 0. Consider $\ell(x) = (1 + \|x\|^\alpha)^{-1}$. Define $\tilde{f}(z) \triangleq a(\hat{\ell}(\zeta)z)^m f_{I(\Phi_o^\zeta)}(z)$, and $\tilde{f}_n(z) \triangleq \frac{F_h(\theta_n\hat{\ell}(\zeta)z)}{\theta_n^m} f_{I(\Phi_o^\zeta)}(z)$, where $f_{I(\Phi_o^\zeta)}(z)$ is the PDF of $I(\Phi_o^\zeta)$. $\{\tilde{f}_n\}$ is a sequence of functions and $\tilde{f}_n \to \tilde{f}$, as $n \to \infty$. Let $g(z) \triangleq A(\hat{\ell}(\zeta)z)^m f_{I(\Phi_o^\zeta)}(z)$. We have that $\tilde{f}_n \leq g$, for all n, and (A.7) indicates $g(z)$ is integrable. By the Dominated Convergence Theorem, we have $\int_0^\infty \tilde{f}(z)dz = \lim_{n\to\infty}\int_0^\infty \tilde{f}_n(z)dz$. Similarly, define $\hat{f}(y) \triangleq \mathbb{E}_{I(\Phi_o^\zeta)}\big[a(\hat{\ell}(\zeta)I(\Phi_o^\zeta))^m\big]f_\xi(y)$, $\hat{f}_n(y) \triangleq \mathbb{E}_{I(\Phi_o^\zeta)}\big[\frac{F_h(\theta_n\hat{\ell}(\zeta)I(\Phi_o^\zeta))}{\theta_n^m}\big]f_\xi(y)$ and $\hat{g}(z) \triangleq Ac_m\hat{\ell}(y)^m f_\xi(y)$. By the Dominated Conver-

gence Theorem, we have $\int_0^\infty \hat{f}(y)dy = \lim_{n\to\infty} \int_0^\infty \hat{f}_n(y)dy$. By the same reasoning, the Dominated Convergence Theorem can also be applied twice for the case $\ell(x) = \|x\|^{-\alpha}$. Thus, for both cases of $\ell(x)$, we obtain that

$$\begin{aligned}\lim_{\theta\to 0} \frac{1 - P_{\mathrm{c}}(\theta)}{\theta^m} &= \int_0^\infty \mathbb{E}_{I(\Phi_o^\zeta)}\left[\lim_{\theta\to 0} \frac{F_h(\theta\hat{\ell}(\zeta)I(\Phi_o^\zeta))}{\theta^m}\right] f_\xi(y)dy \\ &= \int_0^\infty \mathbb{E}_{I(\Phi_o^\zeta)}\left[a(\hat{\ell}(\zeta)I(\Phi_o^\zeta))^m\right] f_\xi(y)dy. \end{aligned} \tag{A.9}$$

Note that by (A.8), (A.9) is finite.

Next, we consider the case when $W > 0$. In (A.6), we only need to replace $I(\Phi_o^\zeta)$ with $(I(\Phi_o^\zeta) + W)$ in the expectation $\mathbb{E}_{I(\Phi_o^\zeta)}(\cdot)$ and the expectation becomes

$$\begin{aligned} H(y) &= \mathbb{E}_{I(\Phi_o^\zeta)}\left[\frac{F_h(\theta\hat{\ell}(\zeta)(I(\Phi_o^\zeta) + W))}{\theta^m}\right] \\ &< \mathbb{E}_{I(\Phi_o^\zeta)}\left[A\hat{\ell}(\zeta)^m(I(\Phi_o^\zeta) + W)^m\right]. \end{aligned} \tag{A.10}$$

By expanding $(I(\Phi_o^\zeta) + W)^m$, we observe that the right-hand side of (A.10) is finite. Analogous to the case when $W = 0$, we can prove that Theorem 1 also holds for $W > 0$. □

