STOCHASTIC GEOMETRY MODELLING AND PERFORMANCE EVALUATION OF CELLULAR NETWORKS

by

Chunlin Chen

A thesis submitted in partial fulfillment of the requirements
for the degree of

Doctor of Philosophy
in
Communications

Department of Electrical and Computer Engineering
University of Alberta

Abstract

Stochastic geometry provides a way of defining and computing macroscopic properties of large scale wireless networks, by averaging over all possible spatial patterns of the network nodes. It abstracts the network as realizations of point process models, and analyzes the network performance in a probabilistic way. While stochastic geometry has its strength in theoretical analysis, statistical questions are seldom discussed and measurement-based validation of certain stochastic assumptions used in the literature is often not given. The aim in most of the related research work in the literature is to show what results can be obtained with stochastic geometry when assumptions of certain point process models are made, without necessarily fully justifying those assumptions. It is critical to find an accurate point process model that best reflects the spatial distribution of the network nodes before any attempts on the theoretical analysis of the underlying point process model. In addition, extensions of the analytical methodology used in Poisson models to more general point process models are often hindered due to the lack of closed-form empty space function and the probability generating functional (PGFL). In view of these problems, the thesis presents and describes a practical technique of statistical validation by fitting stationary and nonstationary point process models to real-life cellular networks using maximum likelihood/pseudolikelihood and minimum contrast methods. We also have studied the distributional properties of the empty space distances in the Matérn hard core point process of Type II, and proposed a piecewise probability density function for the empty space distance, including an exact expression and a heuristic formula, which can be fitted by a Weibull-like function. Fur-

thermore, we have examined the properties of the PGFL for Poisson cluster processes, and studied the downlink coverage performance for a two-tier cellular network.

Preface

The essential content of Chapter 3 of the thesis has been published as

- C. Chen, R. C. Elliott, W. A. Krzymień, and J. Melzer, "Modeling of cellular networks using stationary and nonstationary point processes," *IEEE Access*, 13 August 2018, available under the Early Access on IEEE Xplore, DOI: 10.1109/ACCESS.2018.2865182.

The essential content of Chapter 4 of the thesis has been published as

- C. Chen, R. C. Elliott, and W. A. Krzymień, "Empirical distribution of nearest-transmitter distance in wireless networks modeled by Matérn hard core point processes," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1740–1749, Feb. 2018.

The essential content of Chapter 5 of the thesis has been published as

- C. Chen, R. C. Elliott, and W. A. Krzymień, "Downlink coverage analysis of N-tier heterogeneous cellular networks based on clustered stochastic geometry," in *Proc. 2013 Asilomar Conf. on Signals, Systems, and Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 1577–1581.

For each publication listed above, I carried out the system model development, the mathematical analysis and computer simulations under the supervision of Prof. Witold Krzymień and cooperation of Dr. Robert C. Elliott and Dr. Jordan Melzer (for Chapter 3). I was responsible for the manuscript composition, Dr. Elliott and Prof. Krzymień contributed to manuscript edits.

Acknowledgements

I would like to express my gratitude to my supervisor Prof. Witold A. Krzymień for the amazing opportunity he provided me with. His flexibility, guidance, and encouragement have pushed me far beyond my expectations. I would like to thank Dr. Robert C. Elliott and Dr. Jordan Melzer for stimulating discussions and sincere advice and help.

I would like to express my gratitude to the three most important persons in my life, my parents and my wife. Starting from my mother Xinhong Liang, I want to say, I would have never been in this position without your love, care, and support. My father Xin Chen, you have always been my mentor, you trained me well for the struggle of life. My wife, beautiful, supporting, and loving wife Yan Han, without you, nothing is complete.

I would like to thank my examiners Professors Xiaodai Dong, Chintha Tellambura, Yindi Jing, Hao Liang, Masoud Ardakani, and Hai Jiang. I acknowledge the financial support from NSERC and TRTech. I would like to thank the staff at the Department of Electrical and Computer Engineering. I thank all my group members and lab colleagues, it was fun working with all of you.

Finally, I would like to acknowledge a lot of sincere and loving people, I have been blessed with, that have strongly contributed to this success, including my uncle Xu Chen, aunts Yu Chen and Ping Chen, and my best friends, Jun Yang and Houdao Wu.

Contents

List of Tables

List of Figures

Symbol Notation

$\mathcal{A}$	Subset of $\mathbb{R}^2 \times \mathbb{R}^2$
$B(o, r)$	The circle centered at the origin o with radius r
$\mathcal{B}$	Borel set (bounded subset) of $\mathbb{R}^2$
$\lvert\mathcal{B}\rvert$	Lebesgue measure of B
$\mathrm{b}\left(x_i\right)$	First-order term (or intensity parameter) at x_i
C	Summary statistic
$\hat{C}$	Estimated C
$\bar{C}$	Simulated average of C
$C^L(r)$	Lower bound of $\hat{C}$ at distance r
$C^U(r)$	Upper bound of $\hat{C}$ at distance r
$\mathrm{c}\left(x_i, x_j\right)$	Second-order or pairwise interaction term between x_i and x_j
$\mathrm{d}\left[x_i, x_j\right]$	Distance between x_i and x_j
$\mathcal{D}$	Arc length
$\mathcal{D}'$	Chord length
$\mathrm{e}(x_i, x_j, r)$	Edge correction between x_i and x_j separated by distance r
$\mathrm{e}(u)$	Edge effect bias correction
$F(r)$	Empty space function
$F_R(r)$	Cumulative distribution function of R
$f\left(\phi\right)$	Probability density of pairwise interaction models
$\mathcal{G}(\ell)$	Probability generating functional w.r.t. function $\ell(\cdot)$ r
$G(r)$	Nearest neighbour distance distribution function
$g(r)$	Pair correlation function
$g_{in}(r)$	Inhomogeneous pair correlation function
h_x	Random channel gain between x and the origin o
$h_{x,u}$	Random channel gain between x and u

$\mathrm{h}(x_i, x_j)$	Fraction of the circumference of the circle with center x_i and radius $d[x_i, x_j]$
I_{c}	Cumulative interference
I_{M}	Mean interference
$I_{\mathrm{M}}^{r_w}$	Mean interference measured at the origin over finite networks of radius r_w
$\mathcal{I}$	Set of the locations of the interfering transmitters
$J(r)$	J function
$K(r)$	K function
$K_{in}(r)$	Inhomogeneous K function
$\mathcal{K}$	Intensity function at location x
$\mathrm{k}(u)$	Kernel function at u
$L(r)$	L function
$\ell(\cdot)$	Path-loss function
L^{W}	Side length of sampling window W
$\mathcal{L}_{I_c}$	Laplace transform of I_c
$\mathbf{L}$	The maximum value of the likelihood function $\mathbb{P}(\phi \vert \{\theta, \Phi\})$
$\mathbb{N}$	The family of all sequences
$\mathcal{N}$	The smallest σ-algebra on $\mathbb{N}$
$N\left(\Phi_{\mathrm{P}}^{\mathrm{in}}(\mathcal{B})\right)$	The number of points of Φ in Borel set $\mathcal{B}$
$n(\phi)$	The number of points in ϕ
$P_r(u)$	Received power at u
$P_t(x)$	Transmit power from x
p_c^R	Conditional probability of coverage with a desired link distance R
p, q	Exponents for minimum contrast method
$\mathrm{PL}(\vartheta, \phi)$	Pseudolikelihood given ϑ and ϕ
$\mathbb{R}^2$	Euclidean space
R_{e}	The Earth radius
R	Desired link distance (or empty space distance)
R_{SP}	Interaction radius for Strauss point process

R_{H}	Hard core distance for hard core point process
$S(u)$	First-order term in $\lambda(u, \phi)$
T	SINR threshold
$\mathrm{t}\,(u, \phi)$	The number of points of ϕ that lie within a distance R_{SP} from the location u
$V\,(u, \phi)$	Second-order term in $\lambda(u, \phi)$
W	Study region
$\lvert W \rvert$	Area of W
w	Normalizing constant with regard to Φ
w_{SP}	Normalizing constant with regard to stationary Strauss point process
w_{H}	Normalizing constant with regard to stationary hard core point process
(x_u, y_u)	Spatial coordinates of the point at location u
$Z(u)$	Covariate function
Φ	Point process
ϕ	Sequence of points on $\mathbb{R}^2$
Φ_x	Translated point process Φ with respect to the point x
$\Phi(\mathcal{B})$	$\Phi \cap \mathcal{B}$
Φ_p	Stationary Poisson point process
Φ_p^{in}	Nonstationary Poisson point process
λ	Intensity of point process Φ
$\lambda(x)$	Intensity function at location x
λ_p	Intensity of Φ_p
λ_{MH}	Intensity of the Matérn hard core point process of Type II
$\lambda(u, \phi)$	Conditional intensity at location u given the observed point pattern ϕ
λ_t	Intensity of the Thomas cluster process
$\tilde{\lambda}(u)$	Smoothed version of the true intensity function $\lambda(u)$
$\hat{\lambda}(u)$	Kernel estimator of $\lambda(u)$
$\lambda_\vartheta\,(u, \phi)$	Conditional intensity governed by ϑ
$\Gamma(\cdot)$	Gamma function

$\gamma(\cdot,\cdot)$	Incomplete gamma function
γ_{SP}	Pairwise interaction term of stationary Strauss point process
$\theta(r)$	The fraction of the infinitesimal thin ring of a circle with radius r
θ_{e}	The radius angle between two site locations on earth
ς	Intensity measure of the point process Φ
$\varsigma^{(2)}$	The second-order factorial moment measure of Φ
β_{SP}	Intensity parameter of stationary Strauss point process
β_{H}	Intensity parameter of stationary hard core point process
ϑ	Set of parameters (κ, μ, σ) for the stationary Thomas cluster process
ϑ	Set of regular parameters
$\hat{\vartheta}$	Estimator of θ
$\rho^{(2)}$	The second-order reduced moment measure
$\varrho^{(2)}$	Second-order product density
$\|\cdot\|$	Euclidean norm
η	Path-loss exponent
Ω	Noise power
$\mathbf{1}_{\{\cdot\}}$	Indicator function
κ	Parent intensity of clustered point processes
μ	Offspring intensity of clustered point processes
δ	Hard core parameter of the Matérn hard core point process of Type II
ψ	Regular parameter
ξ	Regular parameter
(α_0, α_1)	Linear coefficients

List of Abbreviations

AIC	Akaike information criterion
BS	Base station
CDF	Cumulative distribution function
CMHC	Complementary Matérn hard core point process
CSMA	Carrier sensing multiple access
CSR	Complete spatial randomness
HCPP	Hard core point process
HetNets	Heterogeneous networks
i.i.d.	Independent and identically distributed
KS	Kolmogorov-Smirnov test
LTE	Long term evolution
MAC	Medium access control
MCP	Matérn cluster process
MHCPP	Matérn hard core point process of Type II
MLE	Maximum likelihood estimation
PCP	Poisson cluster process
PDF	Probability density function
PGFL	Probability generating functional
PPP	Poisson point process
PW	Piecewise
QoS	Quality of service
RMSD	Root mean square deviation
RMSE	Root mean square error
SINR	Signal-to-interference-plus-noise ratio
SP	Strauss point process

TCP	Thomas cluster process
TDMA	Time division multiple access
Wi-Fi	Wireless local area networks based on the IEEE 802.11 standards

Chapter 1

Introduction

1.1 Background

Wireless communication services have experienced dramatic growth over the past 25 years. It was in the 1970s and early 1980s that the first generation cellular systems based on frequency division multiple access and analog FM technology were developed and deployed [1]. Today, digital cellular telephone services are available throughout the world, and have well surpassed fixed-line services both in terms of availability and number of users. While mobile voice telephony drove the past growth of wireless systems and still remains an important application, it is abundantly clear that wireless data applications are driving its future growth. In the past three decades, the Internet transformed from being an academic tool to an indispensable global information network providing a vast array of services and applications – from e-mail to social networking and e-commerce to entertainment.

Due to the scarcity of the radio resources along with the ever-increasing pressure of the rapid proliferation of mobile devices with powerful data-consuming capability, it is not rational in cellular networks to separate concurrent transmissions completely in frequency. Some transmissions will inevitably be made at the same time over the same frequency band, separated only in space. The traditional homogeneous network

expansion techniques via cell splitting cannot cope with the rapid growth of user population and their associated traffic. Moreover, macro base station (BS) deployment necessitates a huge capital expenditure which would be very difficult to recover with the decreasing service cost. In response to these major developments, the industry is driving the standardization bodies to develop new solutions to accommodate the increased capacity demand, such as universal frequency reuse, massive MIMO, and heterogeneous networks (HetNets) [2–6].

A universal frequency reuse will increase the spatial spectrum efficiency and network capacity but at the cost of increased interference from undesired transmitters added to the desired signal at a receiver. There are three main factors that shape the interference statistics in wireless networks, including the network geometry (layout of the network transmitters), medium access control (MAC) protocols (e.g., ALOHA [7–9], carrier-sense multiple access (CSMA) [10–16], time division multiple access (TDMA) [17–19], etc.), and the random channel behaviour due to small-scale fading and shadowing. It is often the combined action of the network geometry and MAC that determines the distribution of concurrently transmitting nodes. For example, even if the network nodes are very randomly (e.g. uniformly) distributed in space, a CSMA scheme [13] will ensure a certain separation between concurrent transmitters by translating the spectrum sensing power threshold into a minimum exclusion distance; hence the spatial distribution of the active transmitters at any given time slot might exhibit some regularity.

1.2 Some Relevant Wireless Performance Measures

In wireless communications, the signal power radiated from a transmitting node decays with Euclidean distance, and hence the geometry of the locations of the transmitting and receiving nodes plays a key role, since it determines the signal-to-interference-plus-noise ratio (SINR) at each receiver. The SINR then determines the information-

theoretic achievable bit rate. Let ℓ be the distance-dependent area-mean of the channel power gain function (inversely proportional to path loss), then the signal power from a given transmitter measured at the location $u \in \mathbb{R}^2$ is given by

$$P_r(u) = P_t(x) h_{x,u} \ell(\|x - u\|) \tag{1.1}$$

where $x \in \mathbb{R}^2$ is the spatial location of the transmitter, $P_t(x)$ is the transmit power, $h_{x,u}$ is a random variable accounting for the random channel (power) gain due to small-scale fading and shadowing between the two locations x and u, and $\|\cdot\|$ is the Euclidean norm. The area-mean channel power gain function can be expressed as [20–22]

$$\ell(\|x - u\|) = \|x - u\|^{-\eta} \tag{1.2}$$

where η is the path-loss exponent depending on different environmental conditions, such as $\eta = 2$ for free space and η ranges between 1.6 and 6.5 in different cellular environments, with 3 to 4 range being more typical for outdoor channels [23]. This is called the unbounded area-mean power gain function due to its singularity at the origin. An alternative area-mean power gain function called the bounded path-loss function [24], is more practical but complicates the theoretical analysis of the system, given as $\ell(\|x - u\|) = \min\{1, \|x - u\|^{-\eta}\}$.

While the choice of the area-mean power gain function may significantly affect the statistics of the received interference power, the impact of the area-mean power gain function on the SINR is smaller. The SINR at each receiver can be calculated as

$$\text{SINR}(u) = \frac{P_t(x_0) h_{x_0,u} \ell(\|x_0 - u\|)}{\Omega + \sum_{x \in \mathcal{I}} P_t(x) h_{x,u} \ell(\|x - u\|)} \tag{1.3}$$

where u is the location of the test receiver, x_0 is the location of the transmitter that sends useful data to the receiver located at u (called the "desired" transmitter[1]), $\mathcal{I} =$

[1]Note that the subscript "0" (i.e. zero) refers to the desired transmitter, while the subscript "o" (see Section 1.4.1) refers to the origin of the plane $\mathbb{R}^2$.

$\{x_1, x_2, \cdots\}$ is the set of the locations of the interfering transmitters, and Ω is the noise power. The summation term $\sum_{x\in\mathcal{I}} \cdots$ is the cumulative (aggregate) interference power at the test receiver. At a generic time instant, the interference experienced by each receiver depends on the location of the desired transmitter, spatial distribution of the interfering transmitters as well as the instantaneous random channel gains. Therefore, given the effect of the network geometry on the interference, the SINR is a random variable that strongly depends on the network geometry and significantly varies from one receiver to another and from one time instant to another.

The SINR distribution for a network is often derived in the form of the probability of coverage. For a given fixed modulation and coding scheme and with the interference treated as noise (by using a simple linear receiver), the user is said to be in coverage if the experienced SINR exceeds a certain threshold T. Under the assumption of the desired transmitter being at a distance R from the receiver, the conditional probability (conditioned on the given distance R) of coverage p_c^R is defined as follows:

$$p_c^R(T) = \mathbb{P}(\text{SINR} \geq T) \tag{1.4}$$

Its complement $1 - p_c^R$ is the conditional outage probability, which is the same as the cumulative distribution function (CDF) of the SINR.

The mean achievable data rate (in bits/s/Hz) of this transmission link averaged over the SINR distribution can be expressed as [25, 26]

$$\mathbb{E}\left\{\log_2(1 + \text{SINR})\right\} = -\int_0^\infty \log_2(1+x)\mathrm{d}p_c^R(x) \tag{1.5}$$

assuming that the interference is treated as Gaussian noise[2].

[2]In general, almost any type of modulation, coding schemes, and receiver structure can be easily treated by adding a gap factor $G \geq 1$ to the rate expression (1.5) [23, 25], i.e., $\mathbb{E}\left\{\log_2(1 + \text{SINR}/G)\right\}$.

1.3 Stochastic Geometry Modelling and Analysis

For the analysis and design of wireless networks, rigorous yet simple and accurate models have long been desired. However, the analytical characterization of network geometry (i.e. the spatial configuration of BSs) has always been a challenging problem even in simple regular models where the BSs are placed deterministically at the center of hexagonal grids [27–30]. These models have been used extensively in academic and industrial communities, but suffer from being both highly idealized and not analytically tractable, so complex system-level Monte Carlo simulations [27] are often used to evaluate coverage probabilities and data rate. Moreover, due to the variation of the capacity demand across different service areas (such as urban and rural areas) and environmental constraints, the BSs will not exactly follow a grid-based model. Instead, realistic wireless networks are more likely to have random topologies, which makes possible the use of stochastic geometry [27, 31–33] as an efficient tool for the modelling, analysis and design of wireless networks [10–13, 21, 22, 24–26, 34–40].

Stochastic geometry is a rich branch of applied probability adapted to the study of random phenomena on the plane or in higher dimensions. It is intrinsically related to the theory of point processes [41, 42]. The theory of point processes was originally developed for applications to biology, astronomy and material sciences. Nowadays, it is also used in image analysis and in the context of communication networks.

In stochastic geometry modelling of wireless networks, the locations of the network nodes are usually seen as the realizations of some point process(es). When the underlying random point process model is assumed to be stationary and ergodic[3], the probabilistic analysis provides a convenient way of estimating spatial averages. Those averages often capture the key dependencies of the network performance characteristics (connectivity and capacity) as functions of a relatively small number of parameters such as the intensities of the underlying point processes and the parameters of

[3]If the point process is ergodic, the spatial averages (across points) equal the ensemble averages (across realizations of the point process).

the MAC protocols involved. In other words, stochastic geometry provides a natural way of defining and computing large-scale properties of such networks, by averaging over all potential geometrical patterns for the nodes, in the same way as queuing theory provides response times or congestion, averaged over all potential arrival patterns within a given parametric class. The spatial average means an empirical average made over a large collection of locations in the domain considered; depending on the cases, these locations could simply be certain points of the domain, or nodes located in the domain, or even nodes on a certain route defined in this domain. Simple examples of such averages are the fraction of nodes which transmit, the fraction of space which is covered or connected, the fraction of nodes which transmit their packet successfully, and the average geographic progress obtained by a node forwarding a packet towards some destination.

The most popular point process model is homogeneous Poisson point process (PPP), where the network nodes are uniformly and independently distributed. Due to its analytical tractability, the homogeneous PPP model has been frequently used in the analysis of cellular [21, 22, 43–48], ad hoc [49–52], and cognitive radio networks [53–56]. The results and analytical methodologies derived using the PPP model have been widely addressed and well understood, serving as the guidelines for design of wireless networks with random user and node locations. They provide answers to such questions as how the interference statistics and outage probabilities are affected by the user density and distribution, the path-loss model, the fading statistics, and power control. In turn, given the network constraints such as minimum probability of coverage or data rate requirements, they allow the tuning of the network parameters for optimum performance[4]. At the expense of reduced tractability, other point processes are also used for

[4]It is well known that considering interference as noise is not the only possible option when analyzing a wireless network. Other options (collaborative or coordinated schemes, successive cancelation techniques) can offer better rates, though at the expense of more algorithmic overhead and/or exchange of more information between network nodes. We believe that the methodology of stochastic geometry has the potential of analyzing such techniques, but for simplicity of analysis, we have only considered the case of treating interference as noise.

modelling of wireless networks, such as the binomial point process [57, 58], Poisson cluster process (PCP) [24], hard core point process (HCPP) [10, 13, 35], and Ginibre point process [36, 37].

1.4 Finding Network Measures Using Stochastic Geometry Approaches

Analytical approaches of stochastic geometry to calculate network measures usually lie in the availability of the sums, products, and the empty space distance distribution properties of the underlying point process. Using sums over a point process, the mean and variance of the interference can be evaluated. Products over a point process generally are used for the calculation of the Laplace transform of the cumulative interference, which is the key to the derivation of the probability of coverage conditioned on a fixed desired link distance R between the receiver and its desired transmitter. Finally, the overall probability of coverage averaged over the area of interest can be obtained by deconditioning the desired link distance using the empty space distance distribution of the underlying point process. The probability of coverage can then be used to find mean data rates, as discussed earlier.

1.4.1 Mean interference

Consider an independent collection of mobile users/receivers, located according to some independent stationary point process. Without any loss of generality, we assume that the mobile user under consideration is located at the origin o over the plane $\mathbb{R}^2$. The total interference is then characterized by

$$I = \sum_{x \in \Phi} h_{x,o} \ell(\|x - o\|) = \sum_{x \in \Phi} h_x \ell(\|x\|) \tag{1.6}$$

where Φ denotes the point process of interest that captures the spatial locations of all the interfering BSs with unit transmit power, $\ell(\cdot)$ is the path-loss function, and h_x denotes the random channel power gain $h_{x,o}$ between x and the origin o. In the case of a homogeneous PPP, it does not matter where the interference is measured. Due to Slivnyak's theorem [41] (see Section 2.1), it could even be measured at a point of the process as long as its contribution to I from that point is not included. In all other cases, however, the measurement location does matter, since the interference seen by a typical point of the point process differs from the interference seen at an arbitrary point of the plane (see Section 2.1 on Palm distributions [41]). From Campbell's theorem [41] (see Section 2.1), the mean interference of a network modelled by a stationary isotropic point process can be obtained by

$$I_{\mathrm{M}} = \mathbb{E}_{\Phi}(I) = \lambda \mathbb{E}(h_x) \int_{\mathbb{R}^2} \ell(\|x\|) \mathrm{d}x = \frac{2\pi\lambda}{\mu_h} \int_0^{\infty} \ell(r) r \mathrm{d}r \tag{1.7}$$

where λ is the intensity[5] of Φ. The specific fading distribution does not matter, as long as $\mathbb{E}(h_x) = 1/\mu_h$ be a constant. For finite-sized networks of radius $r_w > 1$ with a bounded path-loss model $\ell(r) = \min(1, r^{-\eta})$, (1.7) gives [56]

$$I_{\mathrm{M}}^{r_w} = \lambda \mu_h \left(\pi + \frac{2\pi}{\eta - 2} (1 - r_w^{2-\eta}) \right) \tag{1.8}$$

This simple expression offers a guideline on how large to choose the network area in a simulation, where the behaviour of an infinite network is to be explored.

1.4.2 Laplace transform of the cumulative interference

A wireless network modelled by a point process Φ consists of the desired transmitter x_0 and all other interfering transmitters. The cumulative interference measured at the origin o is the summation of the signal power from all interfering transmitters, given

[5] λ can be interpreted as the mean number of points of Φ per unit area. (See also Section 2.1.)

by

$$I_{\mathrm{c}} = \sum_{x \in \Phi \setminus x_0} h_x \ell(\|x\|) \tag{1.9}$$

Generally, there is no known closed-form expression for the probability density function (PDF) (or equivalently, the cumulative distribution function) of the cumulative interference I_{c}. The only two exceptions where the PDF of the interference has a closed-form expression can be found in [59] for deterministic channels, and in [60] for Rayleigh fading channels. Both the closed-form PDFs were obtained under the assumptions of an unbounded path-loss function with exponent $\eta = 2$ or 4, an infinite-sized PPP, and through the application of Euler's product formula [59]. Instead, I_{c} is usually characterized by its Laplace transform, given by

$$\mathcal{L}_{I_{\mathrm{c}}}(s) = \mathbb{E}\left[\exp\left(-sI_{\mathrm{c}}\right)\right] = \mathbb{E}\left[\prod_{x \in \Phi \setminus x_0} \exp\left(-sh_x \ell(\|x\|)\right)\right] \tag{1.10}$$

Stochastic geometry provides a systematic way to obtain the Laplace transform for the cumulative interference associated with the point process of interest. It is called the probability generating functional (PGFL)[6]. In the case of a homogeneous PPP Φ_{P} with intensity λ_p, for a real valued function $f(x) : \mathbb{R}^2 \to [0, 1]$, the PGFL gives

$$\mathbb{E}\left[\prod_{x \in \Phi_{\mathrm{P}}} f(x)\right] = \exp\left(-\lambda_p \int_{\mathbb{R}^2} (1 - f(x)) \mathrm{d}x\right) \tag{1.11}$$

According to Slivnyak's theorem [41], it does not matter if the desired transmitter x_0 is included in the calculation or not, i.e. $\mathbb{E}\left[\prod_{x \in \Phi_{\mathrm{P}} \setminus x_0} f(x)\right] = \mathbb{E}\left[\prod_{x \in \Phi_{\mathrm{P}}} f(x)\right]$. Thus, combining (1.10) and (1.11):

$$\mathcal{L}_{I_{\mathrm{c}}}(s) = \exp\left(-\lambda_p \int_{\mathbb{R}^2} \left(1 - \exp\left(-sh_x \ell(\|x\|)\right) \mathrm{d}x\right)\right) \tag{1.12}$$

[6]The PGFL is the random point process analogue to the probability generating function of a discrete random variable that takes non-negative integer values.

With the Laplace transform, one is able to generate the moments of the cumulative interference. Furthermore, it is the Laplace transform of the cumulative interference that gives the possibility to derive the closed-form network performance metrics such as the probability of coverage, the transmission capacity, and the average achievable data rate [21, 22].

1.4.3 Probability of coverage

Let R be the desired link distance between the receiver and its desired transmitter. In the case of Rayleigh fading, the channel power gain h is exponentially distributed. Assuming $\mathbb{E}(h) = 1/\mu_h$ and a constant transmit power P_t for all network nodes, the conditional probability of coverage p_c^R can be expressed as follows:

$$\begin{aligned} p_c^R(T) &= \mathbb{P}\left(h \geq \frac{T}{P_t \ell(r)}(\Omega + I_{\mathrm{c}})\right) \\ &\overset{(i)}{=} \exp\left(-\frac{\mu_h T \Omega}{P_t \ell(r)}\right) \mathbb{E}_{\Phi}\left[\exp\left(-\frac{\mu_h T I_{\mathrm{c}}}{P_t \ell(r)}\right)\right] \end{aligned} \tag{1.13}$$

where the last step (i) is obtained by taking the expectation with respect to both the random channel power gain h and the underlying point process Φ of interest. Note that the Rayleigh fading assumption can be relaxed at the expense of the tractability of the model. As a result, we may be able to get only approximate solutions or tight bound(s) on the SINR distribution.

Under the assumption of the closest-BS connectivity policy [21, 22], the desired link distance R is actually the distance between a generic location (not part of the point process Φ) and its closest point in the point process Φ. The desired link distance R can be further relaxed or deconditioned using the empty space function $F_R(r)$ (also known as the closest-point distance distribution or contact distance distribution) [41,42] of the point process Φ, thus an average probability of coverage p_c over the whole plane can

be obtained as

$$p_c = \int_0^\infty p_c^R(T)\mathrm{d}F_R(r) \tag{1.14}$$

In the case of the homogeneous PPP with intensity λ_p, the empty space function is $F_R(r) = 1 - \exp\left(-\lambda_p \pi r^2\right)$ [21].