A.3 Proof of Corollary 2

Proof. Consider the worst case, $F_h^{\mathrm{c}}(x) \sim \exp(-ax)$, $x \to \infty$. First, we will show that the Laplace transform of $I(\Phi_o^\zeta)$, denoted by $\mathcal{L}_{I(\Phi_o^\zeta)}(s)$, converges for $s > \tau_0$, where $\tau_0 < 0$. Since $\mathcal{L}_{I(\Phi_o^\zeta)}(s)$ always converges for $s \geq 0$, we only consider the case $s < 0$. To prove the property, we need to derive an upper bound of $\mathcal{L}_{I(\Phi_o^\zeta)}(s)$ that only depends on the Φ^ζ. Similar to the proof of Lemma 1, we can prove the proposition that $\forall s < 0$, there exists a positive $K < \infty$, s.t. $\mathbb{E}_{I(\Phi_o^\zeta)}(\exp(-sI(\Phi_o^\zeta))) \leq K\mathbb{E}_{\hat{I}(\Phi^\zeta)}(\exp(-s\hat{I}(\Phi^\zeta)))$.

Thus, we have

$$\begin{aligned}\mathcal{L}_{I(\Phi_o^\zeta)}(s) &= \mathbb{E}_{I(\Phi_o^\zeta)}(\exp(-sI(\Phi_o^\zeta))) \\ &\leq K\mathbb{E}_{\Phi^\zeta,\{h_x\}}\Big(\prod_{x\in\Phi^\zeta\bigcap B_{\zeta/2}\setminus\{\zeta\}}\exp(-sh_x\ell(x))\Big) \\ &= K\mathbb{E}^{!\zeta}\Big(\prod_{x\in\Phi\bigcap B_{\zeta/2}}\mathbb{E}_h\big(\exp(-sh\ell(x))\big)\Big) \\ &= K\mathbb{E}^{!\zeta}\big(\prod_{x\in\Phi\bigcap B_{\zeta/2}}\mathcal{L}_h(s\ell(x))\big), \end{aligned} \tag{A.11}$$

where $\mathcal{L}_h(s)$ denotes the Laplace transform of h.

Let $k(s,x) \triangleq \mathcal{L}_h(s\ell(x))$. We have that $\mathcal{L}_{\hat{I}(\Phi^\zeta)}(s) = \mathbb{E}^{!\zeta}\big(\prod_{x\in\Phi\bigcap B_{\zeta/2}}k(s,x)\big)$ is finite if and only if

$$\eta(s) = \mathbb{E}^{!\zeta}\big(\sum_{x\in\Phi\bigcap B_{\zeta/2}}|\log k(s,x)|\big) < \infty.$$

Now we show that τ_0 is strictly less than 0. We have

$$\begin{aligned}\eta(s) &= \mathbb{E}^{!\zeta}\big(\sum_{x\in\Phi\bigcap B_{\zeta/2}}|\log k(s,x)|\big) \\ &\overset{(a)}{=} \frac{1}{\lambda}\int_{B_{\zeta/2}}\big|\log k(s,x)\big|\rho^{(2)}(x-\zeta)dx, \end{aligned} \tag{A.12}$$

where (a) follows from the Campbell-Mecke theorem.

Since $F_h^c(x) \sim \exp(-ax)$ for large x, without loss of generality, we assume for some large H_0, the PDF of h is $f_\xi(x) = a\exp(-ax)$ $(x > H_0)$. So,

$$\begin{aligned}k(s,x) &= \int_0^\infty \exp(-sy\ell(x))dF_h(y) \\ &= \int_0^{H_0}\exp(-sy\ell(x))dF_h(y) + \int_{H_0}^\infty a\exp\big(-y(a+s\ell(x))\big)dy. \end{aligned} \tag{A.13}$$

Since $x \in \Phi \bigcap B_{\zeta/2}$, by the Dominated Convergence Theorem, $k(s,x)$ is bounded for all x and $s > -a\ell(\|\zeta\|/2)^{-1}$. Also, for $s \in (-a\ell(\|\zeta\|/2)^{-1}, 0)$, we have $k(s,x) > 1$ and $\log(k(s,x)) \leq k(s,x) - 1$. To show $\eta(s) < \infty$ for $s \in (-a\ell(\|\zeta\|/2)^{-1}, 0)$, we need to prove $\int_{B(o,\omega)^c}(k(s,x)-1)\rho^{(2)}(x)dx < \infty$, for large ω. Since for large $\|x\|$, we have $\rho^{(2)}(x-\zeta) \to \lambda^2$, where λ is the intensity of Φ, we choose ω large enough such that $\rho^{(2)}(x)$ is approximately λ^2 for all $\|x\| > \omega$. So we only need to show that $\int_{B(o,\omega)^c}(k(s,x)-1)dx < \infty$. We have

$$\begin{aligned}&\int_{B(o,\omega)^c}(k(s,x)-1)dx\\ &= \int_{B(o,\omega)^c}\int_0^{H_0}(\exp(-sy\ell(x))-1)dF_h(y)dx + \int_{B(o,\omega)^c}\int_{H_0}^{\infty}(\exp(-sy\ell(x))-1)dF_h(y)dx.\end{aligned} \tag{A.14}$$

For large ω,

$$\int_{B(o,\omega)^c}\int_0^{H_0}(\exp(-sy\ell(x))-1)dF_h(y)dx = \int_{B(o,\omega)^c}\int_0^{H_0}(-sy\ell(x))dF_h(y)dx < \infty, \tag{A.15}$$

and

$$\begin{aligned}&\int_{B(o,\omega)^c}\int_{H_0}^{\infty}(\exp(-sy\ell(x))-1)dF_h(y)dx\\ &= \exp(-aH_0)\int_{B(o,\omega)^c}\left(\frac{-s}{a\ell(x)+s} + \frac{a\ell(x)(\exp(-sH_0\ell(x))-1)}{a\ell(x)+s}\right)dx < \infty,\end{aligned} \tag{A.16}$$

Thus, $\eta(s) < \infty$ and $\mathcal{L}_{I(\Phi_o^\zeta)}(s) < \infty$. Since $I(\Phi_o^\zeta)$ is nonnegative, according the region of convergence (ROC) for Laplace transforms, there exists $\tau < -a\ell(\|\zeta\|/2)^{-1}$, such that $\mathcal{L}_{I(\Phi_o^\zeta)}(s)$ converges for $s < \tau$ and diverges for $s > \tau$. τ is called the abscissa of convergence. By Theorem 3 in [51], it follows that the interference has an exponential tail. Therefore, if the fading has at most an exponential tail, the

interference tail is bounded by an exponential. □

A.4 Proof of Lemma 2

Proof. Since $\tilde{h}$ and $\hat{h}$ are independent, we have

$$\begin{aligned} F_h(t) = \mathbb{P}(\tilde{h}\hat{h} \le t) &= \int_0^\infty \mathbb{P}(\tilde{h} \le \frac{t}{u} \mid \hat{h} = u) f_{\hat{h}}(u) du \\ &= \int_0^\infty F_{\tilde{h}}\Big(\frac{t}{u}\Big) f_{\hat{h}}(u) du \\ &= \int_0^\infty \frac{1}{\sqrt{\pi}\Gamma(m)} \Big(\int_0^{\frac{mt}{u}} w^{m-1} \exp(-w) dw \Big) \frac{V_\sigma}{u} \exp(-V_\sigma^2 (\log u)^2) du, \end{aligned} \tag{A.17}$$

where $V_\sigma = \frac{10}{\sigma\sqrt{2}\log 10}$.

To prove the first property, we have

$$\begin{aligned} \lim_{t\to 0} \frac{F_h(t)}{t^m} &= \lim_{t\to 0} \frac{F_h'(t)}{m t^{m-1}} \\ &= \lim_{t\to 0} \int_0^\infty \frac{V_\sigma m^{m-1}}{\sqrt{\pi}\Gamma(m) u^{m+1}} \exp\Big(-\frac{mt}{u} \Big) \exp(-V_\sigma^2 (\log u)^2) du \\ &\le \int_0^\infty \frac{V_\sigma m^{m-1}}{\sqrt{\pi}\Gamma(m) u^{m+1}} \exp(-V_\sigma^2 (\log u)^2) du. \end{aligned} \tag{A.18}$$

Since as $u \to 0$, $\exp(-V_\sigma^2 (\log u)^2) = o(u^n)$ for any $n \in \mathbb{N}$, (A.18) is bounded. Thus we can apply the Dominated Convergence Theorem and obtain the first property.