1.4.4 Challenges of the stochastic geometry approach

By far, more than 95% or perhaps 99% of the analytical work on wireless network characterization is based on the Poisson model [59]. The complete spatial randomness (CSR) or independence property makes the PPP comparatively easier and tractable to analyze. In particular, when talking about its tractability, we are mainly referring to the closed-form PGFL (for calculating interference) and to the empty space distribution (for calculating the coverage probability and related metrics) of the underlying point process. Unfortunately, both of them only exist for homogeneous PPP models. Therefore, analytical tractability[7] is rarely exactly promised by other point processes to the same extent as for the PPP models. The analysis of the more general point process models requires the use of the Palm theory [25, 26] (see Section 2.1), in particular the conditional probability generating functionals, which are quite involved to calculate. Other techniques such as vulnerability region analysis (i.e. focusing on dominant interferers by region bound or nearest n interferers), approximation of the PDF of the cumulative interference, and inversion of the Laplace transform are also frequently used in the literature [8, 9, 15, 48, 61–63] but at great expense of accuracy, tractability, and practicality tradeoffs.

While stochastic geometry is powerful in theoretical analysis of wireless networks, the most critical first step is to find an accurate point process model that best reflects the spatial distribution of the network nodes. Seldom is the rationale behind the choice of

[7]This is the main advantage of stochastic analysis that leads to simple closed-form equations, and in turn, helps one to understand the behaviour of the wireless systems in response to variations in the design variables.

a particular point process for the model discussed, and rarely is there a measurement-based validation of certain stochastic assumptions used in the literature. This gives rise to questions like: When are Poisson models justified? When should one rather use point processes with some repulsion or attraction between points? When is the stationarity/isotropy assumption valid? The only aim in nearly all the related research work in the literature is to show what can be done with stochastic geometry when assumptions of this kind are made. Clearly, it may not be suitable to use exclusively the PPP as a model, if the network nodes are clustered according to certain social behavior (human activity and residential habits), or separated by some minimum distance due to geographical constraints. Stationarity should also be questioned since the distributional properties of the network nodes may not be the same over space[8]. Instead, the deployment of the network nodes is more or less a compromise of many realistic factors such as the population densities, government regulations, or terrestrial restrictions. Therefore, it is worthwhile to conduct a more extensive and comprehensive analysis of statistical modelling and validation for the spatial structure of the wireless network nodes.

1.5 Motivation and Objectives

Grid-based models are not analytically tractable and also fail to characterize the topological randomness of realistic networks. Stochastic geometry provides a natural way of defining and computing macroscopic properties of wireless networks with random topologies. It is an important tool that can analyze the random wireless networks in a probabilistic way by taking the uncertainties in the locations of network nodes and averaging over all possible network realizations, leading to simple expressions for the network performance metrics as functions of a relatively small number of point process

[8]However, one can assume that the underlying random point process is ergodic, and hence the statistical properties can be deduced simply from sufficiently large number of spatial realizations of the random process.

parameters. Stochastic geometry is an important branch of applied probability and the analysis is intrinsically related to the theory of point processes. The richness and matureness of the mathematical system offers great convenience to the exploration of the interplay between wireless communications and stochastic geometry. Motivated by the potential prospects for stochastic geometry analysis, the main objective of the thesis is to present a comprehensive technique for measurement-based statistical validation of point process models, examine the general framework of the modelling method that accounts for the network geometry, find solutions to some of the analytical challenges discussed in Section 1.4.2, and shed light on the studies of more general point process models.

1.6 Contributions of the Thesis

In this thesis, we will explore fitting of point process models to the observed real-life point patterns[9], study the empty space distance distribution in the Matérn hard core point process of Type II, and derive the downlink probability of coverage in two-tier heterogeneous cellular network using clustered point processes. The contributions of the work presented in this thesis include:

- Development of a practical technique that finds the most appropriate point process models for spatial nodes of wireless networks, from the perspective of statistical measurement and justification. Various point processes such as clustered and regular (or repulsive) models are examined and studied in terms of their own generation mechanisms and statistical properties. Covariate effects are introduced, and spatial inhomogeneity is studied.

- Fitting the empirical probability density function of the closest-point distance in the Matérn hard core point process of Type II to various existing distributions and finding that the Weibull distribution has the best goodness-of-fit. We

[9] A point pattern is a set of points, which can be regarded as a realization of a specific point process.

also propose a better piecewise probability density function for the closest-point distance, including an exact expression and a heuristic formula, which can be approximated by a Weibull-like function.

- Assuming a connected base station exists at the point of reference (or origin) in the tier of interest, we derive an expression for the downlink probability of coverage over a heterogeneous network, wherein the base station locations result from different point processes, such as Poisson point processes and Poisson cluster processes. We also derive lower and upper bounds for the coverage probability in a two-tier heterogeneous network modelled with Poisson point and cluster processes, and evaluate the effect on those bounds when changing various parameter values.

1.7 Organization of the Thesis

In Chapter 2, we present a compact overview on the fundamentals of point process theory, which will be used and extended through the whole thesis. In Chapter 3, we focus on the techniques of statistical modelling and validation for both stationary and nonstationary point process models. Chapter 4 focuses on the study of the empty space distribution for the Matérn hard core point process of Type II, and the formulation of an empirical equation for that distribution. Chapter 5 presents the derivation and analysis of the probability of coverage in heterogeneous networks modelled based on Poisson cluster processes. Finally, we conclude the work done in this thesis with a summary and discussion of future research directions in Chapter 6.

Chapter 2

Fundamentals of Point Processes

Point processes are mathematical models for random point patterns. Informally, a spatial point process Φ on the plane $\mathbb{R}^2$ is a random variable taking values in the measurable space $(\mathbb{N}, \mathcal{N})$ [41]. $\mathbb{N}$ is the family of all sequences $\phi = \{x_1, x_2, \cdots, x_n\}$ of points on $\mathbb{R}^2$ satisfying the local finiteness condition, which says that each bounded subset $\mathcal{B}$ of $\mathbb{R}^2$ contains only a finite number of points. A point process is considered to be simple if there are no duplicated points, i.e. $x_i \neq x_j$ if $i \neq j$. The order of the points x_n is of no interest, only the content of the set $\{x_n\}$ matters. Thus the symbols x_n are dummy variables and have no particular interpretation. For example, the point x_1 need not be the point closest to the origin o of the plane. The σ-algebra $\mathcal{N}$ is defined as the smallest σ-algebra on $\mathbb{N}$ to make all measurable mappings $\phi \to \phi(\mathcal{B})$, for $\mathcal{B}$ running through the bounded Borel sets. A point process Φ is said to be stationary if its characteristics are invariant under translation. That is, the processes $\Phi = \{x_n\}$ and $\Phi_x = \{x_n + x\}$ have the same distribution for all x in $\mathbb{R}^2$. Furthermore, Φ is isotropic if its characteristics are invariant under rotation. Stationarity and isotropy together yield motion-invariance [42].

2.1 Moment Measures

In the classical theory of random variables, the moments (particularly mean and variance) are important statistical measures. Point process theory has analogues to these. However, numerical mean and variance must be replaced by more complicated moment measures. Similar to the mean of a real-valued random variable, the intensity measure ς of a point process Φ is defined as

$$\varsigma(\mathcal{B}) = \mathbb{E}\{\Phi(\mathcal{B})\} \tag{2.1}$$

where $\Phi(\mathcal{B}) = \Phi \cap \mathcal{B}$. So $\varsigma(\mathcal{B})$ is the mean number of points in Borel set $\mathcal{B}$. In practice, the intensity measure is expressed in terms of a non-negative intensity function $\lambda(x)$, i.e.

$$\varsigma(\mathcal{B}) = \int_{\mathcal{B}} \lambda(x) \mathrm{d}x \tag{2.2}$$

One may interpret $\lambda(x)\mathrm{d}x$ as the probability that precisely one point falls in an infinitesimally small region containing x and of area $\mathrm{d}x$. If Φ is stationary then the intensity measure simplifies to be a multiple of the Lebesgue measure $|\mathcal{B}|$, i.e. $\varsigma(\mathcal{B}) = \lambda|\mathcal{B}|$ for some non-negative constant λ, which is called the intensity[1] of Φ.

The second-order factorial moment measure of the point process Φ is defined by [31, 33]

$$\varsigma^{(2)}(\mathcal{A}) = \mathbb{E}\left\{ \sum_{(u,v)\in\Phi}^{\neq} \mathbf{1}_{\{(u,v)\in\mathcal{A}\}} \right\} \tag{2.3}$$

where $\mathcal{A} \subset \mathbb{R}^2 \times \mathbb{R}^2$, $\neq$ means that the sum runs over all pairwise different points (u, v) in Φ, and $\mathbf{1}_{\{\cdot\}}$ is the indicator function. For many important point process models [31, 33], $\varsigma^{(2)}$ is given in terms of an explicitly known second-order product density $\varrho^{(2)}$ as

$$\varsigma^{(2)}(\mathcal{A}) = \int_{\mathcal{A}} \mathbf{1}_{\{(u,v)\in\mathcal{A}\}} \varrho^{(2)}(u, v) \mathrm{d}u\mathrm{d}v \tag{2.4}$$

[1] λ can be interpreted as the mean number of points of Φ per unit area of $\mathcal{B}$.

The term $\varrho^{(2)}(u,v)\mathrm{d}u\mathrm{d}v$ can be interpreted as the probability of observing a point in each of two infinitesimally small areas $\mathrm{d}u$ and $\mathrm{d}v$ and respectively containing u and v.

If Φ is stationary then $\varrho^{(2)}$ depends only on the difference of its arguments, and if furthermore Φ is motion-invariant then $\varrho^{(2)}$ depends only on the distance r between u and v and it is simply written as $\varrho^{(2)}(r)$ [42].

Without using the second-order product density $\varrho^{(2)}$, the second-order factorial moment measure can be expressed in terms of the second-order reduced moment measure $\mathcal{K}$, given by [31, 33]

$$\begin{aligned}\varsigma^{(2)}(\mathcal{B}_1 \times \mathcal{B}_2) &= \lambda^2 \int_{\mathcal{B}_1} \mathcal{K}(\mathcal{B}_2 - x)\mathrm{d}x \\ &= \lambda^2 \int_{\mathbb{R}^2}\int_{\mathbb{R}^2} \mathbf{1}_{\{x\in\mathcal{B}_1\}}\mathbf{1}_{\{(x+v)\in\mathcal{B}_2\}}\mathcal{K}(\mathrm{d}v)\mathrm{d}x\end{aligned} \tag{2.5}$$

The product $\lambda\mathcal{K}$ can be interpreted as the mean number of points in $\mathcal{B}\setminus o$ under the condition that there is a point of Φ at o, but not counting it. The exact definition of $\mathcal{K}$ uses the theory of Palm distributions [33], and $\mathcal{K}$ plays some role in point process statistics in the context of directional analysis [31].

If a second-order product density $\varrho^{(2)}$ exists, then there is the following relationship between $\varrho^{(2)}$ and $\mathcal{K}$:

$$\lambda^2\mathcal{K}(\mathcal{B}) = \int_{\mathcal{B}} \varrho^{(2)}(x)\mathrm{d}x \tag{2.6}$$

The description of the second-order moment measure simplifies even further in the motion-invariant case. It then suffices to consider the second reduced moment function or K function defined as [42]

$$K(r) = \mathcal{K}(B(o,r)) \tag{2.7}$$

where $B(o,r)$ is the circle centered at o with radius r. The quantity $\lambda K(r)$ is the mean number of points of Φ within a circle of radius r centered at the typical point at o, which is itself not counted.

Theoretical analysis of stationary point processes typically is carried out by taking advantage of some useful mathematical tools such as Campbell's theorem [31, 33], the Palm distribution [31, 33] and so on.

Campbell's theorem. Let $f(x) : \mathbb{R}^2 \to \mathbb{R}^+$ be a measurable integrable function, then the expectation of the sum $f(x)$ over the point process Φ is given by the following equation:

$$\mathbb{E}\left[\sum_{x\subset\Phi} f(x)\right] = \int_{\mathbb{R}^2} f(x)\varsigma(\mathrm{d}x) \tag{2.8}$$

Palm distribution. Simply speaking, a Palm distribution is the conditional distribution of a point process under the assumption that there are points of the point process known at a given set of locations. Moreover, in many cases of practical interest we assume a given point at the origin o of the plane; this motivates the interpretation of the Palm distribution $\mathbb{P}_o$ as the corresponding conditional distribution. Another important definition is the reduced Palm distribution $\mathbb{P}_o^!$, which means that the given point is at the origin but is not counted in the distribution. A very popular result of the Poisson point process is the Theorem of Slivnyak [31, 33].

Theorem of Slivnyak: The reduced Palm distribution of a Poisson point process is equivalent to the original distribution, that is

$$\mathbb{P}_o^! = \mathbb{P} \tag{2.9}$$

Hence, a point can be added into or deleted from a Poisson point process while not disturbing the distribution of the original Poisson point process.

2.2 Statistics for Point Processes

The K function is often used in point process statistics and it has many interesting applications. The K function is also closely related to the pair correlation function

[42]. For a motion-invariant point process, the pair correlation function $g(r)$ satisfies

$$g(r) = \frac{1}{2\pi r}\frac{\mathrm{d}K(r)}{\mathrm{d}r} \tag{2.10}$$

It also results from normalization

$$g(r) = \varrho^{(2)}(r)/\lambda^2 \tag{2.11}$$

The forms of these functions correspond to different spatial properties of the underlying point process. Maxima of $g(r)$, or values of $K(r)$ larger than πr^2 for r in specific intervals, indicate frequent occurrences of attraction at such r; minima of $g(r)$ or low values of $K(r)$ indicate inhibition at these distances r.

Other statistics than $K(r)$ and $g(r)$ for investigating interpoint interaction are distance methods, including the empty space function $F(r)$, nearest neighbour distance distribution function $G(r)$, and the second-order characteristics $L(r)$ and $J(r)$. These functional summary statistics are based on distances such as: (1) pairwise distances $s_{ij} = \|x_i - x_j\|$ between all distinct pairs of points x_i and x_j ($i \neq j$) in the point pattern ϕ; (2) nearest neighbour distances $t_i = \min_{j \neq i} s_{ij}$, the distance from each point x_i to its nearest neighbour; (3) empty space distances $\mathrm{d}(u) = \min_i \|u - x_i\|$, the distance from a fixed reference location u in the study region to the nearest point in ϕ.

Assuming Φ is stationary (statistically invariant under translations), we can define the cumulative distribution function of the empty space distance (empty space function) as follows

$$F(r) = \mathbb{P}\left\{\mathrm{d}(u, \Phi) \leq r\right\} \tag{2.12}$$

where $u \in \mathbb{R}^2$ is an arbitrary reference location. If the point process is stationary then this definition does not depend on u.

Assuming the point process Φ is stationary, we can define the cumulative distribution function of the nearest-neighbour distance (nearest neighbour distance distribution

function) for a typical point x in the point process as

$$G(r) = \mathbb{P}\left\{\mathrm{d}(x, \Phi \setminus x) \leq r | x \in \Phi\right\} \tag{2.13}$$

where $\mathrm{d}(x, \Phi \setminus x)$ is the shortest distance from the point x to the point pattern Φ excluding x itself. If the point process is stationary then this definition does not depend on x.

The function $L(r)$ is defined as

$$L(r) = \sqrt{K(r)/\pi} \tag{2.14}$$

and the function $J(r)$ is defined as

$$J(r) = \frac{1 - G(r)}{1 - F(r)} \tag{2.15}$$

All the above summary statistics are powerful in the characterization of a point process. Which one is to be used depends mainly on convenience, but also on the statistical considerations of a certain scenario. Model identification (or justification) and validation may be suggested by parameter estimation through the comparison of empirical summary statistics with model-related theoretical counterparts, which may help one in statistical applications to find a good model for the observed point pattern data. Note that the functional summary statistics above are defined and estimated under the assumption that the point process is stationary (homogeneous), as implied by their use of the distance variable r between points [64, 65].

For nonstationary point processes, a generalized analogue to the K function is called the inhomogeneous K function $K_{in}(r)$ [66]. Suppose Φ is a point process with nonconstant intensity function $\lambda(u)$ at each arbitrary location $u \in \mathbb{R}^2$. The inhomogeneous K function $K_{in}(r)$ is defined to be the expected value of the sum of all terms $1/\lambda(x_j)$ over all points x_j in the process separated from u by a distance less than r,

given that a point of Φ is at the location u. This reduces to the ordinary K function if $\lambda(\cdot)$ is constant.

Given an observed point pattern ϕ in the study region W, $K_{in}(r)$ can be estimated by summing the values $\frac{1}{\lambda(x_i)\lambda(x_j)}$ for all pairs of points x_i, x_j separated by a distance less than r [66]

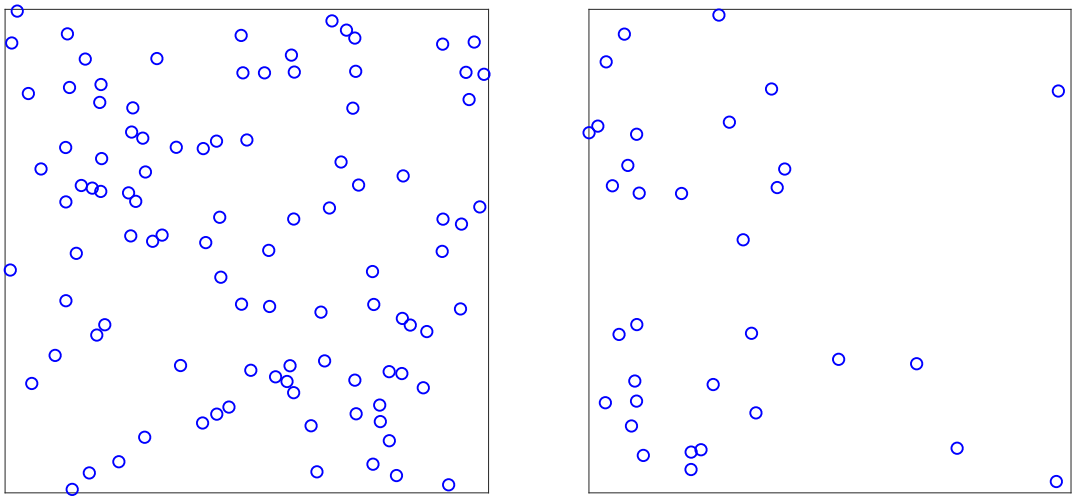
$$K_{in}(r) = \frac{1}{|W|} \sum_{\substack{x_i, x_j \in \phi \\ x_i \neq x_j}} \frac{\mathrm{e}(x_i, x_j, r) \cdot \mathbf{1}_{\{\mathrm{d}[x_i, x_j] \leq r\}}}{\lambda(x_i)\lambda(x_j)} \tag{2.16}$$

where $|W|$ is a constant denominator, $\mathbf{1}_{\{\cdot\}}$ is the indicator function, $\mathrm{d}[x_i, x_j]$ is the distance between points x_i and x_j, and $\mathrm{e}(x_i, x_j, r)$ is an edge correction factor such as border correction, modified border correction, translation correction or isotropic correction [66, 67]. For isotropic correction, the edge correction factor $\mathrm{e}(x_i, x_j, r)$ can be expressed as [66]

$$\mathrm{e}^{-1}(x_i, x_j, r) = |W| \cdot \mathrm{h}(x_i, x_j) \tag{2.17}$$

where $\mathrm{h}(x_i, x_j)$ is fraction of the circumference of the circle with center x_i and radius $\mathrm{d}[x_i, x_j]$ which lies inside the region W.

One use of the K function (both normal and inhomogeneous versions) is for determining if a point process exhibits clustering and/or repulsion among points. In particular, application of the inhomogeneous K function allows one to inspect an observed point pattern for evidence of interpoint interactions, after allowing for spatial inhomogeneity of such point pattern data. Values of $K(r)$ or $K_{in}(r)$ greater than πr^2 suggest a spatial trend of clustering, whereas values less than πr^2 suggest a spatial trend of repulsion.



(a) (b)

Figure 2.1. (a) Realization from a homogeneous PPP with intensity 100 in the unit area; (b) Realization from an inhomogeneous PPP with intensity function $100 \cdot \exp(-3x)$ in the unit area.

2.3 Point Process Models

There is a huge collection of point processes described in [41,42]. Generally, the point processes can be categorized into three main different kinds: Poisson, clustered, and regular. Each specific point process model can be either stationary (with a constant intensity λ) or nonstationary (with a variable intensity $\lambda(\cdot)$). In this section, we will introduce some of the most common point process models in terms of their generation mechanisms and statistical properties.

2.3.1 Poisson point processes

The Poisson point process is the simplest and most important model for random point patterns. It plays a central role as a null model, being the starting point for the construction of many other stochastic models. It is the model for complete spatial randomness (CSR). The location of any point of a PPP is independent of the location of any other point of the process. A stationary (homogeneous) Poisson point process Φ_{P} is charac-

terized by two fundamental properties: (a) the number of points of Φ_{P} in a bounded Borel set $\mathcal{B}$ has a Poisson distribution with mean $\lambda|\mathcal{B}|$ for some constant intensity λ_p; (b) the numbers of points of Φ_{P} in k disjoint Borel sets form k independent random variables, for an arbitrary k. The points of Φ_{P} are uniformly randomly distributed over the area of $\mathcal{B}$.

The nonstationary (or inhomogeneous) Poisson point process $\Phi_{\mathrm{P}}^{\mathrm{in}}$ with intensity function $\lambda(\cdot)$ is however defined by: (a) the number of points of $\Phi_{\mathrm{P}}^{\mathrm{in}}$ in any bounded Borel set $\mathcal{B}$, denoted by $N\left(\Phi_{\mathrm{P}}^{\mathrm{in}}(\mathcal{B})\right)$, is Poisson distributed with mean $\int_{\mathcal{B}} \lambda(u)\mathrm{d}u$, $u \in \mathbb{R}^2$; (b) conditional on $N\left(\Phi_{\mathrm{P}}^{\mathrm{in}}(\mathcal{B})\right) = n$, the n points are independent and identically distributed in $\mathcal{B}$ with density proportional to $\lambda(\cdot)$. The numbers of points in disjoint Borel sets are still independent.

The K function for both stationary and nonstationary PPPs is the same: $K(r) = K_{in}(r) = \pi r^2$. Realizations of a stationary and nonstationary Poisson point process are depicted in Fig. 2.1.

2.3.2 Clustered point processes

Clustering means that some form of attraction exists between points, leading to point patterns aggregated in space. Typical models include the Matérn cluster process (MCP) and the Thomas cluster process (TCP) [41, 42].

A stationary *Matérn cluster process* is formed by taking a pattern of parent locations, generated according to a Poisson process with intensity κ, and around each parent location, generating a random number of offspring points. (Parent locations are not themselves points.) The number of offspring points of each parent location is a Poisson random variable with mean μ, and the offspring points are independently and uniformly distributed inside a circle with radius R_{MC} centered on the associated parent location. The K function of the Matérn cluster process is given by [42]

$$K(r) = \pi r^2 + \ell_1\left(\frac{r}{2R_{MC}}\right)/\kappa \tag{2.18}$$

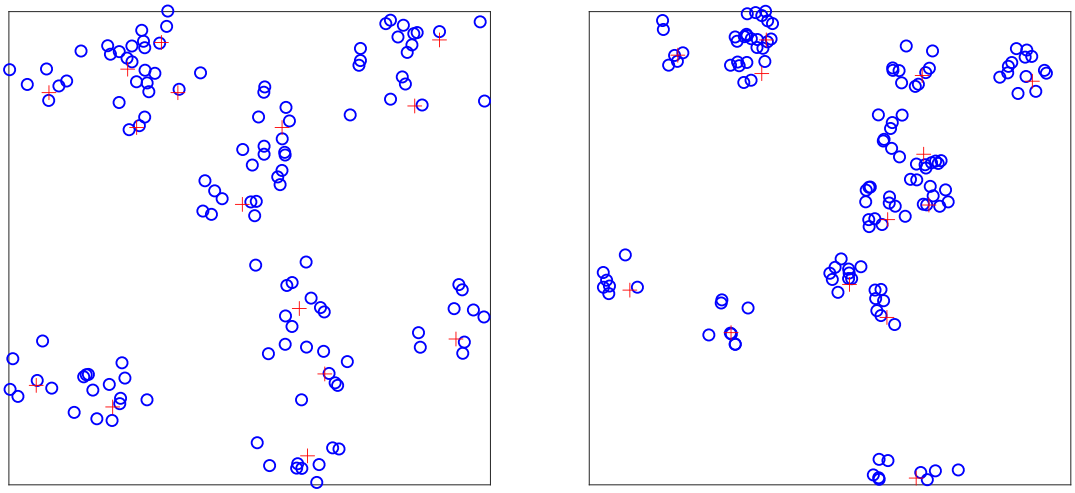


Figure 2.2. Realizations of clustered point processes with parent intensity $\kappa = 10$ and offspring intensity $\mu = 10$ in the unit area (red plus symbols denote the parent locations which are themselves not included). (a) Stationary Matérn cluster process with $R_{MC} = 0.05$; (b) Stationary Thomas cluster process with $\sigma = 0.05$.

where

$$\ell_1(z) = 2 + \frac{1}{\pi}\left(\left(8z^2 - 4\right)\arccos(z) - 2\arcsin(z) + 4z\sqrt{\left(1-z^2\right)^3} - 6z\sqrt{1-z^2} \right) \tag{2.19}$$

for $z \leq 1$, and $\ell_1(z) = 1$ for $z > 1$. The intensity of the Matérn cluster process is $\lambda = \kappa \cdot \mu$.

A stationary *Thomas cluster process* is similar to a Matérn cluster process. The key difference is that the offspring points, rather than being uniformly distributed within a circle, are instead independently and isotropically normally distributed around the parent location with some standard deviation σ. The K function of the Thomas cluster process is given by [42]

$$K(r) = \pi r^2 + \frac{1}{\kappa}\left(1 - \exp\left(-\frac{r^2}{4\sigma^2}\right)\right) \tag{2.20}$$

The theoretical intensity of the Thomas cluster process is $\lambda = \kappa \cdot \mu$. Realizations of a stationary Matérn and Thomas cluster process are depicted in Fig. 2.2.

For nonstationary Matérn and Thomas cluster processes, the intensity function is no longer a constant, but depends on the location of interest. A generalized non- and semi-parametric estimation is still available for inhomogeneous point patterns using the inhomogeneous K function described earlier at the end of Section 2.2.

2.3.3 Regular point processes

Regularity stands for some form of repulsion between points in a point process. Among regular point processes, *Gibbs point processes* make up a large share of them [31, 33]. Gibbs processes are a family of pairwise interaction models, which are generally constructed by specifying their probability densities. Pairwise interaction models have probability densities of the form [68]

$$f(\phi) = \mathrm{w} \left[\prod_{i=1}^{n(\phi)} \mathrm{b}(x_i) \right] \left[\prod_{i<j} \mathrm{c}(x_i, x_j) \right] \tag{2.21}$$

where $\phi = \{x_1, x_2, \cdots, x_n\}$ represent the points of the point pattern, $n(\phi)$ denotes the number of points in ϕ, w is a normalizing constant with regard to point process Φ, $\mathrm{b}(x_i)$ is the first-order term, and $\mathrm{c}(x_i, x_j)$ is the second-order or pairwise interaction term, which must be symmetric, i.e. $\mathrm{c}(x_i, x_j) = \mathrm{c}(x_j, x_i)$. In principle, $\mathrm{b}(\cdot)$ and $\mathrm{c}(\cdot, \cdot)$ may be arbitrary functions provided the resulting probability density is integrable with respect to the unit rate Poisson process, which has intensity $\lambda_p = 1$. Generally, fitting and simulation of Gibbs point processes are implemented using the Monte Carlo method [67]. Depending on different assignments for the first- and second-order terms, there are many different kinds of pairwise interaction models, leading to special cases of Gibbs point processes such as the Strauss point process and the hard core point process.

The stationary *Strauss point process* with interaction radius R_{S} and parameters β_{S} and γ_{S} is the pairwise interaction point process, in which each point contributes a factor β_{S} to the probability density of the point pattern, and each pair of points closer than R_{S} apart contributes a factor γ_{S} to the density. The probability density is given by [68]

$$f(\phi) = \mathrm{w}_{\mathrm{S}} \beta_{\mathrm{S}}^{n(\phi)} \gamma_{\mathrm{S}}^{\mathrm{s}(\phi, R_{\mathrm{S}})} \tag{2.22}$$

where w_{S} is a normalizing constant, $\beta_{\mathrm{S}} > 0$ is the intensity (or abundance) parameter, $n(\phi)$ is the number of points in the pattern, $\gamma_{\mathrm{S}} \in [0, 1]$ is the second-order or interaction term and $\mathrm{s}(\phi, R_{\mathrm{S}})$ is the number of distinct unordered pairs of points that are closer than the interaction radius R_{S} apart. The parameter γ_{S} controls the strength of interaction between points. If $\gamma_{\mathrm{S}} = 1$, the model reduces to a Poisson point process with intensity β_{S}. For values $0 < \gamma_{\mathrm{S}} < 1$, the process is regarded as a soft core process, also exhibiting inhibition (or repulsion) between points. If $\gamma_{\mathrm{S}} = 0$, the model is a hard core point process. An exact tractable analytical form for the K function of the stationary Strauss point process is not available. However, an approximate expression for $K(r)$ is [31, 33]:

$$K(r) = \begin{cases} \gamma_{\mathrm{S}} \pi r^2, & 0 < r \leq R_{\mathrm{S}}, \\ \pi r^2 - (1 - \gamma_{\mathrm{S}}) \pi R_{\mathrm{S}}^2, & r > R_{\mathrm{S}}. \end{cases} \tag{2.23}$$

A stationary *hard core point process* with the hard core distance R_{H} and intensity parameter β_{H} is a pairwise interaction point process, in which distinct points are not allowed to come closer than a distance R_{H} apart. The probability density is zero, if any pair of points is closer than R_{H} units apart, and otherwise equals [68]

$$f(\phi) = \mathrm{w}_{\mathrm{H}} \cdot \beta_{\mathrm{H}}^{n(\phi)} \tag{2.24}$$

Realizations of the stationary Strauss and hard core point processes are depicted in Fig. 2.3.

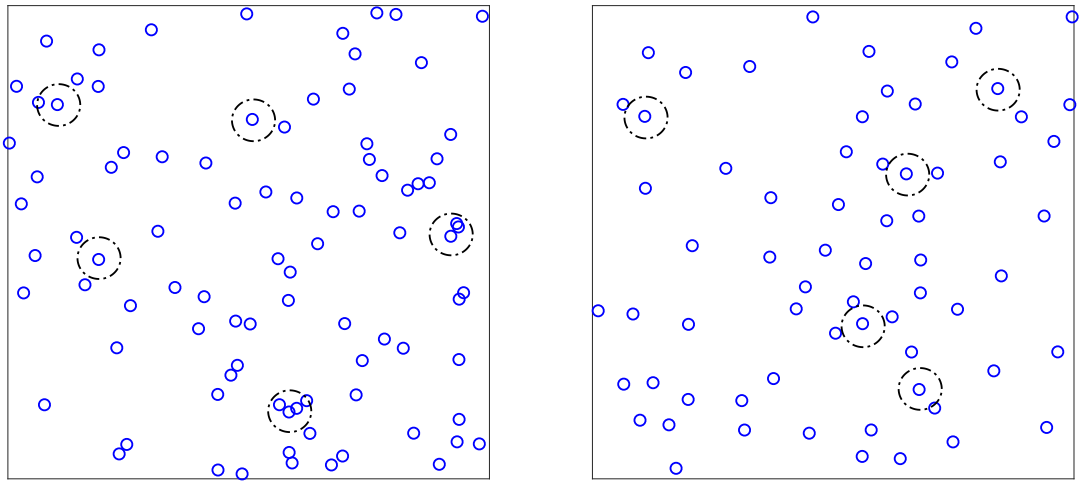


Figure 2.3. (a) Realization from a stationary Strauss point process with $\beta_{\mathrm{S}} = 100$, $\gamma_{\mathrm{S}} = 0.7$, and $R_{\mathrm{S}} = 0.05$ in the unit area; (b) Realization from a stationary hard core point process with $\beta_{\mathrm{H}} = 100$ and $R_{\mathrm{H}} = 0.05$ in the unit area.

For nonstationary Strauss and hard core point processes, the intensity parameter β_{S} or β_{H} is no longer a constant, but depends on the location of interest, which hence results in locally different intensities of points.

There are other types of regular or repulsive point processes, such as Matérn hard core point process of Type II [32] (which we shall call MHCPP for shorthand)[2]. It is constructed by applying dependent thinning to a parent PPP. That is, starting from a parent PPP Φ_{P}, the MHCPP Φ_{MH} is obtained by assigning a random mark uniformly distributed in $[0, 1]$ to each parent point of Φ_{P}, then deleting all the points that coexist within a distance less than the hard-core parameter δ from another point with a lower mark. Hence, only the parent points that have the lowest mark within their neighbourhood distance δ are retained. The MHCPP intensity can be expressed as [32]

$$\lambda_{\mathrm{MH}} = \frac{1 - \exp\{-\lambda_p \pi \delta^2\}}{\pi \delta^2} \tag{2.25}$$

[2]There also exist Matérn hard core point processes of Type I and Type III [32], which differ in how the points are thinned from the parent PPP. We do not consider the Type I and Type III processes in this thesis.

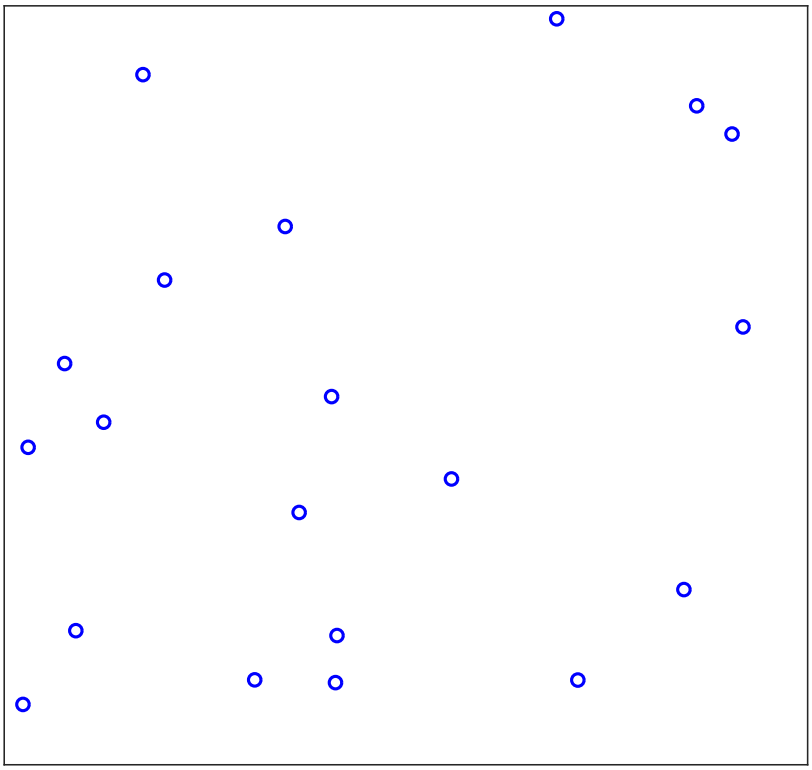


Figure 2.4. Realization from a stationary Matérn hard core point process of Type II with $\lambda_{\text{MH}} = 20$ and $\delta = 0.05$ in the unit area.

where λ_p is the parent PPP intensity and δ is the hard-core parameter[3]. A realization of a stationary MHCPP is depicted in Fig. 2.4.

[3] δ may be in any arbitrary unit of distance, while λ_p and λ_{MH} would be in the same units of distance^{-2}.

Chapter 3

Modelling of Cellular Networks Using Stationary and Nonstationary Point Processes

While stochastic geometry has its strength in theoretical analysis, a measurement-based validation of certain stochastic assumptions is often not given. The aim in most research work is only to show what results can be obtained with stochastic geometry when assumptions of certain point process models are made, without necessarily fully justifying those assumptions. This invites questions like when are Poisson models justified, when should one rather use clustered or repulsive models, and when is the stationary/isotropic assumption valid.

In order to fully characterize the spatial topology of the base stations in cellular networks and identify the most proper and accurate point process model for a given BS point pattern, in this chapter, we present and describe a network-data-supported technique for fitting stationary and nonstationary point process models to real-life cellular networks using maximum likelihood/pseudolikelihood and minimum contrast methods. Nonstationary processes are of particular interest, since real-life wireless networks most often do not have a homogeneous spatial distribution. When fitting with

nonstationary models, both spatial inhomogeneity and covariate effects are considered. We introduce covariates into the point process models as a potential (or secondary) effect that further influences the distribution of wireless nodes, in order to bridge the gaps between the results and measures of stationary models and simulations of real-life cellular networks. The covariates considered account for population densities in urban areas and distance from the base stations to their closest main roads in rural areas. Simulated envelope tests are used for the evaluation of goodness-of-fit. However, such envelope tests are insufficient to conclusively distinguish between the fitted models. Thus we apply other metrics such as the Akaike information criterion and root mean square deviation to differentiate among different fitted models. Additionally, the probability of coverage is also considered as the supplementary criterion for the goodness-of-fit and model selection.

3.1 Introduction

The spatial configuration of base stations (BSs) has a great impact on the performance analysis of cellular networks since received signals and interference mostly depend on the distances between transmitters and receivers. A regular hexagonal grid has long been used as an idealized model for the cellular network layout [27–30]. However, due to the variation of the capacity demand across different service areas (such as urban and rural areas) and environmental constraints, the BSs will not exactly follow a grid-based model. Instead, realistic wireless networks are more likely to have random topologies, which makes possible the use of stochastic geometry as an efficient tool for the modelling, analysis and design of wireless networks [21, 22, 25, 26].

Stochastic geometry [31, 33] is a mathematical study of random spatial patterns that leads to averaging over all network spatial realizations seen from a generic node, weighted by their probabilities of occurrence, for the quantities of interest such as interference, the coverage probability, and average data rate. In stochastic geometry

analysis, the nodes in a wireless network are usually modelled by a convenient point process [41, 42], which reflects their spatial locations. The most popular models are Poisson point processes, which have been frequently used for the analysis of cellular, ad hoc, and cognitive radio networks [34]. The complete spatial randomness or independence property of the points makes the networks modelled by a PPP easy to analyze. However, it may not be suitable to use exclusively the PPP as a model, if the BSs are clustered according to certain social behavior (human activity and residential habits), or separated by some minimum distance due to geographical constraints. Therefore, it is worthwhile to conduct a more extensive and comprehensive analysis by taking more realistic point process models into consideration.

In [24], Poisson cluster processes are used to model wireless ad hoc networks. Distributional properties of the interference and a numerically integrable expression for the outage probability are derived. In [35], Matérn hard core point processes of Type I and Type II are used to model concurrent transmitters in carrier sensing multiple access (CSMA) networks. The mean interference is determined and compared with a PPP model of the same intensity. In [13], a modified Matérn hard core point process model is proposed for the analysis of random CSMA wireless networks in general fading environments. In [36] and [37], Ginibre point processes and determinantal point processes have been investigated as suitable models for wireless networks with spatial regularity. The α-stable distribution [38] has been shown to be an accurate model for characterizing the statistical nature of BS density in cellular networks. In [39] and [40], various repulsive and clustered point processes, such as Strauss and Cox processes, are used for model-fitting of realistic BS deployments.

The BS locations can be interpreted as points, with the set of points forming a spatial point pattern. That pattern can be regarded as a realization of a specific point process. Finding an accurate point process model that is the most likely to yield such a point pattern realization is usually a two-step undertaking. The first step involves determining how the candidate point process models should be selected to be fitted to

the point pattern from some real-life data. The second step is selection and calculation of the statistical metrics used for goodness-of-fit testing, including classic functionary summary statistics (such as the empty space distribution function F and the K function [31, 33]) and performance metrics more relevant to applications such as the coverage probability or average data rate [39]. If a point process model can be found that fits the observed data pattern well, the estimated values of its parameters provide summary statistics that can be used to compare ostensibly similar data sets. More ambitiously, a fitted model can provide an explanation of the underlying physical processes [21, 22].

In this chapter, we aim to describe a network-data-supported statistical technique that finds the most appropriate point process models for the observed BS point patterns obtained from some real-life cellular networks. This technique is based on the standard statistical $\mathbf{R}$ package $\mathrm{spatstat}$ [67], which is designed for analyzing realistic spatial point pattern data. The fitting methods implemented are generally divided into two categories: maximum likelihood/pseudolikelihood and minimum contrast. Both can be applied to stationary and nonstationary point process models. When the observed (BS) point pattern is judged to be spatially homogeneous, stationary point process models can be used for the statistical analysis using classical methods similar to [39,40, 67, 68]. However, real cellular networks are spatially inhomogeneous, with, e.g., very different densities in rural and urban areas, driven by geographic differences in demand and deployment constraints. Standard stochastic geometry methods do not model this inhomogeneity, so must be applied only in homogeneous regions of a network, and tuned separately for each one.

As the main contribution of this chapter, we generalize stochastic geometry methods to inhomogeneous real networks by introducing additional input variables – which are referred to as covariates[1] – to our stochastic geometry models. The covariates help account for how the base station distribution changes over the region of interest. In this chapter, we consider two different kinds of covariate, which respectively account for

[1] Informally, the covariate in statistics is a possible predictive or explanatory variable of the dependent variable in a statistical experiment.

the population densities and for the distance from the BS to its closest main road (modelled as a line segment), and fit the nonstationary point process models to the BS point pattern observed in different areas (urban and rural). To the best of our knowledge, this is the first time that nonstationary point processes with dependence on a covariate have been used for model-fitting of cellular networks.

A simulated envelope test is used to identify point process models that fit the observed data. The selection of the best model out of these candidates is based on other statistics such as the Akaike information criterion (AIC) [69, 70] and root mean square deviation (RMSD) [70], or more relevant wireless metrics such as the probability of coverage. This validation aspect is essential. Some works in the literature provide results assuming a certain point process model for the wireless network, without also doing a quantitative evaluation of whether the model is a match to an actual network layout.

3.2 Methodology

Based on the types of candidate point process models, there are two common methods for model-fitting that are implemented in the **R** package spatstat. When the candidate model is either a Poisson point process or a regular point process (such as Gibbs point processes), the maximum likelihood/pseudolikelihood method based on the conditional intensity can be applied for fitting the model to the observed point pattern data. When the candidate model is a clustered point process, the minimum contrast method based on functional summary statistics is better for model-fitting. Both methods can be applied to either stationary or nonstationary point process models.

To evaluate the goodness-of-fit of a fitted model, a simulated envelope test [67] based on various functional summary statistics (such as the K function) can be used. If more than one candidate model satisfies the requirement of an envelope test, then other criteria or metrics should be used to determine the best fitted model out of all

the candidate models, such as the Akaike information criterion (AIC) [69] and the root mean square deviation (RMSD) [70].

3.2.1 Conditional intensity of point processes

The (Papangelou) conditional intensity [67], denoted by $\lambda(u, \phi)$, is a function of both an arbitrary spatial location u and the observed point pattern ϕ, within the study region W. Informally, the conditional probability of finding a point of the point process inside an infinitesimal neighbourhood of the location u, given the locations of all other points outside this infinitesimal region, is $\lambda(u, \phi)\mathrm{d}u$. Given an observed configuration ϕ in a bounded region, the conditional intensity at a location $u \notin \phi$ for a point process is related to the probability density by

$$\lambda(u, \phi) = \begin{cases} f\left(\phi \cup \{u\}\right) / f\left(\phi\right), & \text{if } f\left(\phi\right) > 0 \\ 0, & \text{if } f\left(\phi\right) = 0 \end{cases} \tag{3.1}$$

In other words, it is the ratio of the probability densities for the configuration ϕ with and without the point u added. A point process is attractive if $\lambda(u, \phi_1) \leq \lambda(u, \phi_2)$, and repulsive if $\lambda(u, \phi_1) \geq \lambda(u, \phi_2)$, whenever $\phi_1 \subset \phi_2$. For the general pairwise interaction process, the conditional intensity is given by [67]

$$\lambda(u, \phi) = \mathrm{b}(u) \prod_{i=1}^{n(\phi)} \mathrm{c}(u, x_i) \tag{3.2}$$

where $\mathrm{b}(u)$ is the first-order term and $\mathrm{c}(u, x_i)$ is the second-order term that describes the interaction between a point at u and x_i. For example, the stationary Strauss process has conditional intensity $\lambda(u, \phi) = \beta_{\mathrm{S}} \gamma_{\mathrm{S}}^{\mathrm{t}(u,\phi)}$, where $\mathrm{t}(u, \phi)$ is the number of points of the given point pattern ϕ that lie within a distance R_{S} from the location u. The Poisson process with intensity function $\lambda(u)$ has conditional intensity $\lambda(u, \phi) = \lambda(u)$, because the points of a Poisson process are mutually independent.

In spatstat, in order to implement the algorithms for the calculation of Poisson regression models, the conditional intensity is split into first-order and second-order terms and must be loglinear[2] in the parameters ϑ [67]. This gives

$$\log \lambda_\vartheta(u, \phi) = \psi \cdot S(u) + \xi \cdot V(u, \phi) \tag{3.3}$$

where $\vartheta = (\psi, \xi)$ are often called regular parameters. $S(u)$ is the first-order term that specifies the spatial trend of the underlying model. If $S(u)$ is a constant, then the point process model is stationary. If $S(u)$ is a function of the spatial coordinates of u, or an observed covariate, or a mixture of both, then the model is nonstationary. $V(u, \phi)$ is the second-order term that describes the interpoint interaction. If $V(u, \phi)$ is absent (or set to be zero), then the model is a Poisson point process. For example, the conditional intensity of a stationary Strauss process can be recast as $\log \lambda(u, \phi) = \log \beta_{\mathrm{S}} + (\log \gamma_{\mathrm{S}})\, \mathrm{t}(u, \phi)$, so that $\vartheta = (\log \beta_{\mathrm{S}}, \log \gamma_{\mathrm{S}})$.

3.2.2 Specification for the spatial trend

In the inhomogeneous case, it is important to specify an intensity function for the fitted model in the first place. Generally, the intensity function of a nonstationary point process can be formulated in terms of spatial coordinates and/or an observed covariate. The relation between the intensity function $\lambda(u)$ and the spatial coordinates is predicted non-parametrically through a so-called kernel-smoothed estimate for the observed point pattern ϕ. The usual kernel estimator of the intensity $\lambda(u)$ is [68]

$$\hat{\lambda}(u) = \hat{\mathrm{e}}(u) \sum_{i=1}^{n(\phi)} \mathrm{k}(u - x_i) \tag{3.4}$$

[2]This constraint on the form of the conditional intensity is imposed by the spatstat package for its calculations; the conditional intensity need not be restricted to this form in general.

where $\mathrm{k}(u)$ is the kernel (an arbitrary probability density function), x_i is the point of ϕ, and

$$\hat{\mathrm{e}}^{-1}(u) = \int_W \mathrm{k}(u-v)\mathrm{d}v \tag{3.5}$$

is an edge effect bias correction. By Campbell's theorem [25], one can find that $\hat{\lambda}(u)$ is an unbiased estimator of $\tilde{\lambda}(u) = \hat{\mathrm{e}}(u) \int_W \mathrm{k}(u-v)\lambda(v)\mathrm{d}v$, which is a smoothed version of the true intensity function $\lambda(u)$. By investigating and understanding the distributional variations of the intensities at different locations, a wide variety of models (forms) for the intensity function $\lambda(\cdot)$ can be constructed in the Cartesian coordinates. Since the true form of the intensity is generally not known, the selection of which intensity function to use is more of an art than a direct scientific derivation, and this often leads to inappropriate assumptions being made for a proper intensity function form.

A point pattern data set may also include covariate information. Similar to the Cartesian coordinates (independent variables) of the intensity function (dependent variable), a covariate is another kind of variable that is measurable and considered to have a statistical relationship with the dependent variable. Covariates may not be of primary interest compared to the independent variables when evaluating the intensity function or intensity measure of a point process model. They arise because the observed point pattern is heterogeneous. One common type of covariate information is a spatial function $Z(u)$ defined at all spatial locations $u \in W$ (i.e. the entire study region and not only at the observed data points). A second common type is another spatial point pattern, or a line segment pattern, e.g. the locations of some fixed objects of interest in the study region. This covariate pattern would be used to define a surrogate spatial function Z as in the first type above. For example, $Z(u)$ may be the distance from u to the fixed covariate locations. In $\mathrm{spatstat}$, a covariate can be applied to improve the accuracy of the schemes implemented for model-fitting because the covariate provides more information for "dummy points" (some other locations than the observed points in the study region). For example, given a data set containing a targeted point pattern ϕ_X, and another covariate point pattern ϕ_Y which may account for some control factors, if one

wants to model ϕ_X as a point process with an intensity proportional to the local intensity of ϕ_Y, then one may formulate the intensity of ϕ_X as $\lambda_X(u) = \exp(\alpha_0 + \alpha_1 Z(u))$, where (α_0, α_1) are linear coefficients and $Z(u)$ is the local intensity value of the point pattern ϕ_Y at the location u. Note that one can use both the Cartesian coordinates and covariate function $Z(u)$ to specify a fitted intensity function of an arbitrary expression, like $\lambda(x_u, y_u) = \exp\left\{\alpha_0 x_u^2 + \alpha_1 \sqrt{Z(x_u, y_u)}\right\}$, where (x_u, y_u) are the spatial coordinates of the point at location u.

3.2.3 Maximum likelihood/pseudolikelihood method

The likelihood function plays a fundamental role in classical approaches to statistical inference [71]. In particular, the homogeneous Poisson point process with intensity λ_p has density

$$f(\phi) = \exp\{(1 - \lambda_p)|W|\}\lambda_p^{n(\phi)} \tag{3.6}$$

where ϕ is the observed point pattern, and $n(\phi)$ is the number of points in the bounded study region W with an area of $|W|$. A maximum likelihood estimator of λ_p can be easily derived as $\hat{\lambda}_p = n(\phi)/|W|$.

Maximum likelihood estimation is in general very difficult and notoriously intractable for most spatial point process models. An alternative method is to maximize the pseudolikelihood, which is specified in terms of the conditional intensity of an underlying point process model. Consider a point process model governed by the parameter set ϑ and having conditional intensity $\lambda_\vartheta(u, \phi)$, $u \in W$. The pseudolikelihood of the point process model can be expressed as [67]

$$\text{PL}(\vartheta, \phi) = \prod_{i=1}^{n(\phi)} \lambda_\vartheta(x_i, \phi) \exp\left\{-\int_W \lambda_\vartheta(u, \phi)\text{d}u\right\} \tag{3.7}$$

The maximum pseudolikelihood estimation of the parameters in ϑ yields the values

that maximize $\mathrm{PL}(\vartheta, \phi)$. By performing a proper quadrature scheme[3], the definite integral on the right hand side of (3.7) can be approximated as

$$\int_W \lambda_\vartheta(u, \phi)\mathrm{d}u \approx \sum_{j=1}^{m} \lambda_\vartheta(u_j, \phi) w_j \tag{3.8}$$

where $u_j \in W$ are quadrature points and $w_j \geq 0$ are the associated quadrature weights for $j = 1, \cdots, m$. Assume the selection of a set of quadrature points $\{u_j\}$ that includes all the observed data points x_i and some other dummy points. Let z_j be a variable that equals 1 if u_j is an observed data point, and 0 if it is a dummy point. Then, the logarithm of the pseudolikelihood can be approximated by

$$\begin{aligned} \log \mathrm{PL}(\vartheta, \phi) &\approx \sum_{j=1}^{m} \left[z_j \log \lambda_\vartheta(u_j, \phi) - w_j \lambda_\vartheta(u_j, \phi) \right] \\ &= \sum_{j=1}^{m} w_j (y_j \log \lambda_j - \lambda_j) \end{aligned} \tag{3.9}$$

where $y_j = z_j / w_j$ and $\lambda_j = \lambda_\vartheta(u_j, \phi)$. Given the observed point pattern ϕ and a candidate model of a specified conditional intensity with the form of (3.3), the algorithm implemented in spatstat first chooses a suitable quadrature rule $\{(u_j, w_j)\}$. It then computes the vector-valued statistic $s_j = (S(u_j), V(u_j, \phi))$, and builds the two-value variable z_j and $y_j = z_j / w_j$. Finally, the algorithm calls generalized linear models to fit the Poisson loglinear regression model $Y_j \sim \mathrm{Poisson}(\lambda_j)$ where $\log \lambda_j = \vartheta \cdot s_j$, and where here "$\cdot$" denotes the dot product. The fitted coefficient vector $\hat{\vartheta}$ is returned as the maximum pseudolikelihood estimate of ϑ. For further explanation see [67]. In spatstat, the command `ppm` is used for fitting Poisson and Gibbs point processes to observed point pattern data.

Irregular parameters, such as the interaction radius R_{S} for the Strauss point process

[3] In numerical analysis, a quadrature scheme is an approximation of the definite integral of a function, which is usually expressed as a weighted sum of function values at specified points within the domain of integration.

and the hard core distance R_{H} for the hard core point process, cannot be estimated directly using the aforementioned algorithm. They must be given *a priori* before model-fitting. The statistical theory for estimating such irregular parameters is unclear. For some special cases, a maximum likelihood estimator of the irregular parameter is available. For example, when fitting a hard core point process, the maximum likelihood estimator of parameter R_{H} is just the minimum nearest-neighbour distance in the observed point pattern [67]. For the Strauss process, the pair correlation function has a jump at R_{S}. This leads to a useful procedure for estimating R_{S} called the "cusp method" [71].

One general strategy available in spatstat for estimating irregular parameters is to maximize the profile pseudolikelihood [67], which means for different values of the irregular parameter, their corresponding fitted models are obtained by the maximum pseudolikelihood method and the irregular parameter value that leads to the largest maximum pseudolikelihood is selected. This can be done by the command `profilepl` in spatstat. The profile pseudolikelihood can be plotted and the best value of the irregular parameter is then indicated.

3.2.4 Minimum contrast method

When the pseudolikelihood is intractable or the conditional intensity is difficult to evaluate, which is typically the case for clustered point processes, an alternative method called the minimum contrast method can be used for model-fitting. The method of minimum contrast [67] is a general technique for fitting a point process model to the observed point pattern data. First, an empirical functional summary statistic[4] is computed from the observed point pattern. Second, the theoretical expected value of this functional summary statistic is derived as an algebraic expression involving the parameters of the candidate model. If the analytical expression does not exist, an av-

[4] In spatstat, the K function and pair correlation function are implemented for the minimum contrast method.

erage estimated version from the simulations of the candidate model could also be used. Then the model is fitted by finding the optimal parameter values that give the closest match between the theoretical and empirical curves. For example, consider a stationary Thomas cluster process and let $K_\vartheta(r)$ denote the K function given in (2.20), where $\vartheta = (\kappa, \mu, \sigma)$. By using the minimum contrast method, the model can be fitted by minimizing

$$D(\vartheta) = \int_{r_0}^{r_1} \left| \hat{K}(r)^q - K_\vartheta(r)^q \right|^p \mathrm{d}r \tag{3.10}$$

where $0 \leq r_0 < r_1$, the strictly positive variables p and q are exponents, $\hat{K}(r)$ is the empirical K function computed from the observed data pattern and $K_\vartheta(r)$ is the theoretical value as a function of the parameters κ and σ. Finally, the parameter μ is inferred from the estimated intensity $\hat{\lambda}$, i.e. $\mu = \hat{\lambda}/\kappa$. The application of the minimum contrast to the candidate point process model is natural as long as an analytical (or even a simulated estimate) version for the functional summary statistic of interest is available. In spatstat, the command `kppm` is used for fitting clustered point processes to observed point pattern data. The default values are $q = 1/4$ and $p = 2$ so that the contrast criterion is the integral of the squared difference between the fourth roots of the two functions.

3.2.5 Simulated envelope test

Even simple point process models for spatial point patterns lead to intractable distributional analysis, and in order to test a fitted model against observed data, we shall make extensive use of Monte Carlo simulations. The most common method is called the simulated envelope test [67], which is based on functional summary statistics of the observed point pattern data together with simulation envelopes to indicate the range of statistical variation. Suppose M independent simulated realizations of the fitted model inside the study region W are run. First, one may compute the estimated summary statistic $\hat{C}$ for each of these realizations, say $\hat{C}^{(j)}(r)$ for $j = 1, \cdots, M$, then obtain the

pointwise lower and upper envelopes of these simulated curves, $C^L(r) = \min_j \hat{C}^{(j)}(r)$ and $C^U(r) = \max_j \hat{C}^{(j)}(r)$. For any fixed value of r, consider the probability that the empirical curve $\hat{C}(r)$ lies outside the envelope $[C^L(r), C^U(r)]$ for the simulated curves. Since $\hat{C}(r)$ and $\hat{C}^{(1)}(r), \cdots, \hat{C}^{(M)}(r)$ are statistically equivalent and independent, this probability is equal to $2/(M+1)$ by symmetry. In other words, if the fitted model is a good assumption for the observed point pattern data, then the empirical curve $\hat{C}(r)$ should lie within the envelope with increasing probability as M increases. Otherwise, the fitted model is rejected. Generally, the simulated envelope test is applied in a hypothetical two-step order for the model-fitting. The first step is the exploratory analysis using the simulated envelope from a Poisson null model, since the Poisson model is usually the dividing line for clustered and regular models. If the empirical curve (from the observed point pattern) is within the simulated Poisson envelope, then the fitted Poisson model will be accepted. If not, the location of the empirical curve will help to indicate a more specific model (either clustered or regular) for fitting, and again, the goodness-of-fit for such a fitted model is evaluated by the simulated envelope test. Values of the empirical curve falling above the envelope indicate a clustered process, while values below the envelope indicate repulsion.