For the second property, on the one hand, for any $n \in \mathbb{N}$,

$$\begin{aligned} \lim_{t\to\infty} \frac{1 - F_h(t)}{t^{-n}} &= \lim_{t\to\infty} \frac{F_h'(t)}{n t^{-n-1}} \\ &= \lim_{t\to\infty} \int_0^\infty \frac{V_\sigma m^m}{\sqrt{\pi}\Gamma(m) u^{m+1} n} t^{n+m} \exp\Big(-\frac{mt}{u} \Big) \exp(-V_\sigma^2 (\log u)^2) du. \end{aligned} \tag{A.19}$$

Assume $H(t) = t^{n+m} \exp\big(-\frac{mt}{u}\big)$. Since $H'(t) = t^{n+m-1}\big(n+m-\frac{mt}{u}\big) \exp\big(-\frac{mt}{u}\big)$, when $t = \frac{u(n+m)}{m}$, $H(t)$ achieves its maximum value and $\max_{t>0} H(t) = \big(\frac{u(n+m)}{m}\big)^{n+m}$.

$\exp(-(n+m))$. Thus,

$$\lim_{t\to\infty}\frac{1-F_h(t)}{t^{-n}} \le \int_0^\infty \frac{V_\sigma u^{n-1}}{\sqrt{\pi}\Gamma(m)n}\frac{(n+m)^{n+m}}{m^n}\exp(-(n+m))\exp(-V_\sigma^2(\log u)^2)du$$
$$< \infty. \tag{A.20}$$

Applying the Dominated Convergence Theorem, we obtain $\lim_{t\to\infty}\frac{1-F_h(t)}{t^{-n}} = 0$ and thus $F_h^{\mathrm{c}}(t) = o(t^{-n})$, as $t\to\infty$, for any $n\in\mathbb{N}$.

On the other hand, for any $a>0$,

$$\lim_{t\to\infty}\frac{1-F_h(t)}{\exp(-at)} = \lim_{t\to\infty}\frac{F_h'(t)}{a\exp(-at)}$$
$$= \lim_{t\to\infty}\int_0^\infty \frac{V_\sigma m^m}{\sqrt{\pi}\Gamma(m)u^{m+1}a}t^{m-1}\exp\left(\left(a-\frac{m}{u}\right)t\right)\exp(-V_\sigma^2(\log u)^2)du. \tag{A.21}$$

For any $a>0$, there exists $\hat{K}>0$, such that for $u>\hat{K}$, $\exp(mt/u) < \exp(at/3)$. Hence, $\lim_{t\to\infty}\frac{1-F_h(t)}{\exp(-at)} = \infty$, for any $a>0$. Thus, $-\log F_h^{\mathrm{c}}(t) = o(t)$, $t\to\infty$. □

A.5 Proof of Lemma 3

Proof. Conditions 1 and 2 in Def. 13 hold for all the three point processes obviously. For Conditions 3 and 4, we treat the three point processes separately.

For the PPP, Condition 3 holds, because the points in Φ are independent; Condition 4 holds, because $\mathbb{P}(\xi>x) = \mathbb{P}(\Phi(b(o,x))=0) = \exp(-\lambda\pi x^2)$.

For the MCP, we first prove that Condition 3 holds. For $y>r_{\mathrm{c}}$, the interference $I(\Phi_o^\zeta)$ consists of two parts. One is the interference from the clusters with center points inside the region $B(o,y+r_{\mathrm{c}})\setminus b(o,y-r_{\mathrm{c}})$, denoted by I_1, and the other part is the interference from the clusters with center points in $B(o,y+r_{\mathrm{c}})^{\mathrm{c}}$, denoted by I_2. I_1 and I_2 are independent. Similarly, $\hat{I}(\Phi^\zeta)$ consists of $\hat{I}_1$ and $\hat{I}_2$, where $\hat{I}_1$ is from the clusters with center points inside $B(o,y+r_{\mathrm{c}})\setminus b(o,y/2)$ and $\hat{I}_2$ is from the clusters

with center points in $B(o, y + r_{\rm c})^{\rm c}$.

Since the parent points are independent, I_2 and $\hat{I}_2$ have the same distribution. For $y \gg r_{\rm c}$, we can easily prove that $\hat{I}_1$ stochastically dominates I_1. As $\mathbb{P}(I(\Phi_o^\zeta) > z) = \mathbb{P}(I_1 + I_2 > z) = \mathbb{E}_{I_2}[\mathbb{P}(I_1 > z - I_2 \mid I_2)]$, we have $\mathbb{P}(I(\Phi_o^\zeta) > z) \leq \mathbb{P}(\hat{I}(\Phi^\zeta) > z)$ for all $z \geq 0$.

Then we prove Condition 4 holds for the MCP. For large y, let $\mathcal{S}$ be the set of the parent points that are in $B(o, y - r_{\rm c})$, i.e., $\mathcal{S} = \{x \in \Phi_{\rm p} : x \in B(o, y - r_{\rm c})\}$ and $\tilde{\Phi}_x$ be the daughter process for the cluster centered at $x \in \Phi_{\rm p}$. We have

$$\begin{aligned}
&\mathbb{P}(\xi > y) = \mathbb{P}(\Phi(B(o, y)) = 0) \\
&\overset{(a)}{\leq} \mathbb{P}(\tilde{\Phi}_x(B(x, r_{\rm c})) = 0, \text{ for all } x \in \mathcal{S}) \\
&= \sum_{k=0}^{\infty} \frac{(\lambda_{\rm p}\pi(y - r_{\rm c})^2)^k \exp(-\lambda_{\rm p}\pi(y - r_{\rm c})^2)}{k!} \exp(-\bar{c})^k \\
&= \exp\big(-(1 - \exp(-\bar{c}))\lambda_{\rm p}\pi(y - r_{\rm c})^2\big), \qquad \text{(A.22)}
\end{aligned}$$

where (a) follows since $\Phi(B(o, y)) = 0$ implies $\Phi(B(x, r_{\rm c})) = 0$, for all $x \in \mathcal{S}$. As $\mathbb{E}(\xi^n) = -\int z^n d\mathbb{P}(\xi > z)$, performing integration by parts, it follows that $\mathbb{E}(\xi^n)$ is bounded.

For the MHP, to prove Condition 3, we consider Φ_o^ζ and Φ^ζ in term of the base PPP $\Phi_{\rm b}$. Conditioned on $\Phi_{\rm b} \bigcap (B(o, y + 2r_{\rm h}) \setminus B(o, y + r_{\rm h}))$, the interference from the region $B(o, y + 2r_{\rm h})^{\rm c}$ in Φ^ζ and that in Φ_o^ζ are i.i.d.. So we only need to consider the region $B(o, y + 2r_{\rm h})$ for large y. As y grows, $\mathbb{E}[\Phi^\zeta(B(o, y) \setminus B(o, y/2))] = \Theta(y^2)$, and $\mathbb{E}[\Phi_o^\zeta(B(o, y + 2r_{\rm h}) \setminus B(o, y))] = \Theta(y)$.[1] It can be proved that the portion of $\hat{I}(\Phi^\zeta)$ that comes from the retained points in $B(o, y + 2r_{\rm h}) \setminus B(o, y/2)$ stochastically dominates the portion of $I(\Phi_o^\zeta)$ that comes from the retained points in $B(o, y + 2r_{\rm h}) \setminus B(o, y)$. Hence, Condition 3 holds.

[1] $f(x) = \Theta(g(x))$, if both $f(x)/g(x)$ and $g(x)/f(x)$ remain bounded as $x \to \infty$.

To prove that Condition 4 holds for the MHP, we use the CCDF of ξ expressed in the form (15.1.5) in [52]:

$$F_{\xi}^{\mathrm{c}}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{B(o,x)} \cdots \int_{B(o,x)} \rho^{(k)}(y_1, \ldots, y_k) dy_1 \cdots dy_k \tag{A.23}$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \alpha^{(k)}[B(o,x)^{\otimes k}], \tag{A.24}$$

where $B(o,x)^{\otimes k}$ is the Cartesian product of k balls and $\alpha^{(k)}$ is the kth-order factorial moment measure. For the MHP, the nth moment density satisfies

$$\rho^{(n)}(z_1, \ldots, z_n) = \lambda^n, \quad \text{for } (z_1, \ldots, z_n) \in S_n(x), \tag{A.25}$$