If the observed point pattern is judged to be spatially homogeneous, then a wide choice of functional summary statistics can be used for the simulated envelope test. Popular examples include estimates for the K function $K(r)$, the empty space function $F(r)$ (also known as the contact distribution function), the nearest neighbour distance distribution function $G(r)$, the pair correlation function $g(r)$, the L function $L(r) = \sqrt{K(r)/\pi}$, and the J function $J(r) = (1 - G(r))/(1 - F(r))$. Among these functional summary statistics, the F, G and J functions pertain to the distance characteristics, whereas the K and L functions pertain to second-order characteristics. When testing models using minimum contrast of second-order statistics, it is natural to use distance-related statistics for complementarity, although it is known that the differences of distance characteristics for different patterns are often small [67, 71]. There

are no explicit guidelines to the selection of summary statistics within the same category (either distance or second-order); much of the time, which statistic to be used depends on the problem at hand and on the inclination of the researcher. However, an analytical expression (or approximation) for the K function tends to be available more often than for the other statistics, which may instead rely on empirical data.

Note that the functional summary statistics above are defined and estimated under the assumption that the point process is stationary (homogeneous), as implied by their use of the distance variable r between points. Therefore if the fitted model is nonstationary, deviation between the empirical and theoretical function values are not necessarily evidence of interpoint interaction of the point pattern data, since they may also be attributable to variations in the intensity functions. If the observed point pattern is judged to be spatially inhomogeneous, then at present only the inhomogeneous K function and inhomogeneous pair correlation function[5] are available in $\mathrm{spatstat}$ for model testing.

3.2.6 Model selection

When comparing point process models fitted by the likelihood-based method to the same data, the Akaike information criterion (AIC) is a measure of the relative quality of these fitted models. Suppose Φ is the model of some given point pattern ϕ. Let k be the number of estimated parameters in the model Φ and $\mathbf{L}$ be the maximum value of the likelihood function for the model, i.e. $\mathbf{L} = \mathbb{P}\left(\phi | \left\{\vartheta, \Phi\right\}\right)$, where ϑ consists of the parameters that maximize the likelihood function. The basic AIC formula is defined as [69]

$$\mathrm{AIC} = 2k - 2\log\left(\mathbf{L}\right) \tag{3.11}$$

Given a set of candidate models for the given observed point pattern, the preferred

[5] In [66], the authors briefly discuss how to define the empty space and nearest neighbour distance distribution functions for inhomogeneous point processes. In [72], an alternative definition of the J function for inhomogeneous point processes is proposed based on the representation in terms of product densities.

model is the one with the minimum AIC value. Note that the AIC does not provide a quality test of a model in the absolute sense. If all the candidate models fit poorly, the AIC will not give any indication of that. The AIC just provides a means for model selection.

A more general measure for comparing different fitted models is the root mean square deviation (RMSD) between the empirical and theoretical functional summary statistic C. The RMSD can be defined as [70]

$$\text{RMSD} = \sqrt{\sum_{i=1}^{n} \left\{ \hat{C}(r_i) - C(r_i) \right\}^2 / n} \tag{3.12}$$

where $\hat{C}(r_i)$ and $C(r_i)$ are the empirical and theoretical values at some distance r_i, respectively. If a theoretical value does not exist, then $C(r_i)$ can be replaced by a simulated average value $\bar{C}(r_i)$. Thus the smaller the RMSD value, the better the fit.

3.3 Model-fitting of Real-life Cellular Networks

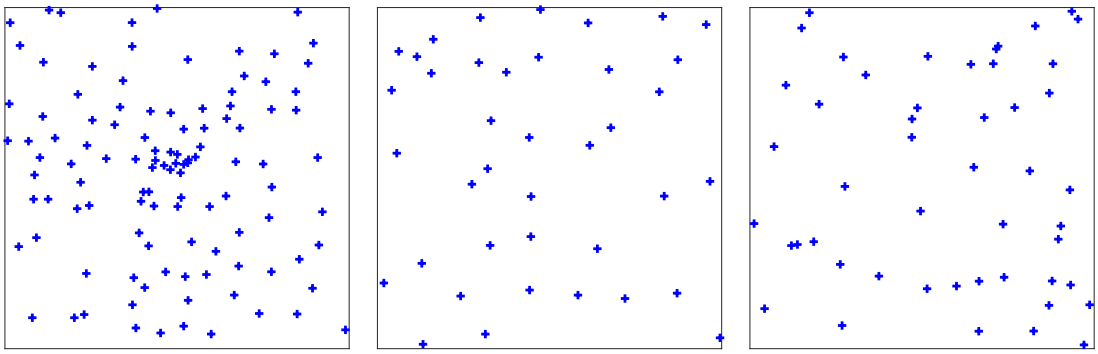
In this section, we fit both stationary and nonstationary point process models to the observed point patterns, which describe the distributions of BSs from a real-life macro-cellular network, although the same methodology can still be applied to micro-cellular and heterogeneous cellular networks. When fitting with stationary point process models, we assume that the observed point pattern is spatially homogeneous. When fitting with nonstationary point process models, the observed point pattern, however, is assumed to be spatially inhomogeneous.

3.3.1 Observed BS point patterns

Our work makes use of real-life data including all BS-related information from TELUS Communications, which is one of the main wireless operators in Canada. The data set

Table 3.1
Details of the Observed BS Point Patterns

	Reference location	Area (km^2)	# of BSs	Intensity (km^{-2})	Nearest neighbour distance (m)
Urban A	53.53° N, 113.51° W	20×20	110	0.275	189
Urban B	49.24° N, 123.09° W	6×6	39	1.083	331
Rural	53.76° N, 111.21° W	150×150	49	0.002	1500



(a) 20×20 km^2 (b) 6×6 km^2 (c) 150×150 km^2

Figure 3.1. The observed BS point patterns: (a) Urban A; (b) Urban B; (c) Rural.

includes all on-air macro BSs of LTE cellular networks, and each record of the BS contains the corresponding location information (i.e. longitude and latitude). The planar coordinates of these real-life BSs are generated by using the Matlab command `distance` (from the Mapping toolbox), which calculates the radius angle θ_{e} of two site locations on earth. Consider a sphere surface, then the distance between the two site locations is the length of the arc $\mathcal{D} = R_{\text{e}} \cdot \theta_{\text{e}}$, where R_{e} is the earth radius (an approximate value of 6371 km is used in this thesis). If we consider a flat plane surface, then the distance between the two site locations should be the length of the chord $\mathcal{D}' = 2R_{\text{e}} \sin(\theta_{\text{e}}/2)$. In this thesis, we use the arc length $\mathcal{D}$ as the distance between any two locations. Given a reference location on the plane (i.e. the origin), the ordinate of each BS is obtained on the assumption that the two locations have the same longitude, whereas the abscissa is obtained assuming that the two locations have the same

latitude[6].

In this thesis, we mainly consider three typical observed BS point patterns accounting for complete spatial randomness, repulsion and clustering, respectively. The detailed information of the selected BS point patterns are summarized in Table 3.1, including the number of points, the area of the study region, the intensity, and the nearest neighbour distance in each data set. From Table 3.1, one can observe that the BSs deployed in urban areas are much more densely spaced than those in the rural area. The high intensity of the BSs in urban areas reflects the great demand for capacity enhancement, whereas the low intensity in rural area shows relatively high requirement for network coverage. BS locations in these areas are depicted in Fig. 3.1. Only square study regions are considered in this thesis, although rectangular or even irregular study regions can also be studied and handled by $\mathbf{R}$ package $\mathrm{spatstat}$.

3.3.2 Fitting with stationary point process models

Most analyses of observed point patterns begin with a test of complete spatial randomness (CSR). Although CSR represents an idealized model, which may be untenable in practice, a test of CSR is used as a means of exploring the observed data sets and indicating a more suitable model for fitting, since CSR acts as a dividing hypothesis to distinguish between observed point patterns that are broadly classified as regular or clustered. A simple test of CSR is called quadrat counting, in which the study region W is divided into congruent rectangular subregions ("quadrats") of equal area. Under the null hypothesis of CSR, the number of points in each quadrat is a Poisson random variable with the same expected value, so the Pearson χ^2 goodness-of-fit test can be used [67]. However, this testing method depends critically on the size of the partition chosen. If there is no natural choice of partition size, then the results can be sensitive

[6]This is largely equivalent to finding the distance parallel to the x-axis between two points in Cartesian coordinates through projection by assuming their y values are the same, and finding the y distance by assuming their x values are the same. The main difference in our case is due to the curvature of the Earth making the exact east-west and north-south distances dependent on the geolocation of the points; however, the small differences that result are negligible on the scale of the areas being considered.

to the chosen partitions.

A more powerful test is the Kolmogorov-Smirnov (KS) test [67], in which a comparison is made between the observed (or empirical) and theoretical (or predicted) cumulative distribution functions of the values of some variables under CSR, such as the empty space function $F(r)$ and K function $K(r)$. If the theoretical distribution is unknown, a KS test can still be carried out, if the theoretical distribution is replaced by an empirical average curve obtained from a number of CSR realizations within the same study region. An envelope test can also be combined with the KS test to assess the significance of departures from CSR for an observed point pattern.

Given the observed point pattern with n points in the study region W, we consider a CSR test based on the K function. Under CSR, the theoretical K function is $K(r) = \pi r^2$. A maximum likelihood estimator of the intensity λ under CSR is $\hat{\lambda} = n/|W|$. The test is performed by comparing the empirical $K(r)$ with its theoretical version. If the observed point pattern is compatible with CSR, the two curves should be roughly overlapped. In this thesis, an envelope test with 99 simulated realizations is used to evaluate the significance of departures from CSR. This means, if the observed point pattern is a sample of a stationary Poisson point process with intensity $\hat{\lambda}$, the empirical K function will be outside the envelope only with the probability $2/(99+1) = 2\%$.

In Fig. $3.2(a)$, inspection of the CSR test shows that the empirical $K(r)$, for the observed rural point pattern, lies close to the curve $K(r) = \pi r^2$ and within the lower and upper envelopes throughout its range, which suggests acceptance of the test for CSR. In Fig. $3.2(b)$, an assessment of the goodness-of-fit based on the nearest neighbour distance distribution function $G(r)$ also demonstrates that a stationary Poisson point process with intensity $\lambda = 0.002$ km^{-2} is a good fit to the observed point pattern. In Fig. $3.3(a)$, a CSR test shows that the spatial features of the observed urban A point pattern are compatible with the existence of an underlying clustering mechanism because the empirical $K(r)$ lies above the upper envelope. Finally, in Fig. $3.3(b)$, for urban B, the empirical $K(r)$ is below the lower envelope of CSR. This provides an

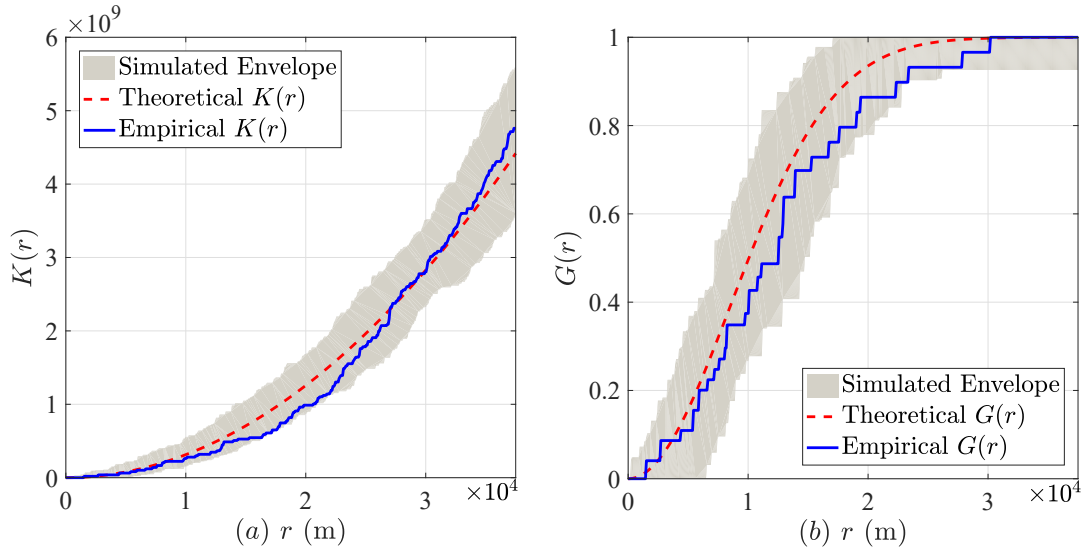


Figure 3.2. Envelope tests of the observed rural BS point pattern and the envelope of 99 realizations of the fitted stationary PPP model. (a) CSR test with K function; (b) Goodness-of-fit test with G function.

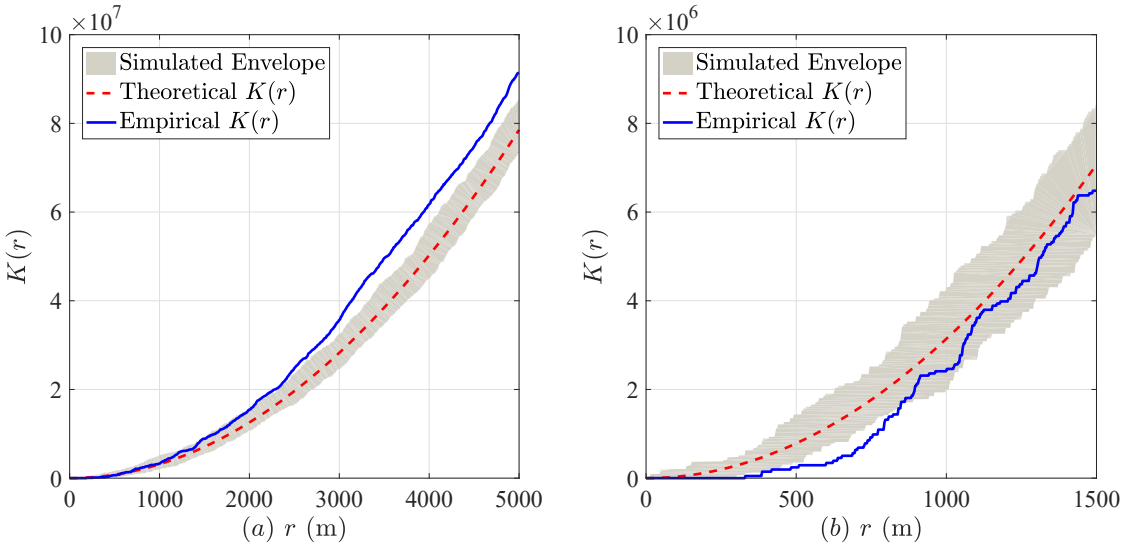


Figure 3.3. CSR test based on K function of the observed BS point patterns and the envelope of 99 realizations of the fitted stationary PPP model. (a) Urban A; (b) Urban B.

explanation for the regularity or repulsive appearance of the observed point pattern.

Based on the CSR test, the observed urban A point pattern is more likely to be a sample of some clustered point process model. We therefore fit Matérn and Thomas cluster processes to the observed point pattern by means of minimum contrast of the

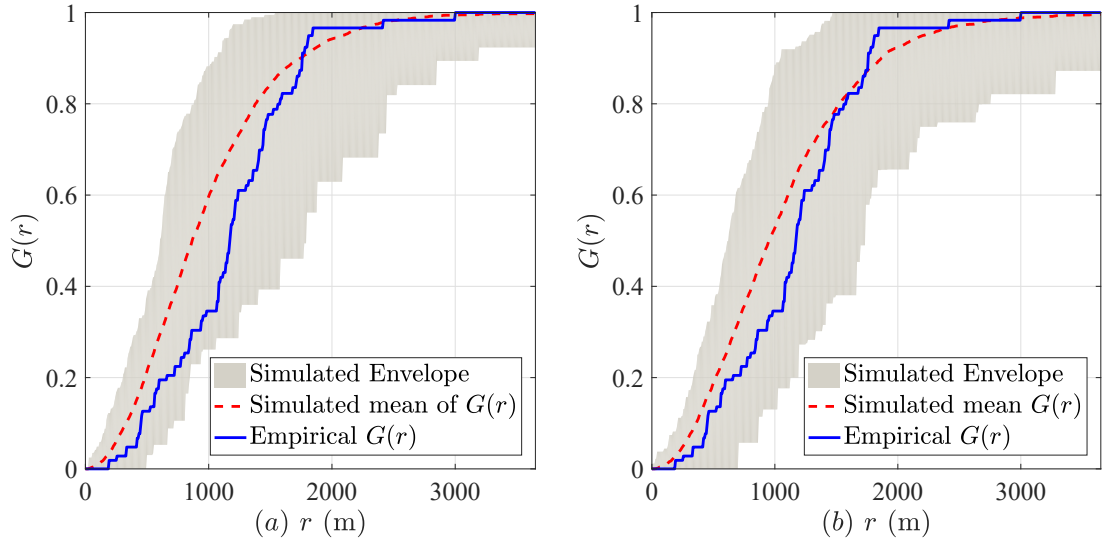


Figure 3.4. Goodness-of-fit test based on G function of the observed BS urban A point pattern and the envelope of 99 realizations of the fitted stationary models. (a) Matérn cluster process; (b) Thomas cluster process.

K function[7]. Other candidate models could also be used for model-fitting, such as a Neyman-Scott cluster process with Cauchy and variance Gamma kernels, as well as a log-Gaussian Cox process [67], [71]. By using the command `kppm` in `spatstat`, the values of the parameters can be easily obtained. Referring back to (2.18), when fitting with a Matérn cluster process, we get estimated parameter values of $\hat{\kappa}_M = 4.09 \times 10^{-4}$ km^{-2} and $\hat{R}_{MC} = 60.6$ km. Since the overall intensity $\lambda_M = \kappa_M \cdot \mu_M$, with $\hat{\lambda}_M = 110/(400 \text{ km}^2) = 0.275$ km^{-2}, we get $\hat{\mu}_M = \hat{\lambda}_M/\hat{\kappa}_M \approx 672.3$. This means that there is an average of 672.3 points in each representative Matérn cluster. When fitting with a Thomas cluster process, referring back to (2.20), we get estimated parameter values of $\hat{\kappa}_T = 1.60 \times 10^{-3}$ km^{-2} and $\hat{\sigma} = 15.5$ km, and hence $\hat{\mu}_T = \hat{\lambda}_T/\hat{\kappa}_T \approx 172.1$. Thus, there is an average of 172.1 points in each representative Thomas cluster. In Fig. 3.4, an assessment of the goodness-of-fit based on the nearest neighbour distance distribution function $G(r)$ shows that either the fitted stationary Matérn or Thomas

[7]One appealing feature of an analysis using the K function is that the mathematical form of $K(r)$ is known, either explicitly or as an integral, for a number of potentially useful classes of spatial point processes. Besides, the K function is also preferred over the pair correlation function, when there is a limited sample of points [71].

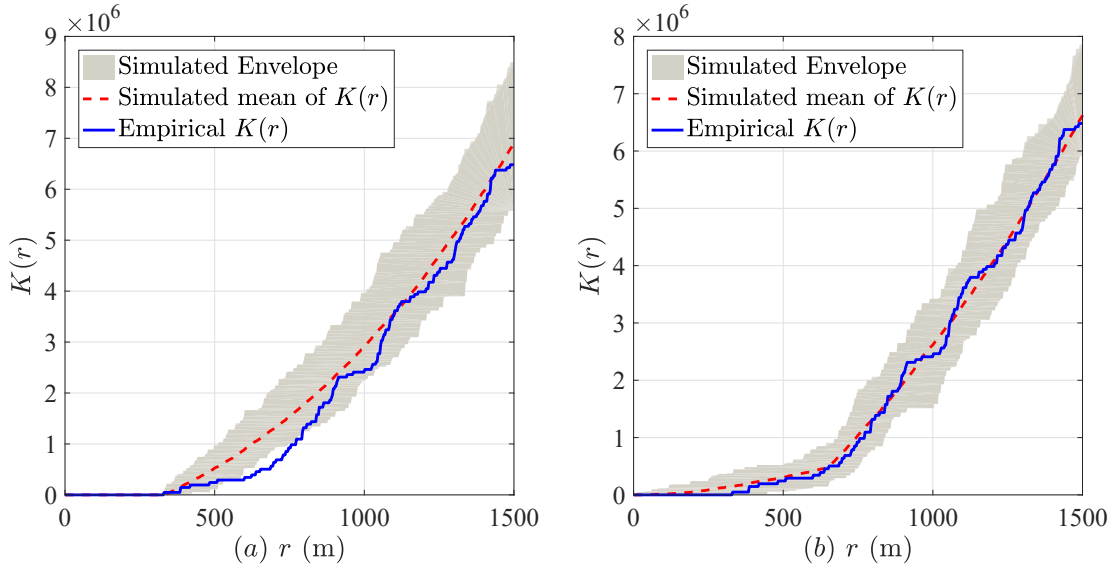


Figure 3.5. Goodness-of-fit test based on K function of the observed urban B BS point pattern and the envelope of 99 realizations of the fitted stationary models. (a) Hard core point process; (b) Strauss point process.

cluster process is a good fit to the observed urban A BS point pattern. According to the RMSD criterion, however, the Thomas cluster process is a better fit than the Matérn cluster process; the fitted Thomas cluster process has an RMSD value of 0.071, whereas the value is 0.098 for the fitted Matérn cluster process.

For the observed urban B BS point pattern, since the CSR test indicates a repulsive interaction between the points, we fit hard core and Strauss point processes to the observed BS point pattern. For the hard core point process, the maximum likelihood estimate of the irregular parameter R_{H} is the minimum nearest neighbour distance in the observed point pattern, thus $\hat{R}_{\mathrm{H}} = 330$ m. The fitted value of β_{H} is $\hat{\beta}_{\mathrm{H}} = 1.491 \times 10^{-6}$. For the Strauss point process, by using the method of maximum profile pseudolikelihood (with the step size equal to 1 m), we obtain the estimated irregular parameter $\hat{R}_{\mathrm{S}} = 650$ m. The corresponding fitted parameter values are $\hat{\gamma}_{\mathrm{S}} = 0.22$, and $\hat{\beta}_{\mathrm{S}} = 4.043 \times 10^{-6}$. In Fig. 3.5, an assessment of the goodness-of-fit based on the K function shows that the Strauss point process is a good fit, whereas the hard core point process is not a proper model for the observed point pattern since its K function

values for the observed data fall outside of the lower simulated envelope. According to the Akaike information criterion (AIC), the fitted Strauss point process has a value of about 719 whereas for the fitted hard core point process the value is about 809, which corroborates that the Strauss point process is a better fit.

3.3.3 Fitting with nonstationary point process models

Although homogeneity is a convenient assumption for the observed point pattern, especially if, as is often the case, only a single sample (or map) is available, in many cases inhomogeneity is present. As for the deployment of BSs, there are many key issues that need to be considered or factored into the decision of a BS (or cell-site) location, such as terrestrial characteristics (which may relate to the propagation conditions), human activities, or even government restrictions. The deployments of different tiers in a multi-tier cellular network are not mutually independent, but rather are of a coordinated nature. More often than not, if the macrocell BSs are deterministically planned to begin with, then small cell BSs are more likely to be deployed away from the macrocells (usually at the macrocell edges), e.g. in order to compensate for coverage gaps. This means the distribution of macrocell BSs will, to some extent, affect the distribution of small cell BSs. Even for an observed macrocell BS point pattern, a large-scale spatial feature of human residential and working environments would also be a potential covariate for determining the distribution of macrocell BSs. In [40], the BS layouts from different wireless operators in a shared cellular network are also shown to be correlated. Therefore the distribution of real-life BS locations simply cannot be described as spatially homogeneous. The use of stationary point process models could be invalid when fitting to an observed BS point pattern, because the deviations between the empirical and theoretical functions are not necessarily evidence of interpoint interaction, since they may also be attributable to variations in intensity.

When fitting with a nonstationary model, it is essential to specify an intensity function $\lambda(\cdot)$ ahead of any fitting methods. The intensity function can be formulated in

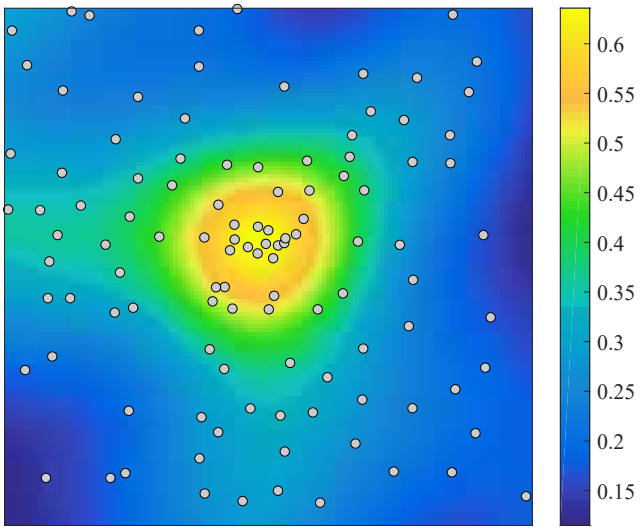


Figure 3.6. Kernel-smoothed intensity estimates (km^{-2}) of the observed urban A point pattern (grey circles).

terms of spatial coordinates, or an observed covariate, or a mixture of both. Given the specified intensity function form of interest, a further examination of the observed point pattern can be carried out through the simulated envelope test based on some functional summary statistics such as the inhomogeneous K function and the inhomogeneous pair correlation function. Similar to the homogeneous case, this is done by assuming that the null model is an inhomogeneous Poisson point process with the specified nonconstant intensity function. If the empirical functional summary statistic of the observed point pattern is within the simulated envelope, then the hypothesis of the observed pattern being a sample from a nonstationary Poisson model is accepted. If the empirical functional summary statistic is not within the simulated envelope, then there are two possibilities. First, it might be due to an inappropriate choice of the intensity function for the nonstationary Poisson model, hence other forms of intensity function could be proposed for verification. Second, a more specific clustered or repulsive model can be used, according to on which side of the simulated envelope the test has failed.

For the observed urban A point pattern, Fig. 3.6 shows a kernel-smoothed esti-

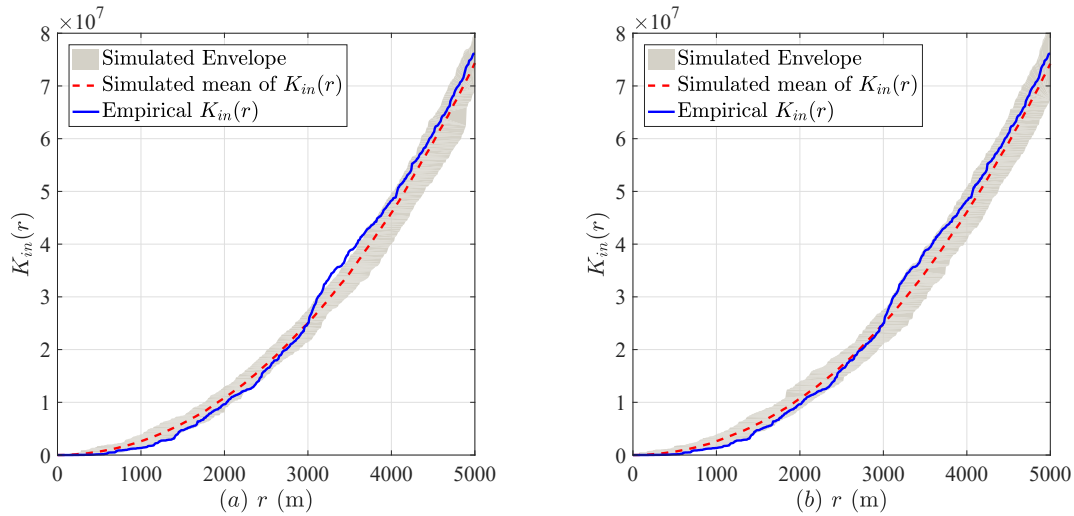


Figure 3.7. Goodness-of-fit test based on inhomogeneous K function of the observed urban A point pattern and the envelope of 99 realizations of the fitted models with intensity function $\log \lambda(x_u, y_u) = \alpha_0 + \alpha_1 \sqrt{x_u^2 + y_u^2}$. (a) Poisson point process; (b) Thomas cluster process.

mate of the intensity function. According to the BS distribution, one can see a large abundance of BSs located in the center, which results in the lightest color in the image, whereas blue colors correspond to locations with fewer base stations. The darker the blue, the less densely the BSs are spaced. The graphical result indicates a spatial trend of the BS intensity that decays with the distance from the study region center. In order to explore and predict a proper intensity form, one can use various regression models to characterize the variation of the local intensity (the response variable) with respect to the distance r (the predictor variable). In this thesis, we will consider a loglinear regression model, that is, $\log \lambda(x_u, y_u) = \alpha_0 + \alpha_1 \sqrt{x_u^2 + y_u^2}$, because it has the best goodness-of-fit in terms of the square of the correlation between the observed intensity values and the predicted intensity values. Note that the regression analysis is only used for predicting a suitable intensity form. The regular parameters (α_0, α_1) will be re-estimated by inserting the intensity form into the `ppm` or `kppm` commands in `spatstat`.