where $S_n(x) \triangleq \{(z_1, \ldots, z_n) \in B(o,x)^{\otimes n} : \|z_i - z_j\| > 2r_{\mathrm{h}}, \forall i \neq j\}$. The complementary set of $S_n(x)$ with respect to $B(o,x)^{\otimes n}$ is $S_n^{\mathrm{c}}(x) = B(o,x)^{\otimes n} \setminus S_n(x) = \{(z_1, \ldots, z_n) \in B(o,x)^{\otimes n} : \exists i \neq j, \text{ s.t. } \|z_i - z_j\| \leq 2r_{\mathrm{h}}\}$. The Lebesgue measure of $S_n^{\mathrm{c}}(x)$ satisfies $\nu(S_n^{\mathrm{c}}(x)) = O(x^{2n-1})$. So, as $x \to \infty$, $\int_{S_n^{\mathrm{c}}(x)} \rho^{(n)}(y_1, \ldots, y_n) dy_1 \cdots dy_n \to 0$. Since (A.23) can be rewritten as

$$F_{\xi}^{\mathrm{c}}(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\int_{S_k(x)} \rho^{(k)}(y_1, \ldots, y_k) dy_1 \cdots dy_k + \int_{S_k^{\mathrm{c}}(x)} \rho^{(k)}(y_1, \ldots, y_k) dy_1 \cdots dy_k \right),$$

it follows that as $x \to \infty$,

$$F_{\xi}^{\mathrm{c}}(x) \sim \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\int_{S_k(x)} \rho^{(k)}(y_1, \ldots, y_k) dy_1 \cdots dy_k + \int_{S_k^{\mathrm{c}}(x)} \lambda^k dy_1 \cdots dy_k \right)$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} (\lambda \pi x^2)^k = \exp(-\lambda \pi x^2).$$

Therefore, $\mathbb{E}(\xi^n)$ is bounded for all n and Condition 4 holds.

□

BIBLIOGRAPHY

1. V. P. Mhatre and C. P. Rosenberg, "Impact of Network Load on Forward Link Inter-Cell Interference in Cellular Data Networks," *IEEE Transactions on Wireless Communications*, Vol. 5, No. 12, pp. 3651-3661, Dec. 2006.

2. P. Charoen and T. Ohtsuki, "Codebook Based Interference Mitigation with Base Station Cooperation in Multi-Cell Cellular Network," *IEEE Vehicular Technology Conference 2011*, pp. 1-5, Sep. 2011

3. F. G. Nocetti, I. Stojmenovic, and J. Zhang, "Addressing and routing in hexagonal networks with applications for tracking mobile users and connection rerouting in cellular networks," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 13, No. 9, pp. 963-971, Sep. 2002.

4. M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 27, No. 7, pp. 1029-1046, Sep. 2009.

5. M. Haenggi, *Stochastic Geometry for Wireless Networks*, Cambridge University Press, 2012.

6. F. Baccelli and B. Blaszczyszyn, *Stochastic Geometry and Wireless Networks*, NOW: Foundations and Trends in Networking, 2010.

7. D. Stoyan, W. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications, 2nd edition*, John Wiley and Sons, 1995.

8. J. G. Andrews, F. Baccelli, and R. K. Ganti, "A Tractable Approach to Coverage and Rate in Cellular Networks," *IEEE Transactions on Communications*, Vol. 59, No. 11, pp. 3122-3124, Nov. 2011.

9. H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and Analysis of K-Tier Downlink Heterogeneous Cellular Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 30, No. 3, pp. 550-560, Apr. 2012.

10. T. T. Vu, L. Decreusefond, and P. Martins, "An analytical model for evaluating outage and handover probability of cellular wireless networks," *15th International Symposium on Wireless Personal Multimedia Communications (WPMC'12)*, pp. 643-647, Sep. 2012.

11. D. Cao, S. Zhou, and Z. Niu, "Optimal Base Station Density for Energy-Efficient Heterogeneous Cellular Networks," *IEEE ICC'12*, pp. 4379-4383, 2012.

12. N. Deng, S. Zhang, W. Zhou, and J. Zhu, "A stochastic geometry approach to energy efficiency in relay-assisted cellular networks," *IEEE GLOBECOM'12*, pp. 3484-3489, 2012.

13. C.-H. Lee and M. Haenggi, "Interference and Outage in Poisson Cognitive Networks," *IEEE Transactions on Wireless Communications*, Vol. 11, No. 4, pp. 1392-1401, Apr. 2012.

14. X. Liu and M. Haenggi, "Towards Quasi-Regular Sensor Networks: Topology Control Algorithms for Improved Energy Efficiency," *IEEE Transactions on Parallel and Distributed Systems*, Vol. 17, No. 9, pp. 975-986, Sep. 2006.