In Fig. 3.7(a), a fitted Poisson model with an intensity $\log \lambda(x_u, y_u) = \alpha_0 + \alpha_1 \sqrt{x_u^2 + y_u^2}$ is considered, where α_0 and α_1 are the regular parameters to be esti-

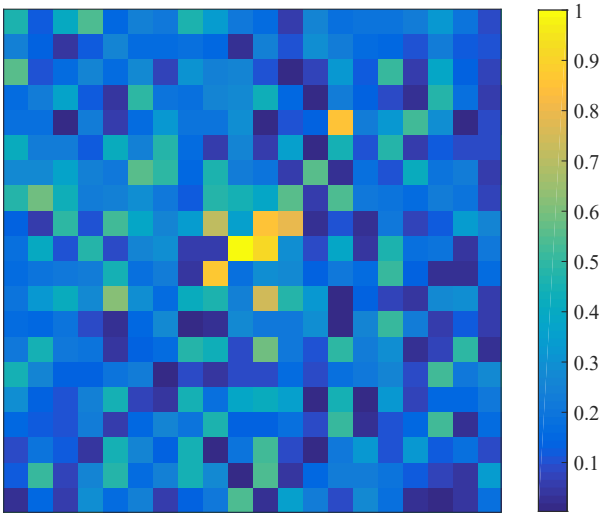


Figure 3.8. Covariate $Z(u)$ for normalized population densities in the urban A area (20 by 20 grid).

mated. By inserting the intensity function into the `ppm` command (which is based on the maximum likelihood method (3.7)–(3.9)), we obtain $\alpha_0 = -14.001$ and $\alpha_1 = -1.582 \times 10^{-4}$ m^{-1}. The negative values suggest a decrease of the intensity exponentially with the distance from the location to the origin. Although the empirical curve is very close to the simulated mean, the fitted model is still statistically rejected because the empirical curve is above the simulated envelope, which indicates attraction between the points. Therefore, in Fig. 3.7(b), we fit a Thomas cluster process with the same intensity form to the observed urban A point pattern. By using the `kppm` command (which is based on the minimum contrast method), we obtain the same estimates for first-order regular parameters $\alpha_0 = -14.001$ and $\alpha_1 = -1.582 \times 10^{-4}$ m^{-1}, whereas for the fitted cluster parameters we have $\hat{\kappa} = 2.17$ km^{-2} and $\hat{\sigma} = 2.44$ km. The envelope test is satisfied and hence the nonstationary Thomas model is accepted as a fit to the observed urban A point pattern.

For the observed urban A point pattern, in Fig. 3.8 we also consider the dependence of the spatial trend on a practical covariate that accounts for the population densities (per square km) based on the 2016 census data collected by Statistics Canada [73, 74]. Intuitively, the more people there are in an area, the more BSs are likely to be in that

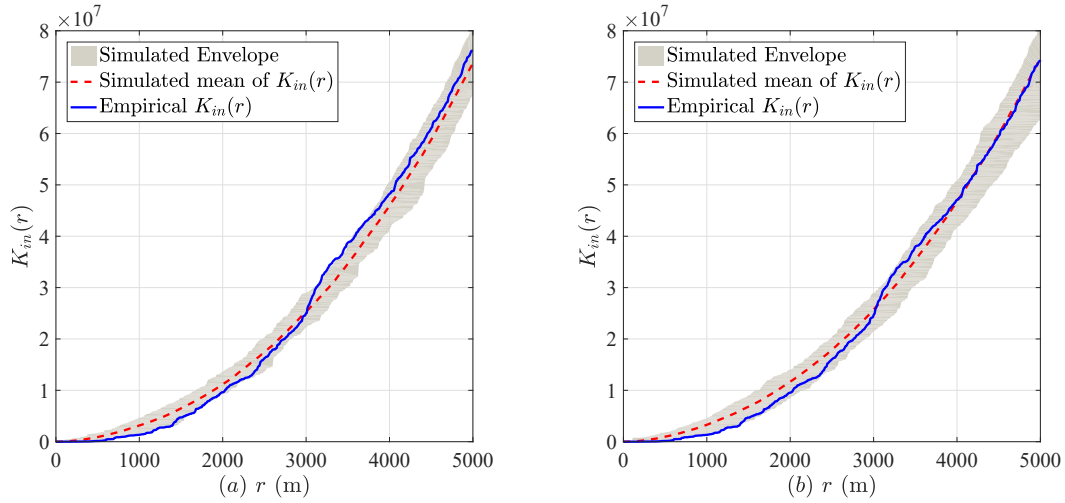


Figure 3.9. Goodness-of-fit test based on inhomogeneous K function of the observed urban A point pattern and the envelope of 99 realizations of the fitted models with intensity $\log \lambda(u) = \alpha_0^{'} + \alpha_1^{'} Z(u)$. (a) Poisson point process; (b) Thomas cluster process.

area. The collection areas for the census data (i.e. "census tracts") have various sizes and are irregularly shaped, with boundaries defined by streets, neighbourhoods, waterways, etc. As an approximation that is more easily manageable for a covariate, we divide the study region into a 20 by 20 grid of 1 km $\times$ 1 km squares. Each square is assigned the population density value (normalized to the maximum 4454 per square km) of the census tract corresponding to the geographic location of the geometric center of that square. We assume that any location within the square has the same covariate value. The intensity form as a function of the covariate is not easy to be specified. Usually, the intensity function is assumed to be loglinear or proportional to the covariate [67]. In this thesis, we will only use an intensity form that is loglinear with the covariate. In Fig. 3.9(a), we fit a Poisson model with an intensity function $\log \lambda(u) = \alpha_0^{'} + \alpha_1^{'} Z(u)$ to the observed urban A point pattern. By calling the fitting command `ppm`, we obtain $\alpha_0^{'} = -16.263$ and $\alpha_1^{'} = 3.747$; these values suggest an increase of the intensity exponentially with the population density. The envelope test is not satisfied due to the curve being outside the upper envelope, and hence the Poisson model with the intensity loglinear to the covariate is rejected. In Fig. 3.9(b), we fit a Thomas model with an intensity function $\log \lambda(u) = \alpha_0^{'} + \alpha_1^{'} Z(u)$ to the observed urban A point pattern. By

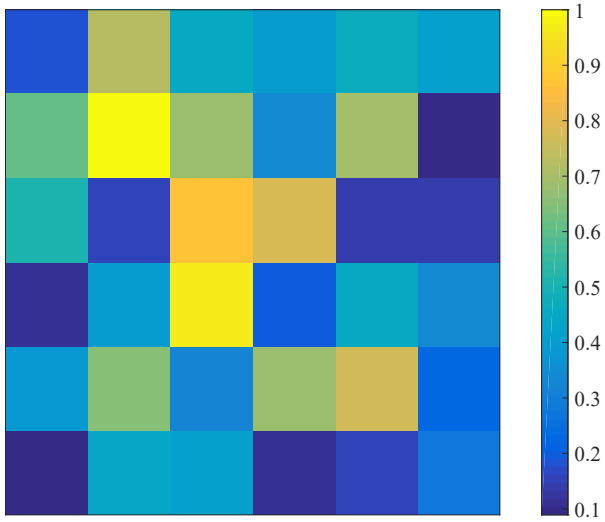


Figure 3.10. Covariate $Z(u)$ for normalized population densities in the urban B area (6 by 6 grid).

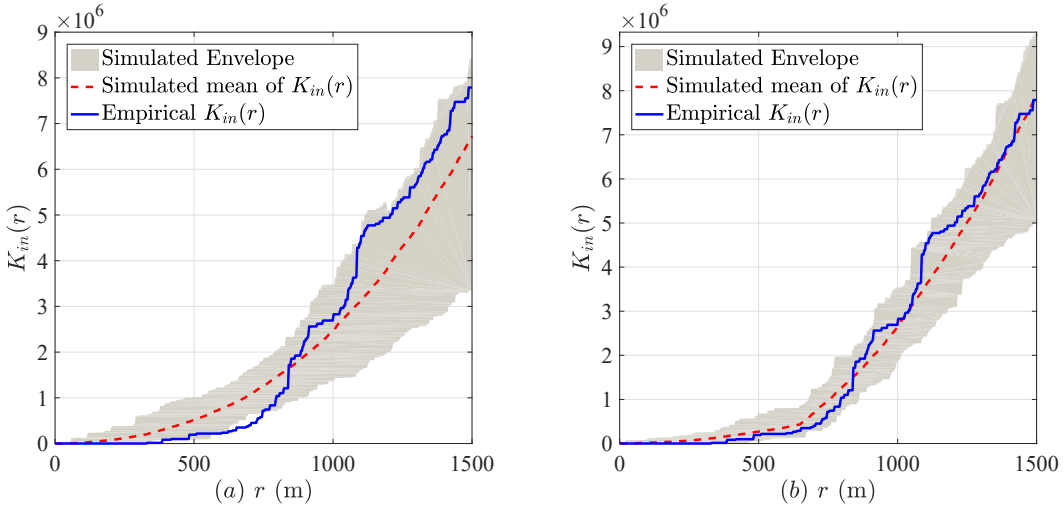


Figure 3.11. Goodness-of-fit test based on inhomogeneous K function of the observed urban B point pattern and the envelope of 99 realizations of the fitted models with an intensity that is loglinear with the population covariate. (a) Nonstationary Poisson process; (b) Nonstationary Strauss process.

calling the fitting command `kppm`, we obtain the same estimates of $\alpha_0^{'} = -16.263$ and $\alpha_1^{'} = 3.747$. For the fitted cluster parameters, we have $\hat{\kappa} = 2.17$ km^{-2} and $\hat{\sigma} = 2.44$ km. The fitted Thomas model is statistically accepted because the empirical curve is within the simulated envelope.

For the observed urban B point pattern, we consider a similar covariate accounting

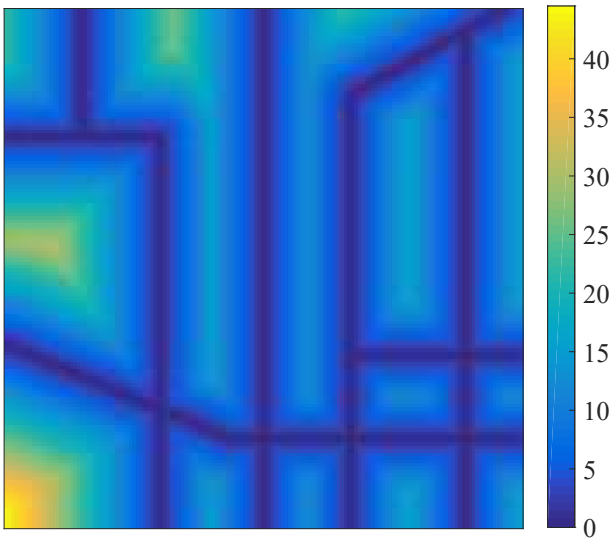


Figure 3.12. Covariate $Z(u)$ for closest distances in the rural area.

for the local population densities (normalized to the maximum 9098 per square km) using a 6 by 6 grid of 1 km $\times$ 1 km squares, shown in Fig. 3.10. The intensity function $\lambda(u)$ of the fitted model is assumed to be $\log \lambda(u) = \alpha_0 + \alpha_1 Z(u)$, where α_0 and α_1 are the regular parameters to be estimated. Fitting the nonstationary PPP and Strauss models to the observed urban B point pattern, using the maximum likelihood method, we obtain $\alpha_0^{\text{PPP}} = -14.959$ and $\alpha_1^{\text{PPP}} = 2.127$, and $\alpha_0^{\text{SP}} = -14.980$ and $\alpha_1^{\text{SP}} = 4.852$, respectively. In Fig. 3.11(a), an envelope test based on $K_{in}(r)$ indicates that with this specified intensity form, although the empirical curve is close to being within the inhomogeneous Poisson envelope, the observed point pattern still has some repulsive interactions among the points. In Fig. 3.11(b), an envelope test for the goodness-of-fit shows that the nonstationary Strauss process is a good fit to the observed point pattern because the empirical $K_{in}(r)$ curve is within the simulated envelope. According to the AIC, the nonstationary Strauss model has a value of about 690, whereas for the fitted stationary Strauss model the value is 719, which shows that the nonstationary Strauss model is a slightly better fit.

For the observed rural point pattern, we consider a different kind of covariate that

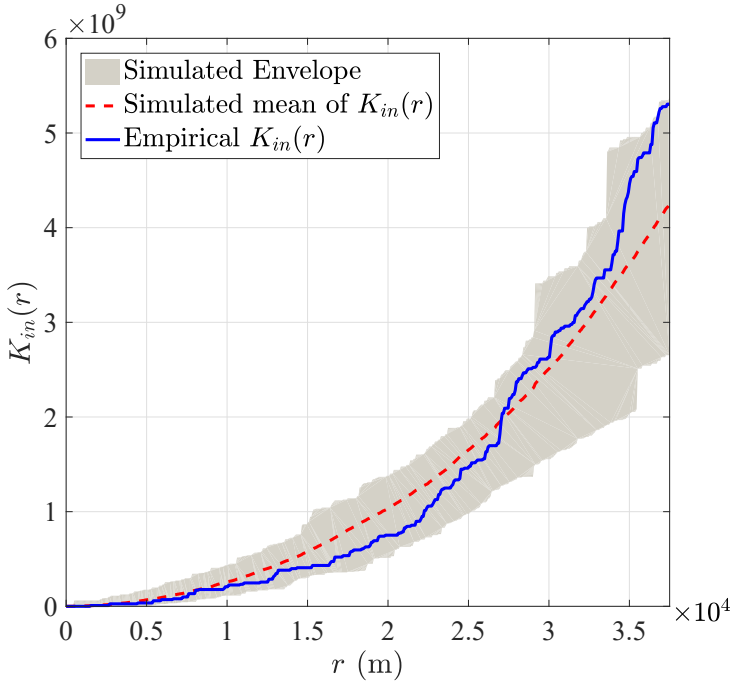


Figure 3.13. Goodness-of-fit test based on inhomogeneous K function of the observed rural point pattern and the envelope of 99 realizations of the fitted Poisson model.

accounts for the distances from the BSs to the main roads across the rural area, because the population and BSs would most likely be located along the lines of travel. These roads are modelled as line segments in the study region. The covariate $Z(u)$ at any location u is calculated to be the distance (in km) from u to its closest line segment, as shown in Fig. 3.12. We fit a Poisson model with an intensity that is loglinear with the covariate $Z(u)$, i.e. $\lambda_\theta(u) = \exp\left(\alpha_0^{\mathrm{R}} + \alpha_1^{\mathrm{R}} Z(u)\right)$. By inserting the intensity form into fitting command `ppm`, we obtain $\alpha_0^{\mathrm{R}} = -19.124$ and $\alpha_1^{\mathrm{R}} = -0.129$ (km^{-1}), which indicates an exponential increase of the intensity with proximity to the roads. In Fig. 3.13, an envelope test shows that the nonstationary Poisson model is also a good fit to the observed rural point pattern. According to the AIC, the nonstationary Poisson model has a value of about 2037 whereas for the fitted stationary Poisson model the value is 2055, which shows that the nonstationary Poisson model is a better fit.

3.4 Wireless Performance Metrics for Fitness Assessment

Since we aim to use point process models to describe wireless networks, it is sensible to use some more relevant metrics that characterize the network performance to evaluate the goodness-of-fit for different fitted point process models, such as the probability of coverage.

In wireless communications, the signal power decays with the distance between the transmitter and the receiver according to the power law

$$P_r(u) = P_t(x)h_{xu}\|x-u\|^{-\eta}, \tag{3.13}$$

where $x \in \mathbb{R}^2$ is the spatial location of a test transmitter, $P_t(x)$ is the transmit power of the transmitter located at x, $u \in \mathbb{R}^2$ is the spatial location of the receiver, h_{xu} is a random variable accounting for the random channel (power) gain due to multi-path fading and shadowing between the two locations x and u, $\|\cdot\|$ is the Euclidean norm, and η is the path-loss exponent. The signal-to-interference-plus-noise ratio (SINR) at the test receiver therefore can be calculated as

$$\text{SINR}(u) = \frac{P_t(x_0)h_{x_0u}\|x_0-u\|^{-\eta}}{\Omega + \sum_{x\in\mathcal{I}} P_t(x)h_{xu}\|x-u\|^{-\eta}}, \tag{3.14}$$

where u is the location of the test receiver, x_0 is the location of the desired transmitter, $\mathcal{I} = \{x_1, x_2, \cdots\}$ is the set of the locations of the interfering BSs, and Ω is the noise power. The summation term $\sum_{x\in\mathcal{I}}\cdots$ is the aggregate interference power at the test receiver. Considering a nearest-BS connectivity policy, the coverage probability P_{c} is defined as the probability that a randomly located user achieves a given SINR threshold T when being served by the closest BS, while the rest of the BSs act as interferers. That is,

$$P_{\text{c}}(T) = \mathbb{P}(\text{SINR}(u) \geq T). \tag{3.15}$$

A common assumption in the literature (that we also use in our simulations) is for h_{xu} to be exponentially distributed (with a mean of 1), which corresponds to a Rayleigh fading scenario. Furthermore, the interference-limited scenario is typically assumed, in which Ω may be considered negligible. Under these assumptions, when all the BSs transmit at the same power level and the path-loss exponent $\eta = 4$, the theoretical average probability of coverage achieved in a homogeneous PPP cellular network is [21]

$$P_{\mathrm{c}}(T) = \frac{1}{1 + \sqrt{T}(\pi/2 - \arctan(1/\sqrt{T}))}. \tag{3.16}$$

In the PPP case, the coverage probability becomes independent of the specific values of the transmit power P_t and the PPP intensity λ, though this is not the case in general. Note that except for the homogeneous PPP model, there are no analytically tractable expressions for the coverage probability. In order to compare different fitted models, we therefore estimate the probability of coverage through Monte Carlo simulations, from which $P_{\mathrm{c}}(T)$ is determined by the average fraction of the whole area where SINR $> T$. For simplicity, we assume all BSs transmit with unit power, $\eta = 4$, and an interference-limited scenario, i.e. $\Omega = 0$. (The SINR values therefore reduce to signal-to-interference ratio (SIR) values.) The network performance is evaluated considering some various-sized central areas of the study region, with the intention of mitigating edge effects. For the observed point pattern, 100000 SIR values are computed at locations distributed uniformly randomly over the central area of the study region. For each fitted model, 1000 realizations are generated and for each realization, the SIR values are evaluated at the same 100000 locations chosen for the observed point pattern.

For the observed urban A point pattern, we have found four fitted models that satisfy the envelope test. In Fig. 3.14, we compare these fitted models in terms of the probability of coverage over two different sized central areas of the study region, i.e. 12 km $\times$ 12 km and 16 km $\times$ 16 km. The graphical results show that both the fitted stationary Matérn and Thomas models have the largest deviation from the probability of

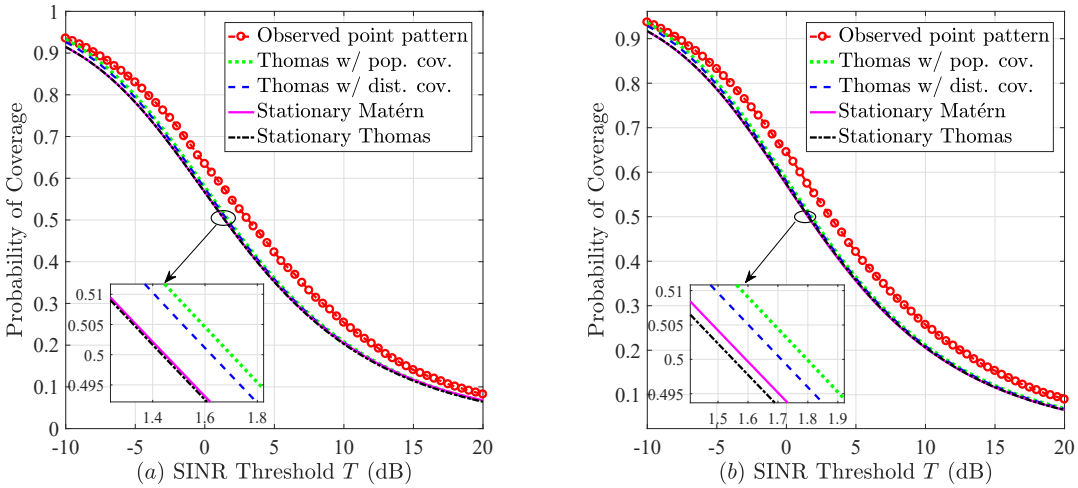


Figure 3.14. Probability of coverage for different fitted point process models for urban A, where the user is uniformly distributed in the central area with the following sizes: (a) 12 km $\times$ 12 km; (b) 16 km $\times$ 16 km.

coverage of the observed point pattern. Although the Thomas model with population covariate has the closest performance to that of the observed point pattern, both the nonstationary models, i.e. Thomas model with population density covariate and distance covariate, are better than the stationary models. When comparing RMSD values with respect to the empirical curve from the observed data, in Fig. 3.14(a), the fitted Thomas model with population covariate has a value of 0.043 whereas the Thomas model with distance has a value of 0.046. Both are smaller than those of the fitted stationary Thomas and Matérn models which are 0.052 and 0.051, respectively. In Fig. 3.14(b), the fitted Thomas model with population covariate has a value of 0.045 and the Thomas model with distance has a value of 0.048, whereas both the fitted stationary Matérn and Thomas models have the same RMSD value of 0.055.

For the observed urban B point pattern, we have found two fitted models that satisfy the envelope test. In Fig. 3.15, we compare these fitted models in terms of the probability of coverage over two different sized central areas (3 km $\times$ 3 km and 5 km $\times$ 5 km) of the study region. Although the two fitted models both satisfy the envelope test, the graphical results show that the fitted nonstationary Strauss model is a better fit than the fitted stationary Strauss model because the coverage curve of the fitted non-

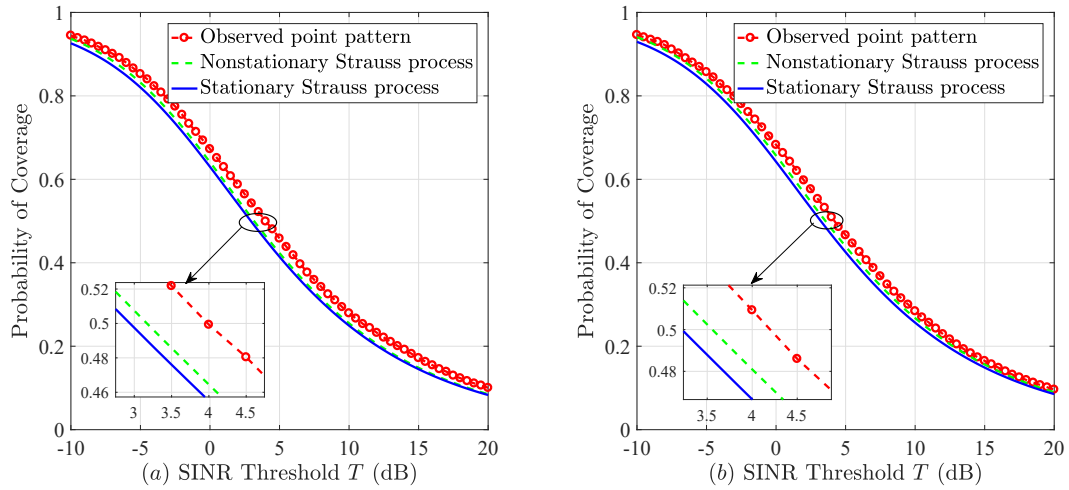


Figure 3.15. Probability of coverage for different fitted point process models for urban B, where the user is uniformly distributed in the central area with the following sizes: (a) 3 km $\times$ 3 km; (b) 5 km $\times$ 5 km.

stationary Strauss model is closer to the curve of the observed data. This agrees with the results seen earlier when comparing the fitness of the two models using their AIC values. When comparing the RMSD values with respect to the empirical curve from the observed data, in Fig. 3.15(a), the fitted nonstationary Strauss model has a value of 0.026, whereas the stationary Strauss model has a value of 0.029. In Fig. 3.15(b), the fitted nonstationary Strauss model has a value of 0.019, whereas the stationary Strauss model has a value of 0.026.

For the observed rural point pattern, we have found two fitted models that satisfy the envelope test. In Fig. 3.16, we compare these fitted models in terms of the probability of coverage evaluated over 90 km $\times$ 90 km and 120 km $\times$ 120 km sized central areas of the study region. The graphical results show that the coverage curve of the fitted nonstationary Poisson model is closer to the empirical coverage curve of the observed pattern than that of the fitted stationary Poisson model. This reflects the earlier fitness comparison of the two models using their AIC values. When comparing the RMSD values with respect to the empirical curve from the observed data, in Fig. 3.16(a), the fitted nonstationary Poisson model has a value of 0.048, whereas the sta-

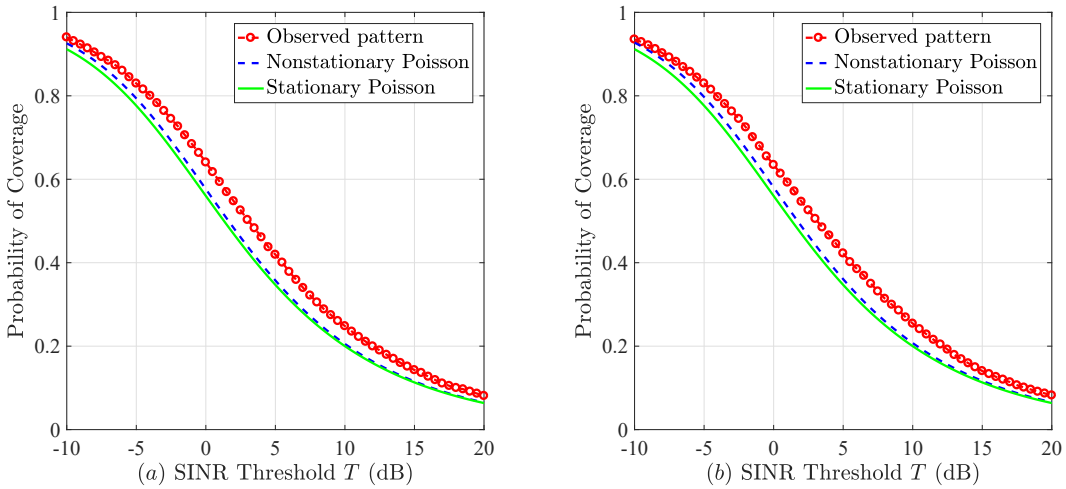


Figure 3.16. Probability of coverage for different fitted point process models for the rural network, where the user is uniformly distributed in the central area with the following sizes: (a) 90 km $\times$ 90 km; (b) 120 km $\times$ 120 km.

tionary Poisson model has a value of 0.055. In Fig. 3.16(b), the fitted nonstationary Poisson model has a value of 0.047, whereas the stationary Poisson model has a value of 0.056.

As one can observe, there remain differences in the probability of coverage between the fitted nonstationary models and the observed point pattern. This might be due to the incompleteness of the covariate information. For example, the population density covariate is based on residences and does not account for industrial areas, shopping centers, parks, sports complexes, etc., which would also require cellular coverage. Other covariates, such as the land zoning/usage or terrain features, may also play an important role in determining the distribution of BSs.

3.5 Summary

In this chapter, we have presented a comprehensive analysis of the spatial modelling of real-life Canadian cellular networks. Part of the analysis uses stationary point process models, similarly to [39], but we have also considered inhomogeneous models. Here,

secondary covariate information, accounting for population densities and distance from the BSs to their closest main roads, has for the first time been introduced as a controlling factor for the intensity function of a fitted spatially inhomogeneous model. The first step of the analysis of the model-fitting is a simulated envelope test using functional summary statistics of the point process, such as the K function (or its inhomogeneous analogue, $K_{in}(r)$); this test examines for deviations from a baseline Poisson point process exhibiting independence between points. Such deviations can be the result of inhomogeneous variations in intensity rather than interactions between points. Therefore, if covariate information is available or inhomogeneity is thought to be present, it should be incorporated into the model to account for this. If the empirical curve of the observed data is within the simulated envelopes, then a fitted stationary PPP model is accepted; otherwise, the PPP model is rejected. When empirical curve is above the envelopes, clustered point process models are instead considered; when below, regular point process models are used. For each examined real-life network scenario, multiple different fitted models have been found to satisfy the conditions of an envelope test. Statistics such as the AIC and RMSD values, and more relevant wireless metrics such as the probability of coverage, have then been used for the selection of the best model among the candidates.

The stochastic geometry modelling offers an alternative way to visualize, quantify, and simulate the network. Numerical evaluations of the performance expressions derived from the stochastic geometry models can usually be calculated faster than running a full Monte Carlo simulation on the network. When it comes to a green field (from scratch) design with a predefined QoS requirement, stochastic geometry modelling has more profound effects on decisions related to, for example, how many BSs should be deployed, where to put them, and how far or how close a distance should be between the BSs. In addition, the coverage requirement is also associated with traffic-loading requirements, which in turn rely on the stochastic model chosen to determine the traffic distribution or off-loading from an existing cell site to new cell sites. More-

over, if additional covariate information on a green field is available, then the network could be designed and deployed more thoughtfully and efficiently.

We have seen that going from the lack of a covariate to the inclusion of one has improved the accuracy of the model. However, there still remain gaps, e.g. between the results of the model and the simulations of the real-life cellular network. Future research work can be foreseen such that more appropriate intensity functions can be specified for nonstationary models with the use of other covariate information, such as the propagation characteristics or traffic models derived from technical reports. For the population covariate, a finer grid could be used, or other irregular shapes more closely resembling the census areas could be adopted. Furthermore, several of these covariates could be used simultaneously. Marked point processes and multi-variate point processes [31, 33] can also be used to model other networks such as multi-tier or cognitive radio networks. The covariates themselves could also potentially be modelled as random point or line processes instead of being deterministic.

Chapter 4

Empirical Distribution of Nearest-Transmitter Distance in Wireless Networks Modelled by Matérn Hard Core Point Processes

The availability of the distribution of the distance between a generic location and the closest point to it from a point process (also called the empty space function) is crucial for a tractable performance analysis of wireless networks modelled by such a point process. This distribution is key to the methodology used with the homogeneous Poisson point process models [21, 22] or more general vulnerability region analysis [13].

As a step toward applying these methodologies to more general point process models, in this chapter, we fit the empirical probability density function of the closest-point distance in the Matérn hard core point process of Type II to various existing distributions, and find that the Weibull distribution has the best goodness-of-fit among all other distributions examined (e.g. the gamma, log-normal and Rayleigh distributions). We also propose a better piecewise probability density function for the closest-point distance, including an exact expression and a heuristic formula, which can be fitted by

a Weibull-like function. Simulation results show that the proposed piecewise model has a very close goodness-of-fit to the empirical data.

4.1 Introduction

Stochastic geometry [31, 41] has been widely used for the modelling and evaluation of wireless networks [25, 26, 34, 75], providing spatial averages over a large number of points at different locations for the quantities of interest, such as average coverage probability and average data rate. Most of the analysis in the literature is based on the assumption that the locations of base stations are stationary and distributed according to a Poisson point process [21, 22]. The complete spatial randomness and independence between points makes the PPP less difficult to deal with. In particular, when talking about its tractability, we are mainly referring to the closed-form probability generating functional (PGFL) [41] and to the distribution of the distance between a generic location (not part of the point process) and its closest point in a PPP model. This distance can be interpreted as the length of the desired link between the so-called "typical user" and its associated BS in a wireless network, under the assumption of the closest-BS connectivity model. From the PGFL, the SINR distribution can be derived conditioned on a fixed desired link length (again, if the closest-BS connectivity model is assumed). By de-conditioning the fixed desired link length with the distance distribution, an average performance metric can be obtained. Unfortunately, both the PGFL and closest-point distance distribution in analytical closed form only exist for PPP models[1]. Therefore, theoretical analysis is rarely performed using other spatial point processes, such as a hard core point process (HCPP) [25, 26, 34, 75]. An HCPP is a repulsive point process where no two points of the process can coexist within a distance less than a hard-core parameter; i.e., an HCPP correlates the locations of the points by conditioning on a minimum distance separating them. In the literature,

[1] Although an analytical PGFL is available for Poisson cluster processes in [24], it is very involved to perform a numerical calculation using it.

PPP-based analysis of networks has been widely addressed and is well understood. However, due to the increased complexity and reduced tractability, only a few works on repulsive network models exist [10,12,13,35,76]. Almost all the analysis for HCPP models is based on the assumption that an HCPP model behaves asymptotically like a PPP model, either because of small parent PPP intensities or small hard-core parameter values (or both).

In this chapter, we focus on the empirical distribution of the distance between a generic location (distributed uniformly over the area of the network) and its closest point in an HCPP model (specifically, involving a Matérn hard core point process of Type II [32], which we shall refer to as "MHCPP" for short). We compare the empirical distribution with various known distributions, including gamma, log-normal, Rayleigh and Weibull distributions. Among the examined distributions, the Weibull distribution most closely fits the closest-point distance distribution. We also propose a piecewise probability density function (PDF) model, which fits even better than the Weibull distribution. These findings not only help us further understand the specific hard core point process, but also provide a step toward the numerical analysis of HCPP-based wireless networks.

Independently and in parallel with our work on this topic, the authors of [77] derived an analytical expression for the contact distance between an MHCPP point and a point from an independent homogeneous PPP, which is an equivalent scenario to ours. However, that derivation uses a PPP approximation for the void ball around the location (point) of interest, i.e. assuming the parent PPP of the MHCPP is devoid of points in the ball rather than the MHCPP itself. We shall show that our empirical expression not only fits better to the empirical data, but is also both simpler and tractable in terms of the distance variable. Our derivation of the expression also provides an alternative yet complementary way of approaching and analyzing the problem, compared to the one in [77].

4.2 Background

The most popular point processes in stochastic geometry are PPPs. In general, a point process Φ_{P} is regarded as a PPP if and only if the number of points in any compact region $\mathcal{B} \subset \mathbb{R}^2$ is a Poisson random variable, and the numbers of points in disjoint regions are mutually independent. The points are independently and uniformly distributed over the area of $\mathcal{B}$. The probability distribution for the number $N\left(\Phi_{\mathrm{P}}(\mathcal{B})\right)$ of PPP points falling into any region $\mathcal{B}$ is given by

$$\mathbb{P}\left\{N\left(\Phi_{\mathrm{P}}(\mathcal{B})\right)=n\right\}=\frac{(\lambda_p|\mathcal{B}|)^n}{n!}\exp(-\lambda_p|\mathcal{B}|) \tag{4.1}$$

where λ_p denotes the intensity of Φ_{P}, $|\mathcal{B}|$ is the area of the region $\mathcal{B}$, and $\lambda_p|\mathcal{B}|$ is the average number of points in the region $\mathcal{B}$.

We are interested in the distribution of the distance R between a generic location (not part of the point process) and its closest point in the point process of interest. For the PPP case, the probability that $R > r$ simply equals the null probability of a 2-D Poisson process (i.e. that no points exist in a circular area of πr^2), given by [21]

$$\mathbb{P}[R>r]=\exp\{-\lambda_p\pi r^2\} \tag{4.2}$$

Therefore, the cumulative distribution function of the distance R is $F_R(r) = 1 - \exp\{-\lambda_p\pi r^2\}$, and the probability density function can be obtained as

$$f_R(r)=\frac{\mathrm{d}F_R(r)}{\mathrm{d}r}=2\pi\lambda_p r\exp\{-\lambda_p\pi r^2\} \tag{4.3}$$

As an example of the application of $f_R(r)$, in [21], given the conditional probability of coverage $p_c^R(T)$ (which is derived from the PGFL of a PPP), where R is the desired link distance and T is the SINR threshold (i.e. the minimum SINR to be in coverage),

the average probability of coverage p_c can be obtained as

$$p_c = \int_0^\infty p_c^R(T) f_R(r) \mathrm{d}r \tag{4.4}$$

The above equation can actually be applied to whatever arbitrary kind of point process that is being considered, if the closest-point distance distribution $f_R(r)$ and the PGFL of that point process are known (again using the latter to derive $p_c^R(T)$). Unfortunately, a tractable expression for the PGFL of an MHCPP is presently unknown.

In this thesis, we investigate a specific model of HCPPs, that is, the Matérn hard core point process of Type II [32]. It is constructed by applying dependent thinning to a parent PPP, as described in Section 2.3.3. The MHCPP intensity can be expressed as [32]

$$\lambda_{\text{MH}} = \frac{1 - \exp\{-\lambda_p \pi \delta^2\}}{\pi \delta^2} \tag{4.5}$$

where λ_p is the parent PPP intensity and δ is the hard-core parameter[2].

For ease of reference, we here provide the analytical expression for the CDF of the MHCPP contact distance R as found in [77]:

$$\begin{aligned} &F_R^{\text{PPP}\to\text{MHC}}(r_0|\delta) = \\ &1 - \exp\left(- \int_0^{r_0} 2\pi r \frac{1 - \exp\left[\lambda_p(\ell_2(r,\delta) - \pi\delta^2)\right]}{\pi\delta^2 - \ell_2(r,\delta)} \mathrm{d}r \right) \end{aligned} \tag{4.6}$$

where $\ell_2(r,\delta)$ is

$$\ell_2(r,\delta) = \begin{cases} \pi r^2, & 0 < r < \delta/2 \\ r^2 \cos^{-1}\left(1 - \dfrac{\delta^2}{2r^2}\right) + & r \geq \delta/2 \\ \delta^2 \cos^{-1}\left(\dfrac{\delta}{2r}\right) - \dfrac{1}{2}\delta\sqrt{4r^2 - \delta^2}, & \end{cases} \tag{4.7}$$

[2] δ may be in any arbitrary unit of distance, while λ_p and λ_{MH} would be in the same units of distance^{-2}.

4.3 Single-function model

Because of the correlations among the MHCPP points, it is exceedingly difficult to analytically derive a complete exact expression for the distribution of distance R like the PPP model has. (Even the analytical expression in [77], which is the closest-to-exact expression we are aware of, uses a PPP approximation during its derivation.) Therefore, we instead begin by using various existing distributions to fit the set of empirical samples of R generated from Monte Carlo simulations in order to find an appropriate fitted model for $f_R(r)$. Since the distance R cannot be negative, we only consider distributions that are supported on semi-infinite intervals, usually $[0, \infty)$. When a data set is said to satisfy a specific distribution, it means the data set is consistent with this hypothetical distribution and follows its corresponding properties.

4.3.1 Distributions

In the following, we list the existing distributions that will be examined as the fitting models in this thesis.

Gamma distribution: A gamma random variable X with finite shape parameter $\varepsilon > 0$ and finite scale parameter $\mathrm{n} > 0$ has a probability distribution function

$$f_X(r) = r^{\varepsilon-1} \frac{e^{-\frac{x}{\mathrm{n}}}}{\mathrm{n}^{\varepsilon}\Gamma(\varepsilon)} \tag{4.8}$$

for $r \geq 0$, where $\Gamma(\cdot)$ is the gamma function. The CDF is

$$F_X(r) = \frac{\gamma(\varepsilon, r/\mathrm{n})}{\Gamma(\varepsilon)} \tag{4.9}$$

where $\gamma(r_1, r_2) = \int_0^{r_2} r^{r_1-1} e^{-r} \mathrm{d}r$ is the lower incomplete gamma function.

Log-normal distribution: Given a log-normal distributed random variable X and two parameters μ and σ that are, respectively, the mean and standard deviation of X's

natural logarithm, the PDF of X is defined as

$$f_X(r) = \frac{1}{r\sigma\sqrt{2\pi}} \exp\left\{-\frac{(\ln r - \mu)^2}{2\sigma^2}\right\} \tag{4.10}$$

Rayleigh distribution: The PDF of the Rayleigh distribution is

$$f_X(r;k) = \frac{r}{k^2} \exp\left\{-\frac{r^2}{2k^2}\right\} \tag{4.11}$$

where k is the scale parameter of the distribution. Note especially that if we relate equation (4.11) with (4.3), it can be seen that the distance distribution in a PPP model is actually Rayleigh distributed with $k^2 = 1/\left(2\pi\lambda_p\right)$.

Weibull distribution: The PDF of the Weibull distribution is defined as

$$f_X(r;\tau,k) = \frac{k}{\tau}\left(\frac{r}{\tau}\right)^{k-1} \exp\left\{-\left(\frac{r}{\tau}\right)^k\right\} \tag{4.12}$$

for $r \geq 0$, where $k > 0$ is the shape parameter and $\tau > 0$ is the scale parameter. The CDF is

$$F_X(r;\tau,k) = 1 - \exp\left\{-\left(\frac{r}{\tau}\right)^k\right\} \tag{4.13}$$

for $r \geq 0$. The Weibull distribution is related to a number of other probability distributions. In particular, it interpolates between the exponential distribution ($k = 1$) and Rayleigh distribution ($k = 2$).

4.3.2 Simulation results

In this section, we present the fitting results to the empirical data set. The simulation methodology is as follows. To begin, there is a square sampling window with a side length of L^{W}, centered at the origin. In order to mitigate edge effects, the inner sampling window is then enlarged to a bigger simulation window, with sides multiplied by a factor of ν. Thus, if $\nu = \sqrt{2}$, the whole simulation area will be twice as large,

with on average half the total points falling inside the inner window, and half outside. PPP points are generated within the larger area and thinned to an MHCPP. L^{W} is chosen as $\max(20, 20\delta, 15/\sqrt{\lambda_{\mathrm{MH}}})$ to obtain a reasonably large sampling window with, on average, a minimum of 225 MHCPP points within. Then, we generate 100 locations uniformly over the area of the inner window, and measure the distance to the nearest BS. That sample is kept if the nearest BS is also located within the inner window; otherwise, a new random location is generated until 100 samples are kept. This is repeated for 10000 realizations of the MHCPP, for every pair of (λ_p, δ).

There are various approaches to estimating the parameters of a distribution, such as maximum likelihood estimation (MLE) and linear regression. In this thesis, MLE is used to estimate the parameters for different distributions and the confidence interval is set to be 0.95. The best fitted curve is chosen based on the method of least squares, i.e. choosing the parameters for the curve that yield the least sum of squared deviations between the curve and data.

In Fig. 4.1, we compare the empirically measured PDF with various existing distributions. The gamma, log-normal, Weibull and Rayleigh distributions are plotted with their parameter values chosen to fit as closely as possible to the distribution of the empirical data. As is apparent, among all the examined distributions, the Weibull distribution has the best fit to the empirical PDF $f_R(r)$. Similar results were found for other values of λ_p and δ. A numerical comparison is provided in Table 4.1, in terms of the root mean square error (RMSE). Indeed, the Weibull distribution has the minimum RMSE values while the log-normal has the maximum ones, which strengthens the conclusion of choosing the Weibull distribution.

4.4 Piecewise model

Although the Weibull distribution fits the best to the empirical data out of the distributions examined thus far, there is still a visible deviation between the Weibull function

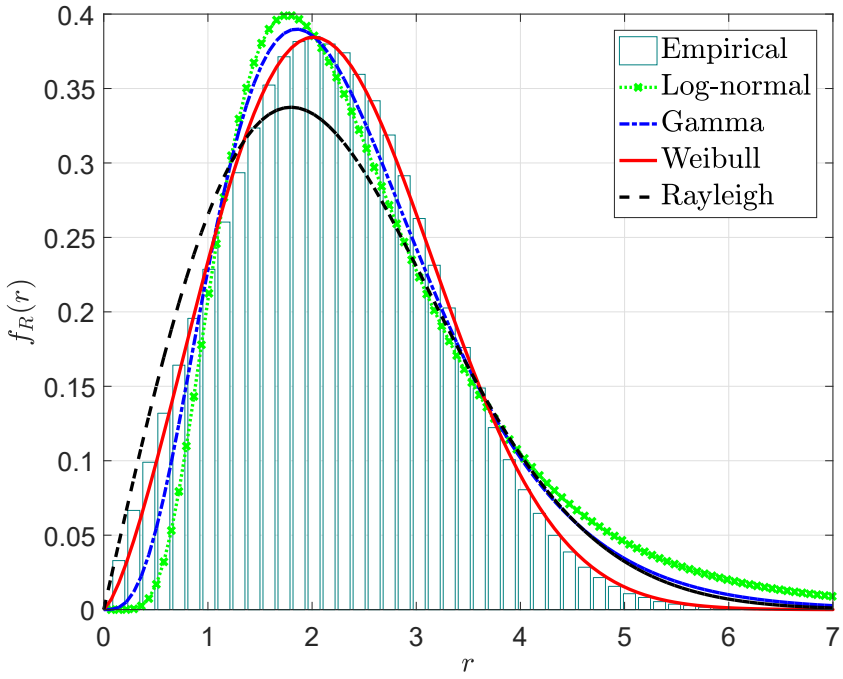


Figure 4.1. The fitting results for the PDF $f_R(r)$ with $\lambda_p = 1$ and $\delta = 3$.

Table 4.1
RMSE Comparison for Different Values of δ, $\lambda_p = 1$

δ	0.2	0.4	0.6	0.8	1	2	3
Weibull	6.8×10^{-3}	9.0×10^{-3}	0.012	0.016	0.016	0.010	7.1×10^{-3}
Rayleigh	7.4×10^{-3}	0.026	0.059	0.077	0.072	0.044	0.029
Gamma	0.057	0.054	0.060	0.067	0.059	0.035	0.023
Log-normal	0.120	0.113	0.111	0.111	0.094	0.054	0.036

and the empirical PDF. However, it turns out we can do even better than simply using the Weibull distribution.

4.4.1 Derivation

Due to the repulsive distance δ, the analytical expression for $f_R(r)$ should be piecewise, i.e. have at least two parts to it[3]. Let us consider a general BS of interest. When

[3]This is bolstered by the fact that the CDF (4.6) from [77] was also found to be piecewise, in the expression for ℓ_2 in (4.7).

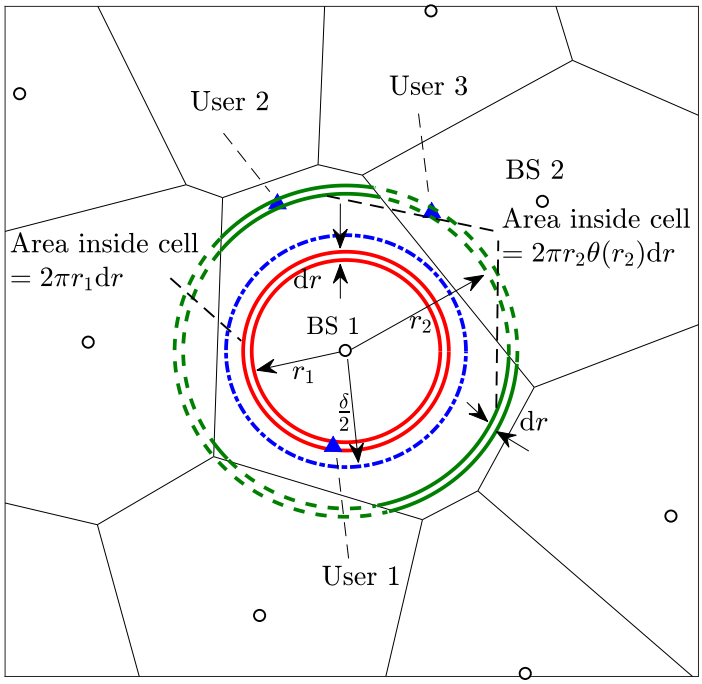


Figure 4.2. Interpretation of the closest distance R between the users and BSs for a typical cell in the MHCPP model.

a user is located at a distance $0 \leq r \leq \delta/2$ from the BS, because of the repulsion of the point process, it is certain that the BS in question is the closest one to the user, as shown by user 1 being closest to BS 1 in Fig. 4.2. Thus, the cumulative distance distribution $F_R(r)$ should be directly related to the probability that the user is located within that distance. With a uniform distribution over the region for the user, that probability should be directly proportional to the area πr^2 of a circle of radius r. Alternatively speaking, the PDF of the user being specifically at the distance r from the BS, where $0 \leq r \leq \delta/2$, is related to the probability of being located within the area of an infinitesimally thin ring of inner radius r and outer radius $r + \mathrm{d}r$. The area of that ring is $2\pi r\mathrm{d}r$, and so $f_R(r) \propto 2\pi r$.

When $r > \delta/2$, things become more complicated. Consider the Voronoi cells

formed by using the points of the MHCPP as generators, and more specifically the cell containing the BS of interest. It is no longer the case that the infinitesimal ring of radius r will be completely located within that cell; segments of the ring can be located in other cells, meaning that for those segments, the closest BS to the user is actually a BS in another cell. This is illustrated in Fig. 4.2 in which user 2 is closest to BS 1 while user 3 is closest to BS 2. The PDF $f_R(r)$ therefore must now become a function of what portion of the ring still remains closest to the BS of interest. Let $\theta(r) \in [0, 1]$ denote the fraction of the infinitesimally thin ring of radius r centered at the BS of the typical cell, wherein a user located in the ring is still the closest to the BS of interest. Note that this fraction may be split across multiple ring segments, as pictured in Fig. 4.2. $\theta(r)$ can be interpreted in several ways: a) the fraction of the infinitesimally thin ring of radius r that is located within the typical cell of interest; b) the probability that the distance from the BS to the border of the typical cell in a random direction is greater than r; c) the probability that, if a user is located at distance r from the BS, then it is located inside the typical cell of interest; d) the probability that, for a given BS of interest, a user at distance r would be associated with that BS in a closest-BS association scheme. The total area of the ring inside the cell is $2\pi r\theta(r)\mathrm{d}r$, and we thus in general have $f_R(r) \propto 2\pi r\theta(r)$. We may define $\theta(r) = 1$ for $0 \leq r \leq \delta/2$. Since $\theta(r)$ is a measure of the typical cell, it can be interpreted as the average fraction of the infinitesimally thin ring of radius r that is located inside the cell, averaged over all possible cells and all possible realizations of the MHCPP.

For one particular realization of an MHCPP and considering one Voronoi cell from that realization, we can sum up all the areas of the infinitesimal ring segments. This is a Riemann sum, and since the ring widths are infinitesimal, in the limit as the widths go to zero and the number of rings goes to infinity, the sum converges to a Riemann integral as follows:

$$\int_0^\infty 2\pi r\hat{\theta}(r)\mathrm{d}r = \hat{A} \tag{4.14}$$

$\hat{\theta}(r)$ is the fraction of the ring of radius r that lies inside this particular cell, and $\hat{A}$ is

the total area of this Voronoi cell. Averaging over all possible realizations and cells of those realizations to obtain an expression for the typical cell, we get

$$\begin{aligned}\mathbb{E}\left\{\int_0^\infty 2\pi r\hat{\theta}(r)\mathrm{d}r\right\} &= \int_0^\infty 2\pi r\mathbb{E}\{\hat{\theta}(r)\}\mathrm{d}r \\ &= \int_0^\infty 2\pi r\theta(r)\mathrm{d}r = \mathbb{E}\{\hat{A}\} = \bar{A}\end{aligned} \tag{4.15}$$

where $\bar{A}$ is the mean area of the typical cell. We do not know the exact distribution of the MHCPP cell area, but fortunately, we do know its average value. For any stationary point process of positive intensity λ, the Voronoi tessellation formed from that point process is also stationary, and λ is also the mean number of cells per unit area; hence the mean area per cell is λ^{-1} [31, Ch. 9]. Thus, for the MHCPP, $\bar{A} = \lambda_{\text{MH}}^{-1}$.

As stated above, we have in general $f_R(r) \propto 2\pi r\theta(r)$, or $f_R(r) = 2\pi cr\theta(r)$ for some c that is not dependent on r. Furthermore, from the properties of a PDF, we must have $\int_0^\infty f_R(r)\mathrm{d}r = \int_0^\infty 2\pi cr\theta(r)\mathrm{d}r = 1$. Comparing this with (4.15), and knowing that $\bar{A} = \lambda_{\text{MH}}^{-1}$, it is clear that $c = \lambda_{\text{MH}}$. We can therefore conclude that:

$$f_R(r) = \begin{cases} 2\pi\lambda_{\text{MH}}r, & 0 \le r \le \delta/2 \\ 2\pi\lambda_{\text{MH}}r\theta(r), & r > \delta/2 \end{cases} \tag{4.16}$$

The probability of two MHCPP points being separated by a distance of exactly δ is 0 almost surely. Similarly, the probability of two points being separated by a distance between δ and $\delta + 2\epsilon$ (so the border between the cells is located between $\delta/2$ and $\delta/2 + \epsilon$) goes to zero as $\epsilon \to 0$. Thus, $f_R(r)$ and $\theta(r)$ should be continuous at $r = \delta/2$, meaning $\lim_{r\to\left(\frac{\delta}{2}\right)^-} \theta(r) = \lim_{r\to\left(\frac{\delta}{2}\right)^+} \theta(r) = \theta\left(\frac{\delta}{2}\right) = 1$, and $\lim_{r\to\left(\frac{\delta}{2}\right)^-} f_R(r) = \lim_{r\to\left(\frac{\delta}{2}\right)^+} f_R(r) = f_R\left(\frac{\delta}{2}\right) = \pi\lambda_{\text{MH}}\delta$.

We strongly suspect, though have not completely formally proven it, that the function in (4.16) is in fact the exact analytical expression for $f_R(r)$ in the range of $0 \le r \le \delta/2$. For $r > \delta/2$ though, we must still resort to an empirical approximation for

$f_R(r)$ and $\theta(r)$.

When $f_R(r)$ is considered piecewise, we have a (presumably) exact PDF of R over the range $[0, \delta/2]$. However, there is no exact expression for $f_R(r)$ when $r > \delta/2$. Yet, interestingly, from simulations we have found the latter piece can be very closely described by a generalized Weibull-like function with the form of $\alpha r^{\beta-1} \exp\{-\varphi r^{\beta}\}$ with positive real parameters α, β, and φ. Compared to (4.16), $\theta(r)$ therefore can be approximated as a function with the form $\theta(r) = \alpha/(2\pi\lambda_{\text{MH}}) r^{\beta-2} \exp\{-\varphi r^{\beta}\}$.

Proposition 4.1: The PDF $f_R(r)$ for the closest distance R, when $r > \delta/2$, can be fitted to the function $\alpha r^{\beta-1} \exp\left(-\varphi r^{\beta}\right)$. The PDF also has the requirements of continuity of value at $r = \delta/2$ and that $\int_0^{\infty} f_R(r) \mathrm{d}r = 1$. Given these factors, the PDF $f_R(r)$ can be approximated using only the parameter β and the MHCPP parameters λ_{MH} and δ as:

$$f_R(r) = \begin{cases} 2\pi\lambda_{\text{MH}} r, & 0 \le r \le \delta/2 \\ 2\pi\lambda_{\text{MH}} r \left(\dfrac{2r}{\delta}\right)^{\beta-2} \times & r > \delta/2 \\ \quad \exp\left\{\dfrac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4 - \pi\lambda_{\text{MH}}\delta^2)} \left(1 - \left(\dfrac{2r}{\delta}\right)^{\beta}\right)\right\}, & \end{cases} \tag{4.17}$$

and therefore

$$\theta(r) = \begin{cases} 1, & 0 \le r \le \delta/2 \\ \left(\dfrac{2r}{\delta}\right)^{\beta-2} \times & r > \delta/2 \\ \quad \exp\left\{\dfrac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4 - \pi\lambda_{\text{MH}}\delta^2)} \left(1 - \left(\dfrac{2r}{\delta}\right)^{\beta}\right)\right\}, & \end{cases} \tag{4.18}$$

$Proof$: See Appendix A.1.