15. A. Busson, G. Chelius, and J. M. Gorce, "Interference Modeling in CSMA Multi-Hop Wireless Networks," *Tech. Rep. 6624, INRIA*, Feb. 2009.

16. A. Busson and G. Chelius, "Point Processes for Interference Modeling in CSMA/CA Ad Hoc Networks," *Sixth ACM International Symposium on Performance Evaluation of Wireless Ad Hoc, Sensor, and Ubiquitous Networks (PE-WASUN09)*, pp. 33-40, Oct. 2009.

17. M. Haenggi, "Mean Interference in Hard-Core Wireless Networks," *IEEE Communications Letters*, Vol. 15, No. 8, pp. 792-794, Aug. 2011.

18. H. Q. Nguyen, F. Baccelli, and D. Kofman, "A Stochastic Geometry Analysis of Dense IEEE 802.11 Networks," *IEEE INFOCOM'07*, pp. 1199-1207, 2007.

19. C. Geyer, "Likelihood inference for spatial point processes", in *Current Trends in Stochastic Geometry and Applications*, ed. by O. E. Barndorff-Nielsen, W. S. Kendall, and M. N. M. Lieshout, London: Chapman and Hall, pp. 141-172, 1999.

20. J. Riihijarvi and P. Mahonen, "A spatial statistics approach to characterizing and modeling the structure of cognitive wireless networks," *Ad Hoc Networks*, Vol. 10, No. 5, pp. 858-869, 2012.

21. P. Mitran and C. Rosenberg, "On fractional frequency reuse in imperfect cellular grids," *IEEE Wireless Communications and Networking Conference (WCNC) 2012*, pp. 2967-2972, Apr. 2012.

22. A. Baddeley and R. Turner, "Practical Maximum Pseudolikelihood for Spatial Point Patterns," *Australian and New Zealand Journal of Statistics*, Vol. 42, No. 3, pp. 283-322, 2000.

23. P. J. Diggle and R. J. Gratton, "Monte Carlo Methods of Inference for Implicit Statistical Models", *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 46, No. 2, pp. 193-227, 1984.

24. A. Baddeley, *Analysing Spatial Point Patterns in R*, Version 3, CSIRO, 2008.

25. The R Project for Statistical Computing, `http://www.r-project.org` [Accessed March 2014].

26. A. Baddeley and G. Nair, "Fast approximation of the intensity of Gibbs point processes," *Electronic Journal of Statistics*, Vol. 6, pp. 1155-1169, 2012.

27. R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert W function," *Advances in Computational Mathematics*, Vol. 5, No. 4, pp. 329-359, 1996.

28. A. Baddeley and R. Turner. "spatstat: An R Package for Analyzing Spatial Point Patterns," *Journal of Statistical Software*, Vol. 12, Iss. 6, pp. 1-42, Jan. 2005.

29. K. K. Berthelsen and J. Moller, "A Primer on Perfect Simulation for Spatial Point Processes," *Bulletin of the Brazilian Mathematical Society*, Vol. 33, No. 3, pp. 351-367, 2002.

30. J. S. Ferenc and Z. Neda, "On the size distribution of Poisson Voronoi cells," *Physica A: Statistical Mechanics and its Applications*, Vol. 385, Iss. 2, pp. 518-526, Nov. 2007.

31. M. Tanemura, "Statistical Distributions of Poisson Voronoi Cells in Two and Three Dimensions," *Forma*, Vol. 18, No. 4, pp. 221-247, 2003.

32. S. Lin and D. J. Costello, *Error Control Coding*, 2nd ed., Englewood Cliffs, NJ: Prentice-Hall, 2004.

33. J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A primer on spatial modeling and analysis in wireless networks," *IEEE Communications Magazine,* Vol. 48, No. 11, pp. 156-163, Nov. 2010.

34. T. D. Novlan, R. K. Ganti, A. Ghosh, and J. G. Andrews, "Analytical Evaluation of Fractional Frequency Reuse for OFDMA Cellular Networks," *IEEE Transactions on Wireless Communications,* Vol. 10, No. 12, pp. 4294-4305, Dec. 2011.

35. H. S. Dhillon, R. K. Ganti, F. Baccelli and J. G. Andrews, "Modeling and Analysis of K-Tier Downlink Heterogeneous Cellular Networks," *IEEE Journal on Selected Areas in Communications*, Vol. 30, No. 3, pp. 550-560, Apr. 2012.