Corollary 4.1: The CDF $F_R(r)$ of the distance R can be approximated as:

$$F_R(r) = \begin{cases} \pi\lambda_{\text{MH}} r^2, & 0 \le r \le \delta/2 \\ 1 - \left(\dfrac{4 - \pi\lambda_{\text{MH}}\delta^2}{4}\right) \times & r > \delta/2 \\ \quad \exp\left\{\dfrac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4 - \pi\lambda_{\text{MH}}\delta^2)}\left[1 - \left(\dfrac{2r}{\delta}\right)^{\beta}\right]\right\}, & \end{cases} \tag{4.19}$$

Proof: For $0 \le r \le \delta/2$, $F_R(r) = \int_0^r 2\pi\lambda_{\text{MH}} r \mathrm{d}r = \pi\lambda_{\text{MH}} r^2$. Let $C = \frac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4-\pi\lambda_{\text{MH}}\delta^2)}$ and $t = \frac{2r}{\delta}$. For $r > \delta/2$, we have

$$\begin{aligned} F_R(r) &= \int_0^{\delta/2} f_R(r)\mathrm{d}r + \int_{\delta/2}^{r} f_R(r)\mathrm{d}r \\ &= \frac{\pi\lambda_{\text{MH}}\delta^2}{4} + \int_1^{\frac{2r}{\delta}} \pi\lambda_{\text{MH}}\delta \exp\{C\} t^{\beta-1} \exp\left\{-Ct^{\beta}\right\} (\delta/2)\mathrm{d}t \\ &= \frac{\pi\lambda_{\text{MH}}\delta^2}{4} - \frac{\pi\lambda_{\text{MH}}\delta^2}{2C\beta} \exp\left\{C\left(1 - t^{\beta}\right)\right\}\Big|_1^{2r/\delta} \\ &\overset{(i)}{=} 1 - \left(\frac{4 - \pi\lambda_{\text{MH}}\delta^2}{4}\right) \exp\left\{\frac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4 - \pi\lambda_{\text{MH}}\delta^2)}\left[1 - \left(\frac{2r}{\delta}\right)^{\beta}\right]\right\} \end{aligned} \tag{4.20}$$

where the last step (i) is simply obtained from the substitution of C and the rearrangement of terms.

4.4.2 Simulation results

When simulating $\theta(r)$, for every realization, we first choose 100 BSs at random that are located within the inner window, each with equal probability of being chosen. For each one, we then find the distance d in 360 directions, in evenly-spaced degree increments, from the BS to the cell border. After all samples are taken, the function $\theta(r)$ is equivalent to the probability those distances are greater than the value of r. An alternative method is to take a ring of 360 points around the BS, evenly-spaced in angle, with radius r, then counting the number of points located within the cell. The fraction

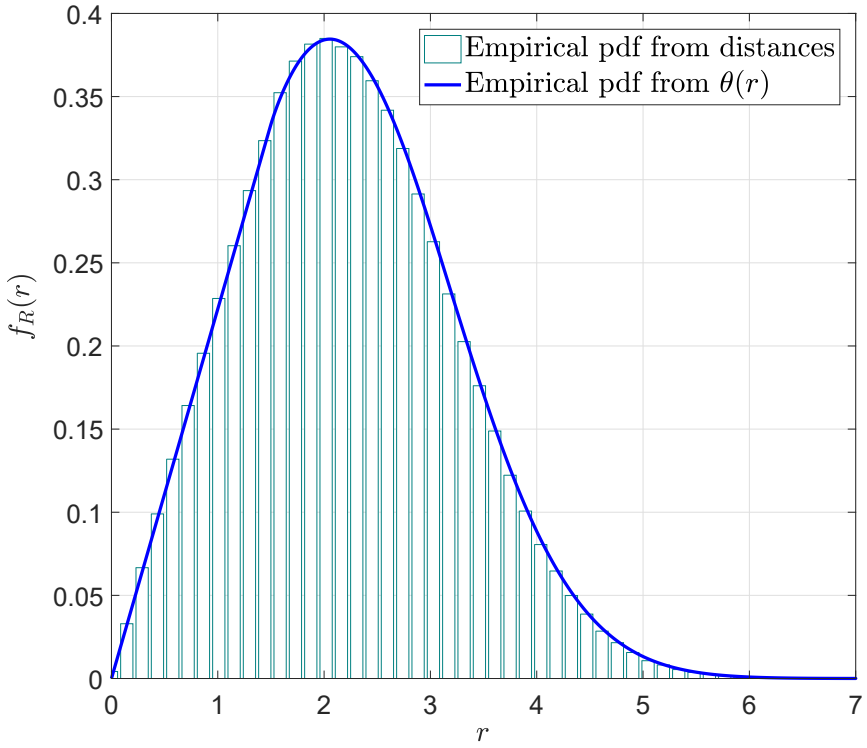


Figure 4.3. Comparison between empirical PDFs found from measuring distances and from measuring $\theta(r)$, $\lambda_p = 1$ and $\delta = 3$.

of ring points that are inside the cell is $\theta(r)$. Both methods give identical results.[4] In Fig. 4.3, we compare the empirical PDF as found directly by distance measurements, and as found by measurements of $\theta(r)$ and plotting $2\pi\lambda_h r\theta(r)$. It can be seen that the empirical PDF essentially coincides for both methods. The same holds true for other values of λ_p and δ, which demonstrates the validity of the formulation for $f_R(r)$ using $\theta(r)$.

In Fig. 4.4, we compare the PDF curves between the empirical data and the piecewise model (4.17) given the best-fit values of β. It shows that the differences between the empirical PDFs (where $f_R(r)$ is now based on the measured data for $\theta(r)$) and the corresponding piecewise models are very small. There still remains a small deviation

[4]For both $f_R(r)$ and $\theta(r)$, taking multiple samples per realization is also important to avoid the "zero-cell" biasing effect [31, Ch. 9.3.3]. Taking just a single cell per realization, e.g. one containing the origin or some random location, means that cell on average will be larger than the typical cell, since it is more likely for the location to be within a larger cell than a smaller one.

Table 4.2
RMSE Values $\left(\times 10^{-4}\right)$ Comparison for Different Values of λ_p and δ

(λ_p, δ)	$(10^{-4}, 0.2)$	$(0.2, 0.2)$	$(0.2, 0.6)$	$(0.2, 3)$	$(0.6, 0.2)$	$(0.6, 0.6)$	$(0.6, 3)$	$(3, 0.2)$	$(3, 0.6)$	$(3, 3)$	$(10^3, 1)$
Single Weibull	0.11	2.2	14	55	8.2	53	55	84	250	55	170
Piecewise model	0.12	2.6	16	4.1	9.1	48	4.6	95	14	3.8	13

between the empirical curve and the piecewise model around the peak of the curve, but the overall deviation is not as large as for a single-function Weibull distribution, especially at the leftmost part of the curve. Therefore, given the values of the parameter β, the PDF $f_R(r)$ can be very conveniently and reasonably accurately calculated using the proposed piecewise model (4.17). For a numerical comparison, in Table 4.2, we compare the goodness-of-fit between the single Weibull-fit PDF and piecewise model PDF in terms of RMSE for different values of λ_p and δ. As expected, the proposed piecewise model is generally even more closely fitted to the empirical data than the single Weibull distribution. However, for small parameter values (especially for δ), the single Weibull function can still have a slightly smaller RMSE, since under these conditions, the Weibull distribution (and the MHCPP itself) reduce closely to the parent PPP. The leftmost and rightmost columns respectively represent the cases of extremely sparse and extremely dense parent processes, and demonstrate that the piecewise model still works well under those extreme settings.

In Fig. 4.5, we show the estimated values of the parameter β fitted from the empirical data for different MHCPP parameter values. With a fixed λ_p, the value of β increases as the hard-core parameter δ grows, eventually reaching an asymptotic value[5] around 2.467. The larger the value of λ_p, the sooner β approaches its asymptotic value. With a fixed δ, the value of β increases as the value of λ_p grows. When λ_p is sufficiently large, β remains almost unchanged even if λ_p continues to grow, which means the MHCPP model becomes a somewhat more regularly-spaced model[6] and the

[5]Interestingly, this is very close to the value $\pi^2/4$, though this likely may simply be coincidental.

[6]Although there are always variations in the distances between BSs even when λ_p is sufficiently large, the distribution of distance R between a generic location and its closest MHCPP point remains almost unchanged with changing λ_p.

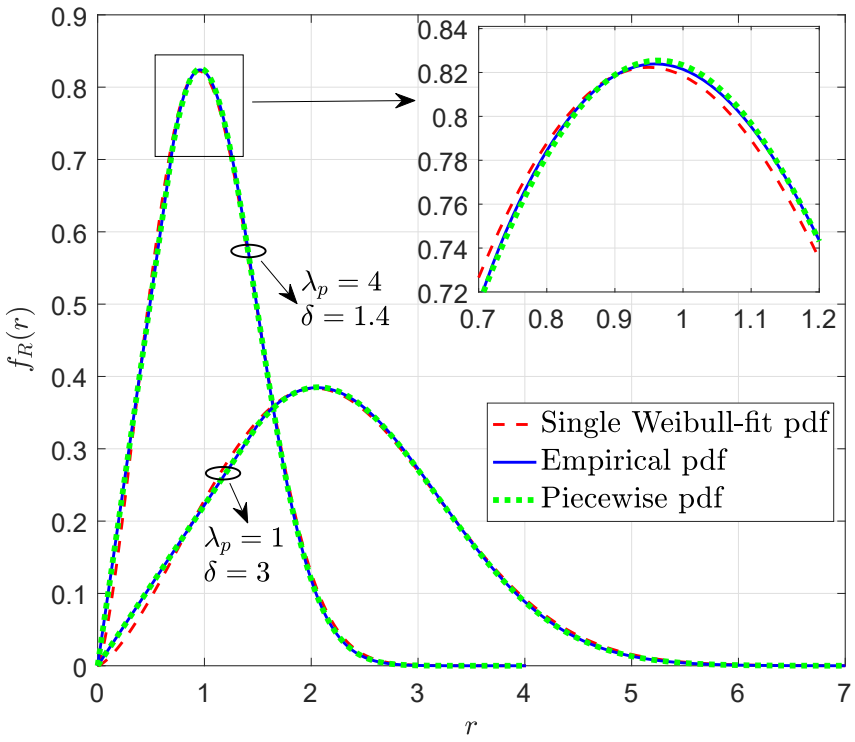


Figure 4.4. Comparison between the empirical data and the piecewise model given the values of β.

MHCPP intensity λ_{MH} becomes asymptotically constant at $1/\pi\delta^2$ as the numerator of λ_{MH} goes to 1. When δ and λ_p tend towards 0, the value of β approaches 2.

Remark 4.1: As β approaches 2 for small values of λ_p and δ, the piecewise distribution reduces to a Rayleigh distribution, with approximately the same scale parameter value as for a PPP distribution. That is, $\lambda_{\text{MH}} \approx \lambda_p$, and $f_R(r)$ in (4.17) reduces to (4.3) for all values of r.

Proof: See Appendix A.2. For small values of λ_p and δ (or more specifically, the product $\lambda_p\pi\delta^2$), almost all the parent points of the MHCPP are retained since they are, for the most part, separated by distances of at least δ to begin with. Thus, the MHCPP asymptotically acts very much like its parent PPP. This result is consistent with that found in previous work, e.g. [10]. However, based on the RMSE values in Table 4.2, the piecewise function may converge to the Rayleigh distribution slightly weaker than

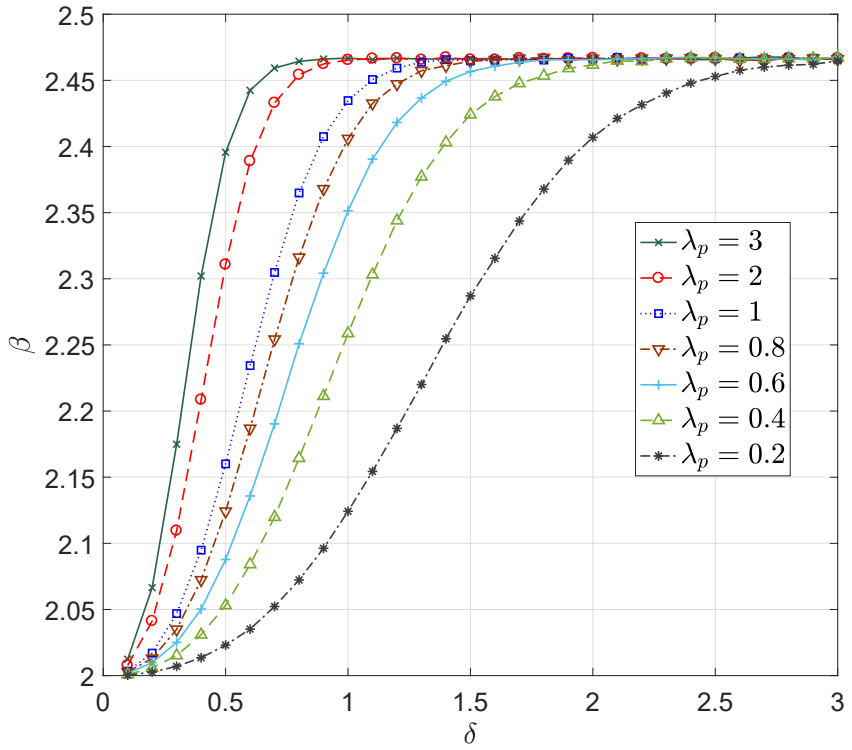


Figure 4.5. Estimated values of the parameter β for the piecewise function in (4.17).

a single Weibull function does.

In Fig. 4.6(a), we plot the best-fit values of β vs. $\lambda_{\text{MH}}\pi\delta^2$. Interestingly, we find that the curve can be reasonably closely described by the following quadratic polynomial model:

$$\beta \approx 0.3686\chi^2 + 0.0985\chi + 2, \text{ where } \chi = \lambda_{\text{MH}}\pi\delta^2 \tag{4.21}$$

The RMSE between the β values and the quadratic polynomial model is calculated to be 1.7×10^{-3}. With this approximation, the variable β can be replaced, and the distribution $f_R(r)$ in (4.17) can be expressed using solely the MHCPP parameters λ_{MH} and δ.

In Fig. 4.6(b), we plot the best-fit values of β vs. $\log_{10}(\lambda_p\pi\delta^2)$. This graph simply graphically demonstrates and confirms the asymptotic conditions for the β values that we have numerically described earlier. The lower asymptote of 2 occurs for values

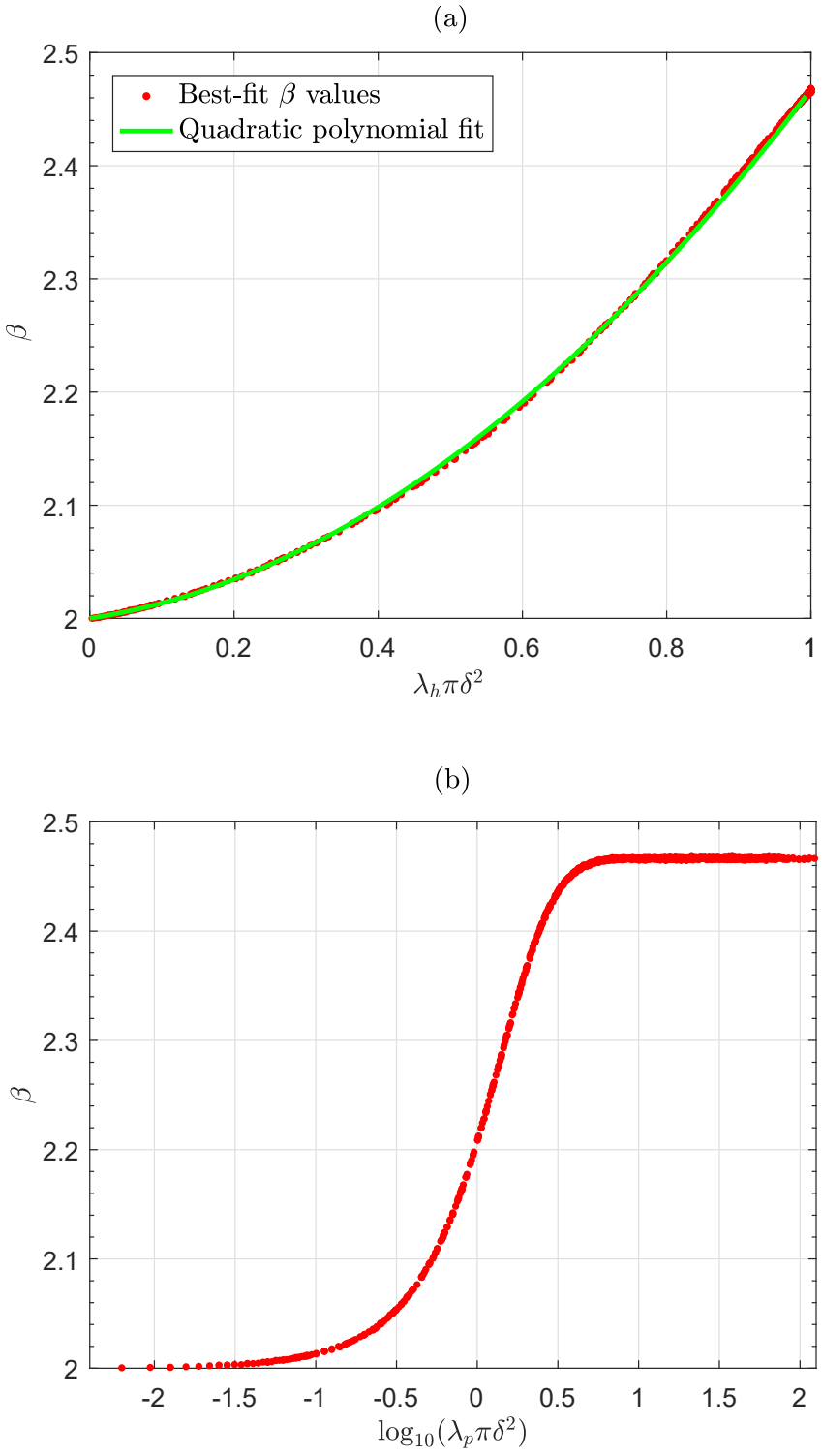


Figure 4.6. Best-fit β vs. $\lambda_{\text{MH}}\pi\delta^2$ and $\log_{10}(\lambda_p\pi\delta^2)$, respectively.

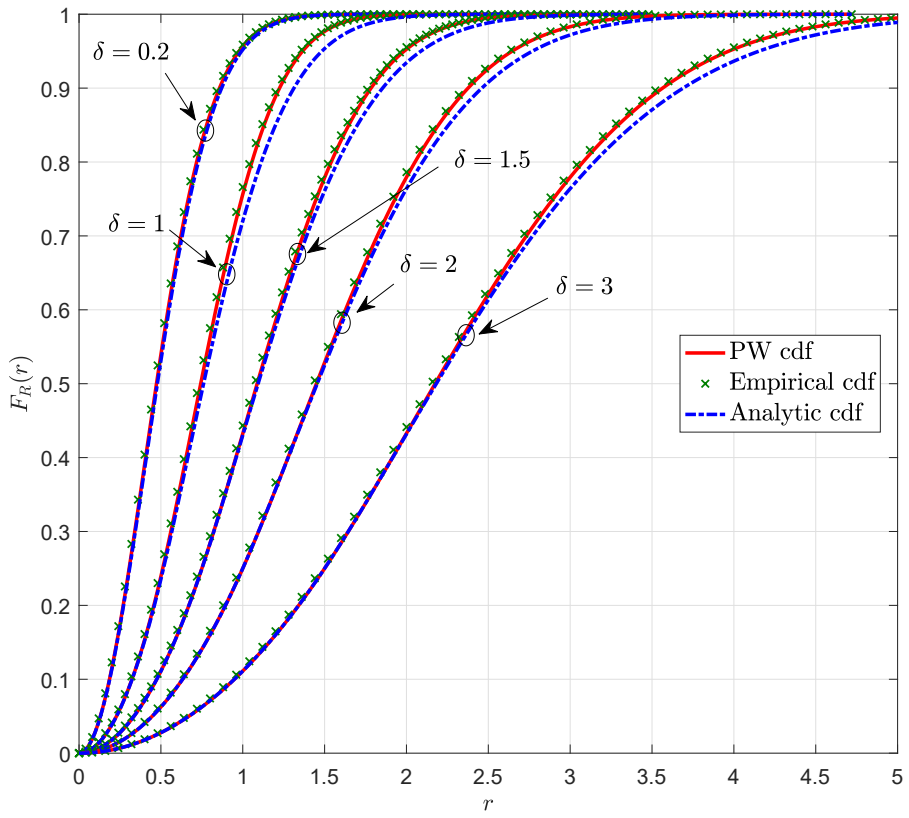


Figure 4.7. Comparison of empirical CDF with piecewise (PW) model CDF in (4.19) and analytical CDF from [77], λ_p = 1.

of $\lambda_p \pi \delta^2 \ll 1$ (e.g. starting around 2 orders of magnitude smaller than 1), which is reflected in the graph for $\log_{10}(\lambda_p \pi \delta^2)$ around -2 (or slightly larger). Meanwhile, the upper asymptote occurs when the numerator of λ_{MH} approaches 1, or $\exp\{-\lambda_p \pi \delta^2\} \ll 1$. This is equivalently roughly around $\lambda_p \pi \delta^2 > -\ln(0.01)$, or $\log_{10}(\lambda_p \pi \delta^2) > 0.6632$, as seen in the graph. As an aside, Fig. 4.6(b) could also be fit reasonably well using a generalized logistic function, although we have declined to do so here, due to the simpler fitting vs. $\lambda_{\text{MH}} \pi \delta^2$ in Fig. 4.6(a).

In Fig. 4.7 and Table 4.3, we compare our piecewise CDF in (4.19) and the analytical CDF from [77] in (4.6)-(4.7) with the empirical CDF from simulations. It can be seen that our piecewise equation fits the empirical data closer than the analytically-

Table 4.3
RMSE Comparison for PW Model CDF and Analytical CDF, $\lambda_p = 1$

δ	0.2	1	1.5	2	3
Piecewise CDF	3.3×10^{-3}	2.5×10^{-3}	1.8×10^{-3}	1.2×10^{-3}	6.9×10^{-4}
Analytical CDF	7.1×10^{-3}	0.019	0.010	8.5×10^{-3}	8.3×10^{-3}

derived expression, especially as λ_p and δ increase. The PPP approximation in the conditional thinning probability [77, eq. (14)] neglects the effect of the points thinned from the parent PPP to form the MHCPP. (The thinned points are referred to as the complementary Matérn hard core (CMHC) process.) This results in an underestimation of the thinning probability and thus the overestimation of the distances in the distribution (or, equivalently, an underestimation of the distribution at a given distance). The conditional thinning of an MHCPP point given a CMHC point is indeed weak, as stated in [77], which is reflected in the analytical CDF being close to the empirical CDF. However, the effect of CMHC points in total is still sufficient to, for instance, add up to about a 5-10$\%$ overestimation of the distance at the 90th percentile of the CDF in some cases[7]. Note also that the MHCPP retention probability (given by $\lambda_{\text{MH}}/\lambda_p$ [31], [77, eq. (12)]) goes to zero as λ_p and δ increase, meaning that points that are thinned from the parent PPP (i.e. the CMHC points) make up by far the vast majority of the parent PPP points.

We also note that, along with fitting better to the empirical CDF, our piecewise CDF of (4.19) is both tractable and simpler than the analytical expressions of (4.6)-(4.7). Recall that the second part of our piecewise PDF in (4.17) is based on, and thus reduces to, an equation of the form $\alpha r^{\beta-1}\exp\left\{-\varphi r^{\beta}\right\}$, with α and φ being positive constants involving the parameters λ_{MH}, δ, and β. Similarly, the second part of (4.19),

[7]Interestingly, for the main result of [77], i.e. the distribution of the contact distance between an MHCPP point and its nearest neighbouring point from the same process, the tightness of the fit between the derived analytical expression and the empirical CDF appears much better. We speculate that this is because the conditional thinning probability in this case [77, eq. (10)] is a fraction, with the PPP approximation performed in both the numerator and denominator. Hence, the effect of the approximation might be close to proportional in the numerator and denominator, and thus mostly cancel itself out.

although appearing complex, also simply reduces to an expression of the form $1 - \mathrm{c}_1 \exp\left\{-\mathrm{c}_2 r^{\beta}\right\}$, with c_1 and c_2 again being positive constants involving λ_{MH}, δ, and β. (In both cases, β can be further approximated by (4.21).)

4.5 Summary

In this chapter, based on empirical data from simulations, we have found that the Weibull distribution is the best fit (among the other existing distributions examined, namely gamma, log-normal and Rayleigh) to the distribution of the distance between a generic location and its closest point in an MHCPP of Type II, if a single function is used for the fitting. We have also proposed a better piecewise model for the PDF, including an analytically derived formula for $r \in [0, \delta/2]$ and an approximate Weibull-like function for $r \in (\delta/2, \infty)$. It is highly likely that the piecewise equation for $r \in [0, \delta/2]$ is the exact analytical expression for the PDF in that range of r, though we have not completely formally proved this. The piecewise model expresses the distance distribution in terms of the MHCPP parameters λ_{MH} and δ, along with a third parameter β. Simulation results show that the proposed piecewise model has a very close goodness-of-fit to the empirical PDF. We have also plotted the best-fit values of the parameter β against different values and combinations of the variables λ_{MH}, λ_p and δ. It turns out the value of β can be closely fitted to a quadratic polynomial model. Therefore, a more general expression for the distance distribution and/or all the parameters of the piecewise model PDF can be expressed in terms of any arbitrary values of only the parameters δ and λ_{MH} (or λ_p). Both the piecewise PDF and its CDF version are tractable and reduce to relatively simple functions of the distance variable. The CDF version of the model also fits better to the empirical data than the analytical expression derived in [77].

Chapter 5

Downlink Coverage Analysis of N-Tier Heterogeneous Cellular Networks Using Clustered Stochastic Geometry

In the analysis of cellular networks, the underlying BS distribution is almost ubiquitously assumed to be the homogeneous Poisson point process. However, in more practical scenarios, especially for small cell BSs, the distributions are more likely to be based on clustered processes. For example, the urban A point pattern in Chapter 3 was found to be best modelled using a Thomas cluster process. The clustering of nodes may be due to geographical restrictions or artificially induced by MAC protocols. This motivates the need to extend the abundant work assuming homogenous PPPs to other point process models.

In this chapter, under the assumption of a connected base station at the point of reference (or origin) in the tier of interest, we derive an expression for the downlink probability of coverage over a heterogeneous network, wherein the base station locations result from different point processes, such as Poisson point and Poisson cluster processes. Numerical results show increasing the base station density in each tier lowers the coverage probability, but not as much as increasing the coverage threshold does.

We also provide lower and upper bounds for the coverage probability in a two-tier heterogeneous network modelled with Poisson point and cluster processes, and evaluate the effect on those bounds when changing the various parameter values. These results are of potential use for future cellular heterogeneous network designs.

5.1 Introduction

Cellular networks are approaching an important transition from homogeneity to heterogeneity due to the rapid proliferation of lower-power and low-cost radio access nodes that operate in both licensed and unlicensed spectra, like picocells, femtocells, microcells, metrocells, and Wi-Fi nodes. A major challenge in deploying cellular heterogeneous networks (HetNets) is how to characterize the statistics of cumulative interference, from which performance metrics like coverage probability (or outage probability), capacity, and throughput can be derived. For the analysis and design of interference management techniques in HetNets, rigorous yet tractable interference models are required. However, interference modelling has always been an arduous task even in simple traditional single-tier networks with base stations (BSs) following a regular hexagonal grid, not to mention HetNets with more topological randomness and diverseness in each tier.

A new modelling method has been widely used for interference analysis in heterogeneous networks by adopting stochastic geometry and averaging over all network topologies seen from a generic node weighted by their probability of occurrence. It not only captures the topological randomness, but also leads to tractable analytical results. In stochastic geometry, locations of BSs are abstracted to a matching point process according to the network type as well as MAC layer behavior. The most common point processes used in modelling and analysis of wireless communication systems are Poisson point processes (PPPs) [21, 22, 34, 78], hard core point processes [14, 15], and Poisson cluster processes (PCPs) [24]. Among these point processes, the PPP is

the most important due to the fact that PPPs not only provide the baseline model for different point processes, but also this model is the most tractable. One approach of interest for system analysis is presented in [21], where system performance metrics like coverage probability and ergodic data rate are obtained by averaging all spatial scenarios for a single-tier network. The analysis is extended to multi-tier cellular networks in [22, 78]. PPPs have been used for modelling co-channel interference in large-scale ad-hoc networks in [25, 26, 79] as well as other generic settings in [80, 81].