36. W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput Optimization, Spectrum Allocation, and Access Control in Two-Tier Femtocell Networks," *IEEE Journal on Selected Areas in Communications,* Vol. 30, No. 3, pp. 561-574, Apr. 2012.

37. Y. Zhong and W. Zhang, "Multi-Channel Hybrid Access Femtocells: A Stochastic Geometric Analysis," *IEEE Transactions on Communications,* Vol. 61, No. 7, pp. 3016-3026, Jul. 2013.

38. S. P. Weber, X. Yang, J. G. Andrews, and G. de Veciana, "Transmission capacity of wireless ad hoc networks with outage constraints," *IEEE Transactions on Information Theory,* Vol. 51, No. 12, pp. 4091-4102, Dec. 2005.

39. S. Bandyopadhyay and E. J. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," *Proc. IEEE INFOCOM 2003,* Vol. 3, pp. 1713-1723, Apr. 2003.

40. R. K. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Transactions on Information Theory*, Vol. 55, No. 9, pp. 4067-4086, Sep. 2009.

41. H. ElSawy and E. Hossain, "A Modified Hard Core Point Process for Analysis of Random CSMA Wireless Networks in General Fading Environments," *IEEE Transactions on Communications,* Vol. 61, No. 4, pp. 1520-1534, Apr. 2013.

42. M. Haenggi and R. K. Ganti, "Interference in Large Wireless Networks," *Foundations and Trends in Networking*, Vol. 3, No. 2, pp. 127-248, 2008.

43. R. K. Ganti, J. G. Andrews, and M. Haenggi, "High-SIR Transmission Capacity of Wireless Networks with General Fading and Node Distribution," *IEEE Transactions on Information Theory*, Vol. 57, pp. 3100-3116, May 2011.

44. R. Giacomelli, R. K. Ganti, and M. Haenggi, "Outage Probability of General Ad Hoc Networks in the High-Reliability Regime," *IEEE/ACM Transactions on Networking*, Vol. 19, pp. 1151-1163, Aug. 2011.

45. R. K. Ganti and M. Haenggi, "Interference in Ad Hoc Networks with General Motion-Invariant Node Distributions," *in 2008 IEEE International Symposium on Information Theory (ISIT'08)*, (Toronto, Canada), Jul. 2008.

46. R. K. Ganti and J. G. Andrews, "A new method for computing the transmission capacity of non-Poisson wireless networks," *IEEE ISIT'10*, Jun. 2010.

47. N. Deng, W. Zhou, and M. Haenggi, "The Ginibre Point Process as a Model for Wireless Networks with Repulsion," *IEEE Transactions on Wireless Communications*, 2014. Submitted. Available at `http://www3.nd.edu/~mhaenggi/pubs/twc14c.pdf` [Accessed April 2014].

48. F. Hansen and F. I. Meno, "Mobile fading-Rayleigh and lognormal superimposed," *IEEE Transactions on Vehicular Technology,* Vol. 26, No. 4, pp. 332-335, Nov. 1977.

49. A. M. D. Turkmani, "Probability of error for M-branch macroscopic selection diversity," *IEE Proceedings I Communications, Speech and Vision,* Vol. 139, No. 1, pp. 71-78, Feb. 1992.

50. D. Xenakis, N. Passas, L. Merakos, and C. Verikoukis, "Energy-efficient and interference-aware handover decision for the LTE-Advanced femtocell network," *IEEE ICC'13*, Jun. 2013.

51. K. Nakagawa, "Application of Tauberian theorem to the exponential decay of the tail probability of a random variable," *IEEE Transactions on Information Theory*, Vol. 53, pp. 3239-3249, Sep. 2007.

52. D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure (Vol. 2)*, Springer, second edition, 2007.

53. S. Weber and J.G. Andrews, "A Stochastic geometry approach to Wideband Ad Hoc Networks with Channel Variations", *Workshop on Spatial Stochastic Models for Wireless Networks,* Apr. 2006.

This document was prepared & typeset with pdfLaTeX*, and formatted with* NDdiss2$_\varepsilon$ *classfile (v3.2013[2013/04/16]) provided by Sameer Vijay and updated by Megan Patnott.*