PPP-based models are well studied and understood in the literature of wireless communication systems. Nevertheless, in reality the locations of BSs are either clustered or not uniformly distributed. For instance, picocells are more likely to be clustered in crowded streets and femtocells are more likely to be deployed in a building. This motivates us to extend the abundant collection of results available for PPPs to other node distributions, such as a PCP [31, 41], which provides an important model for spatially aggregated point patterns. In general, a PCP generation mechanism involves "parent" and "offspring" points, and it consists of two stages: (a). Parent locations (not included as actual points) are generated as a PPP with intensity κ in the plane $\mathbb{R}^2$; (b). Offspring points are generated i.i.d. for each parent location and scattered around the parent locations according to some density function $f(x)$.

In Fig. 5.1, two kinds of distributions for a two-tier HetNet are depicted. In both cases, macrocell BS locations (denoted by red squares) are modelled according to a PPP. On the left, picocell BSs (blue dots) are also distributed by an independent PPP (i.e. a PPP-PPP model); on the right, they are distributed according to a PCP (i.e. a PPP-PCP model).

5.2 System Model

In this thesis, we modify the downlink model presented in [22]. Locations of BSs are not only determined by PPP, but also by other spatial point processes. For the i^{th} tier

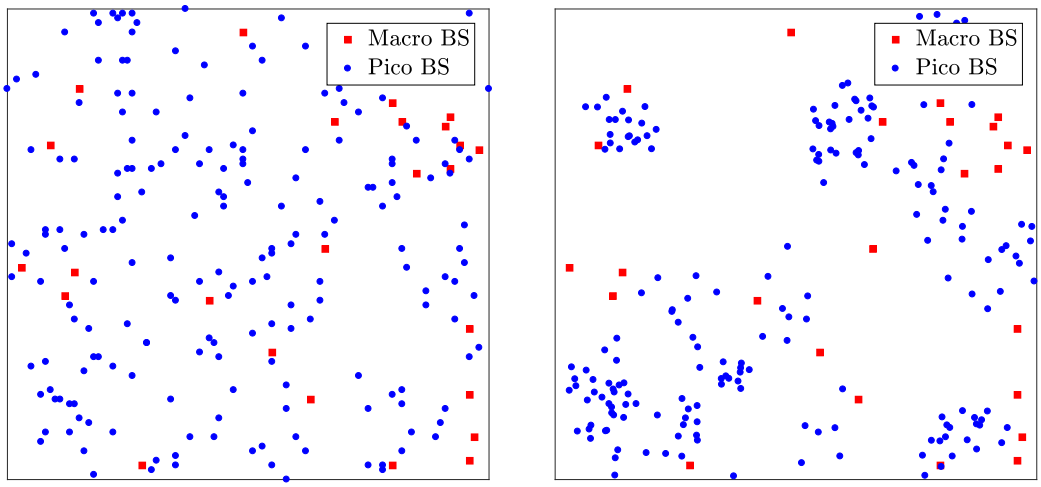


Figure 5.1. Two distribution models for a two-tier HetNet. Left: PPP-PPP model. Right: PPP-PCP model.

Φ_i of interest we assume a BS at the point of reference (or origin) o in the network connected with a mobile user at a location z; hence $\|z\|$ denotes the distance from the BS to its connected user. Similarly to [22], the resulting SINR at the mobile user receiver at the location z in the i^{th} tier can be expressed as:

$$\text{SINR}_i(z) = \frac{P_i h_z \ell(z)}{\Omega + I_{\text{c}}} = \frac{P_i h_z \ell(z)}{\Omega + \sum_{j=1}^{N} \sum_{x \in \{\Phi_j \setminus o\}} P_j h_x \ell(x - z)} \tag{5.1}$$

where h_z and h_x denote the exponentially distributed power gain with mean $1/\mu_h$ due to Rayleigh fading between the BSs and the user at the location z, P_j denotes the transmit power of each BS in the j^{th} tier, $\ell(x) = (1 + \|x\|^{\eta})^{-1}$ models the signal attenuation due to path loss [24], Ω is the additive noise power, and I_{c} denotes the cumulative interference power from all other BSs except the connected BS at the origin o. I_{c} is a stochastic process that depends on the locations of the interferers captured by the point process and random channel gains.

A user can connect to a BS in the i^{th} tier if $\text{SINR}_i(z) > T_i$, where T_i denotes

the SINR threshold of the i^{th} tier. The user is said to be in coverage, if it is able to connect to at least one tier out of N available. Hence, the probability of coverage (now dropping the variable z from SINR_i for ease of notation) is:

$$p_c = 1 - \prod_{i=1}^{N} \left(1 - \mathbb{P}\left(\text{SINR}_i > T_i\right)\right) \tag{5.2}$$

The probability $\mathbb{P}\left(\text{SINR}_i > T_i\right)$ can be expressed as [22]:

$$\mathbb{P}\left(\text{SINR}_i > T_i\right) = \mathcal{L}_{I_c}\left(\frac{\mu_h T_i}{P_i \ell(z)}\right) \cdot \exp\left(-\frac{\mu_h T_i \Omega}{P_i \ell(z)}\right) \tag{5.3}$$

where $\mathcal{L}_{I_c}$ stands for the Laplace transform of the cumulative interference. Let $s = \mu_h T_i / P_i \ell(z)$, then similarly to [22]:

$$\mathcal{L}_{I_c} = \mathbb{E}_o^{!}\left(\prod_{x \in \{\Phi_i \setminus o\}} \frac{1}{1 + s P_i \ell(x)}\right) \cdot \prod_{j=1, j \neq i}^{N} \mathbb{E}_{\Phi_j}\left(\prod_{x \in \{\Phi_j\}} \frac{1}{1 + s P_i \ell(x)}\right) \tag{5.4}$$

where $\mathbb{E}_o^{!}$ denotes the expectation with respect to the reduced Palm measure [31,41]. It is the conditional expectation for point processes given there is a point of the process at the origin, but without including that point.

5.3 Analysis of Probability of Coverage

For the following work, we consider two-tier HetNets with two specific topologies: PPP-PPP and PPP-PCP. Both are assumed to be interference-limited (i.e. Ω is negligible in comparison to I_c). The resulting probability of coverage can be expressed as:

$$p_c = \mathbb{P}\left(\text{SINR}_1 > T_1\right) + \mathbb{P}\left(\text{SINR}_2 > T_2\right) - \mathbb{P}\left(\text{SINR}_1 > T_1\right)\mathbb{P}\left(\text{SINR}_2 > T_2\right) \tag{5.5}$$

For the PPP-PPP scenario, macrocell and picocell BS locations are both modelled

as independent PPPs with different parameter tuples $\left\{\lambda_p^{(1)}, T_1, P_1\right\}$ and $\left\{\lambda_p^{(2)}, T_2, P_2\right\}$, respectively. According to Slivnyak's theorem [31, 41], the reduced Palm distribution of a PPP (not including the point at the origin) is equal to that of the original distribution (which includes that point). Thus, the probability generating functional (PGFL) [31, 41] $\mathcal{G}(\ell)$ and the conditional PGFL $\mathcal{G}^*(\ell)$ of a PPP with density λ_p with regard to the real-valued function $f(x)$: $\mathbb{R}^2 \to [0, 1]$ are given by:

$$\mathcal{G}(f(x)) = \mathcal{G}^*(f(x)) = \exp\left(-\lambda_p \int_{\mathbb{R}^2} [1 - f(x)]\, \mathrm{d}x\right) \tag{5.6}$$

Let $s = \dfrac{\mu_h T_i}{P_i \ell(z)}$ and $\|z\| = R$, then similarly to [21, 82], the Laplace transform of the single-tier interference I_c^i can be expressed as

$$\mathcal{L}_{I_c^i}(s) = \mathbb{E}\left[\exp\left(-sI_c^i\right)\right] = \mathbb{E}_{\Phi_i}\left[\prod_{x\in\Phi_i} \mathbb{E}_{h_x}\left\{\exp\left(-sP_i h_x \ell(x - z)\right)\right\}\right] \tag{5.7}$$

By taking the average over the exponentially fading distribution, we have

$$\mathcal{L}_{I_c^i}(s) = \mathbb{E}_{\Phi_i}\left[\prod_{x\in\Phi} \frac{\mu_h}{\mu_h + sP_i\ell(x - z)}\right] \tag{5.8}$$

Since Φ is a Poisson point process, we obtain

$$\begin{aligned}
\mathcal{L}_{I_c^i}(s) &= \mathbb{E}_{\Phi_i}\left[\prod_{x\in\Phi} \frac{\mu_h}{\mu_h + sP_i\ell(x - z)}\right] \\
&\overset{(i)}{=} \exp\left(-2\pi\lambda_p \int_0^\infty \left\{1 - \frac{\mu_h}{\mu_h + sP_i\ell(r)}\right\} r\mathrm{d}r\right) \\
&= \exp\left(-\csc\left(2\pi/\eta\right) \frac{2\pi^2\lambda_p s P_i \eta^{-1}}{\left(\mu_h + sP_i\right)^{1-2/\eta}}\right)
\end{aligned} \tag{5.9}$$

where the step (i) is obtained from (5.6) and transformation to polar coordinates.

Hence

$$\mathbb{P}\left(\text{SINR}_i > T_i\right) = \prod_{i=1}^{2} \mathcal{L}_{I_{\text{c}}^i}\left(\frac{\mu_h T_i}{P_i \ell(z)}\right) \tag{5.10}$$

For the PPP-PCP scenario, macrocell BS locations are still distributed according to a PPP denoted by Φ_p with parameter tuple $\left\{\lambda_p^{(1)}, T_1, P_1\right\}$, while picocell BS locations are distributed as a specific category of PCP, i.e., a Thomas cluster process (TCP) denoted by Φ_T with parameter tuple $\left\{\lambda_t^{(2)}, T_2, P_2\right\}$. For a TCP, the number of offspring points in each cluster is Poisson-distributed with mean μ, and the points are isotropically scattered around the parent locations according to a symmetric normal distribution with standard deviation σ [31,41]:

$$f(x) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|x\|^2}{2\sigma^2}\right) \tag{5.11}$$

The PGFL and the conditional PGFL of a PCP are obtained from [24, 31, 41], respectively, and can be expressed as

$$\mathcal{G}(\ell) = \mathbb{E}\left(\prod_{x\in\Phi_{\text{PCP}}}\right) = \exp\left(-\kappa \int_{\mathbb{R}^2} \left[1 - \mathcal{M}\left(\int_{\mathbb{R}^2} \ell(x+y) f(y) \mathrm{d}y\right)\right] \mathrm{d}x\right) \tag{5.12}$$

$$\begin{aligned}\mathcal{G}^*(\ell) = \mathbb{E}_o^!\left(\prod_{x\in\Phi_{\text{PCP}}}\right) = \exp\left(-\kappa \int_{\mathbb{R}^2} \left[1 - \mathcal{M}\left(\int_{\mathbb{R}^2} \ell(x+y) f(y) \mathrm{d}y\right)\right] \mathrm{d}x\right) \\ \times \int_{\mathbb{R}^2} \mathcal{M}\left(\int_{\mathbb{R}^2} \ell(x-y) f(x) \mathrm{d}x\right) f(y) \mathrm{d}y\end{aligned} \tag{5.13}$$

where κ is the parent intensity, and $\mathcal{M}$ is the moment generating function of the number of points generated in any cluster of a PCP and $\ell(x)$ is the signal attenuation function. Due to the complexity of the multiple integrals, it is not easy to calculate an exact value of $\mathbb{P}\left(\text{SINR}_i > T_i\right)$ for a TCP-based distribution of BSs. Consequently, we employ a lower bound based off of Lemma 5 in [24] and an upper bound from Lemma 6 in [24]

to facilitate the computation. From the two lemmas, we get:

$$\mathbb{P}_{\text{PPP}}(\lambda_t)\mathbb{P}_{\text{PPP}}(\mu \cdot f^*) \leq \mathbb{P}_{\text{PCP}} \leq \mathbb{P}_{\text{PPP}}(\lambda_t/(1+\mu \cdot \mathcal{T})) \tag{5.14}$$

where $\mathbb{P}_{\text{PPP}}(\lambda_t)$ denotes the probability of coverage when a PPP with the same intensity of the TCP is applied. In general, we have $f^* \leq 1/\left(4\pi\sigma^2\right)$ for a TCP [24] and $\mathcal{T}$ is derived through Hölder's inequality [83].

Hölder's inequality. Let $1/p + 1/q = 1$, with $p, q > 1$. Then Hölder's inequality for integrals states that

$$\int_a^b \|f(x)g(x)\|\mathrm{d}x \leq \left[\int_a^b \|f(x)\|^p\right]^{1/p} \left[\int_a^b \|g(x)\|^q\right]^{1/q} \tag{5.15}$$

with equality when

$$\|g(x)\| = c\|f(x)\|^{p-1} \tag{5.16}$$

where $c \geq 0$. If $p = q = 2$, Hölder's inequality becomes Schwarz's inequality [83].

Let $T = T_1 = T_2$, and

$$\mathcal{P} = \int_{\mathbb{R}^2} \frac{\ell(x)}{T^{-1}\ell(\|z\|) + \ell(x)}\mathrm{d}x \tag{5.17}$$

By Hölder's inequality, the calculation of moment generating function $\mathcal{M}$ can be relaxed by [83]

$$\mathcal{T} \leq \min\{1, \hat{f} \cdot \mathcal{P}\} \tag{5.18}$$

where $\hat{f} = \sup_{x\in\mathbb{R}^2} f(x)$.

5.4 Numerical Results

For the numerical evaluation, we consider BS transmitters and receivers each with only a single antenna. Macro BS transmitters have 50 W transmit power and pico BSs have

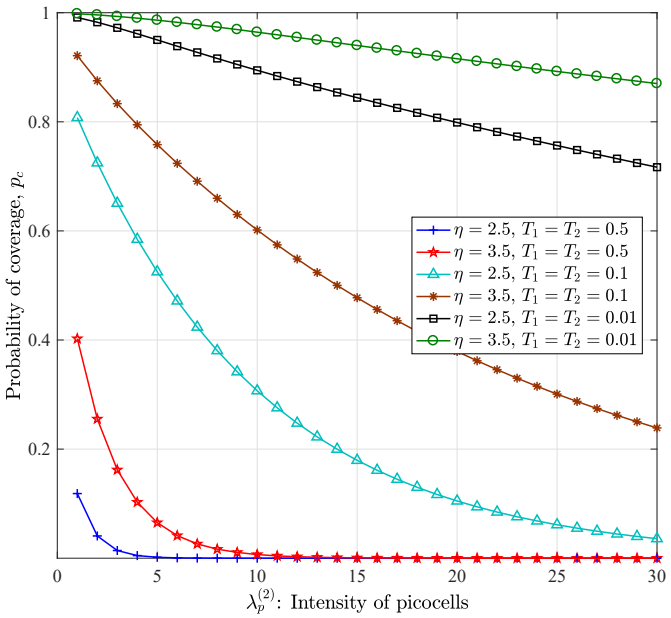


Figure 5.2. Probability of coverage in PPP-PPP HetNet against the intensity of picocells, with the desired link distance $\|z\| = 1$ (normalized) and $\lambda_p^{(1)} = 2$.

2 W. Rayleigh fading with $\mu_h = 1$ is assumed and the distance between the given user and its connected BS is $\|z\|$. For the TCP, μ is set to be $\lambda_t^{(2)}/\lambda_p^{(1)}$. Other parameter values are varied for the results of the various figures.

In Fig. 5.2, numerical results show that the probability of coverage decreases as the intensity of picocells increases, mainly because of the increased interference. However, that decrease is not as drastic as when it is caused by the increase of the coverage thresholds. These results differ from [22], which concludes the intensities of the tiers are irrelevant to the probability of coverage if each tier has the same threshold. This is partially because of our different signal attenuation compared to [22], and partially since we fix $\|z\| = 1$ in our evaluation.

Figure 5.3 shows the lower and upper bounds on coverage probability in a PPP-PCP case, for several values of thresholds T_1 and T_2. It can be seen that lower SINR coverage thresholds and shorter distances both lead to a higher probability of coverage, much the same as for the PPP-PPP case. In Fig. 5.4, we compare the two cases under

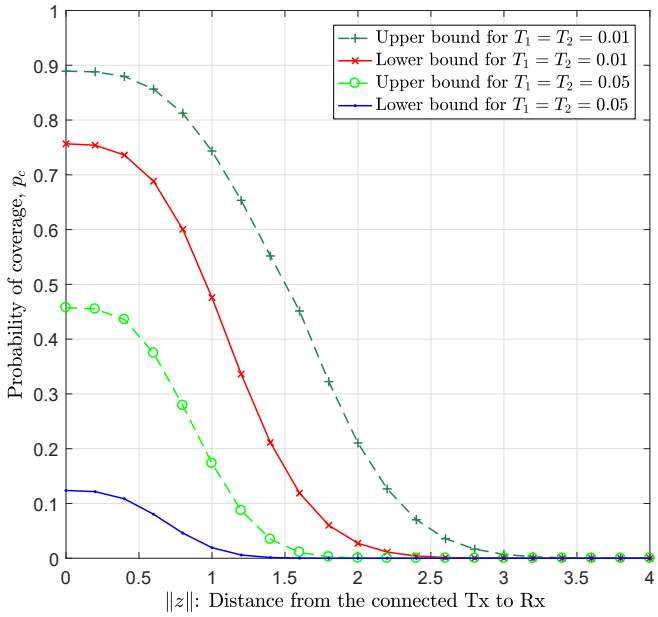


Figure 5.3. Probability of coverage in PPP-PCP HetNet against the desired link distance $\|z\|$, with $\eta = 3$, $\lambda_p^{(1)} = 2$, $\lambda_t^{(2)} = 6$, and $\sigma = 0.25$.

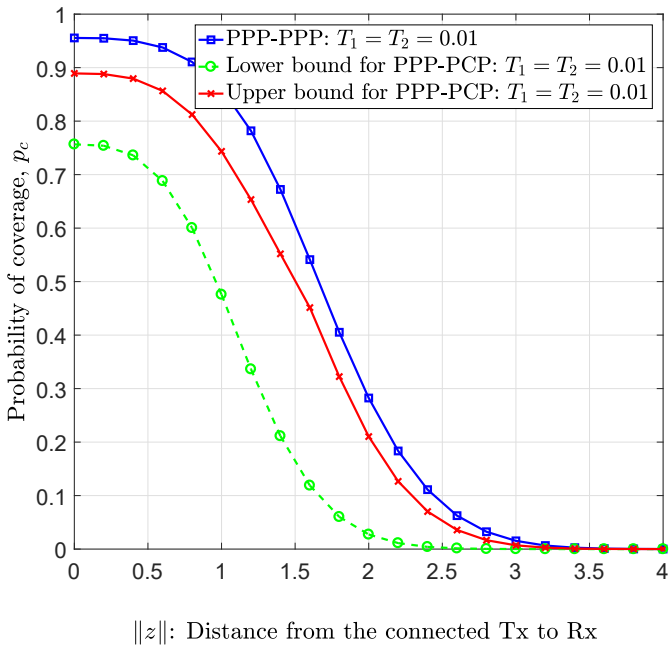


Figure 5.4. Coverage Comparison between PPP-PCP and PPP-PPP HetNet against the desired link distance $\|z\|$, with $\eta = 3$, $\lambda_p^{(1)} = 2$, $\lambda_t^{(2)} = 6$, and $\sigma = 0.25$.

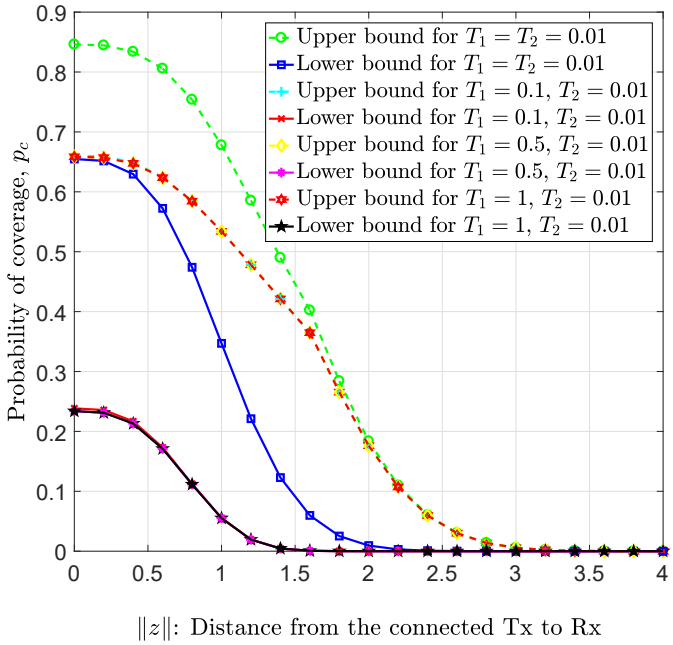


Figure 5.5. Probability of coverage in PPP-PCP HetNet with different macro tier thresholds, with $\eta = 3$, $\lambda_p^{(1)} = 2$, $\lambda_t^{(2)} = 8$, and $\sigma = 0.25$

the same parameter values. Clearly, PPP-PPP outperforms PPP-PCP, especially at smaller distances between the user and the target transmitter, since nodes clustered near the connected BS cause the worst interference and are dominant in the SINR. However, at longer distances the performance difference between the PPP-PCP case and the PPP-PPP case diminishes gradually, since the effect of clustering is weakened by the increasing distance between the given user and its connected BS transmitter.

We also show the effects of changing other parameter values on the probability of coverage for a PPP-PCP HetNet in Figs. 5.5 and 5.6. In Fig. 5.5, for $T_1 \geq T_2$, it can be seen that the lower SINR threshold dominates the probability of coverage; the plots overlap as T_1 increases. A mobile user receiver is more likely to connect to the tier with the smaller threshold, eventually effectively ignoring the coverage provided by the other tier. Figure 5.6 shows the gap between the lower and upper bounds decreases along with the degree of aggregation of the PCP points (i.e. a decrease in how much they are clustered together). This is because the bounds are based off of the coverage

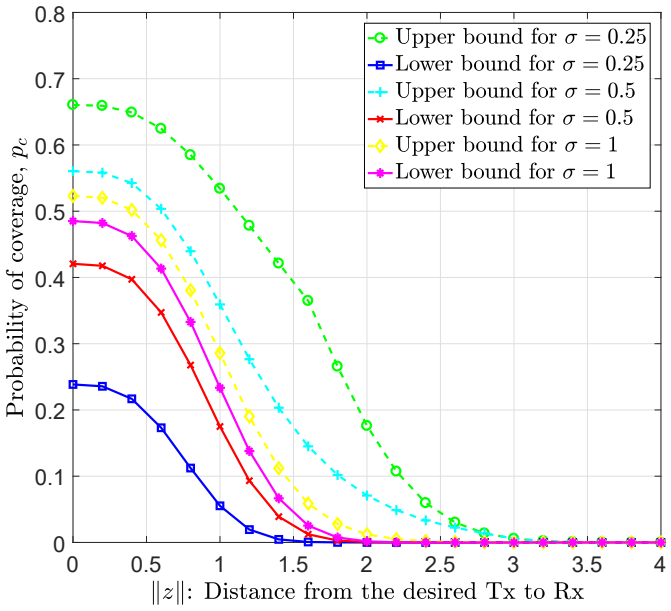


Figure 5.6. Probability of coverage in PPP-PCP HetNet with different standard deviation σ, $\eta = 3$, $\lambda_p^{(1)} = 2$, $\lambda_t^{(2)} = 8$, $T_1 = 0.1$ and $T_2 = 0.01$.

probability of a PPP, and the more spread out the PCP points are (the higher the value of σ), the more closely they resemble a PPP.

5.5 Summary

In this chapter, through the use of stochastic geometry, we have modelled N-tier clustered HetNets by different point processes, namely Poisson point processes and Poisson cluster processes. In the specific case of two tiers, we have formulated and evaluated the probability of coverage for PPP-PPP and PPP-PCP distributions of BSs. Notably, it appears the coverage probability for a PPP-PCP HetNet is always lower than that of a PPP-PPP HetNet with the same parameter values. The numerical results have shown the effects of the distance between the BS and its connected mobile user transceiver, the path loss, the coverage thresholds, the intensity of picocell clusters, and the spread of pico BSs within a cluster on the coverage probability. This examination gives useful insights for the design of future cellular HetNets.

Chapter 6

Conclusion and Future Directions

6.1 Conclusion

Stochastic geometry provides a natural way of defining and analyzing macroscopic properties of large scale wireless networks, by averaging over all possible spatial patterns for the network nodes. It abstracts the network as a realization of a random model over the whole plane, and allows the network to be analyzed in a probabilistic way. The key performance metrics of the network are translated and/or characterized as functions of a relatively small number of stochastic parameters, such as the intensities of the underlying point processes.

While stochastic geometry has its analytical strengths, the most critical step is to find an accurate point process model that best reflects the spatial distribution of the network nodes. "Accurate" in this context also means providing network performance measures (e.g. coverage probability, SINR distribution, rates, etc.) that most closely match the measures given by the real-life network. In the literature of stochastic geometry, measurement-based validation of certain stochastic assumptions used in the literature is often not provided. For example, when are Poisson models justified? When should one rather use point processes with some repulsion or attraction? When is the stationarity/isotropy assumption valid? The only aim in nearly all the research work is

to show what can be done with stochastic geometry when assumptions of certain point process models are made.

6.1.1 Summary of contributions

In Chapter 3, we present and describe a practical technique for fitting stationary and nonstationary point process models to realistic cellular networks using maximum likelihood/pseudolikelihood and minimum contrast methods. The use of covariates (accounting for population densities in urban areas and distances from the base stations to their closest main roads in rural areas) in general opens a door to new possibilities for stochastic geometry analysis. It allows the analysis to better reflect the realities, and even more ambitiously, helps bridge the gap between the stochastic model and realistic simulations.

Although PPP models result in tractable analytical expressions most of the time, extensions of the same methodology to more general point process models do not yield the same tractability, mainly due to the lack of closed-form empty space function $F(r)$ and the probability generating functional (PGFL). Therefore, in Chapter 4, we have studied the distributional properties of the closest-point distance in the Matérn hard core point process of Type II, and found that the Weibull distribution has the best fit to the empirical curve among the other existing distributions such as gamma, log-normal and Rayleigh. We have also proposed a piecewise probability density function for the closest-point distance, including a more general expression for the distance distribution from which all the parameters of the piecewise model PDF can be expressed in terms of any arbitrary values of only the parameters θ and λ_{MH} (or λ_p).

Finally, in Chapter 5, we have examined the properties of the PGFL for Poisson cluster processes. We have formulated and evaluated the downlink coverage probability for a two-tier heterogeneous cellular network consisting of PPP-PPP and PPP-PCP distributions of BSs, respectively. Through simulations, we have shown the effects of the link distance, the path loss, the coverage thresholds, the density of picocell clusters,

and the spread of pico BSs within a cluster on the coverage probability.

6.2 Future Directions

Emerging from the already completed research work, there are two possible main future research directions for stochastic geometry modelling and analysis. The first direction is to exploit more accurate covariates in the context of nonstationary point process models. Another possible future direction is to apply the piecewise PDF model in the analysis of more general performance metrics such as the average probability of coverage in MHCPP-based wireless networks. A detailed discussion of these potential research directions is provided below.

6.2.1 Covariates

The current covariates are not completely accurate or fully indicative of BS placement. In order to bridge the gaps between the results and measures of stationary models and simulations of real-life cellular networks, other covariate information, such as the propagation characteristics or traffic models derived from technical reports, could also be considered. For the population covariate, a finer grid could be used. Instead of regular squares, we could consider covariates in Poisson-Voronoi tessellations, or use the topography already inherent in the shapes of the census tracts. We are currently using deterministic covariates, but it is also possible to model the covariates themselves as random point or line processes. That might aid in the derivation of equations for the performance measures. Coverage itself could also be a covariate in some multi-tier or iterative model, much like how the coverage of a macro BS network can be considered as a covariate for the distribution of small cell BSs. Coverage could not be used as a covariate as part of a first tier/iteration, since one needs an actual realization of points to calculate coverage in the first place, and the covariate partially determines the placement of that realization. However, later iterations or tiers of the model could

potentially refine the covariate, coverage, and/or the accuracy of the model. There would almost certainly be sacrifices in the analytical tractability and ease of calculation of numerical expressions, though. Moreover, several of these covariates could be used simultaneously and might help further in modelling of networks with nonstationary models.

6.2.2 Future work with the piecewise model

The findings in Chapter 4 are important for the understanding of repulsive BS deployments, and also crucial when one wants to numerically calculate average performance metrics in wireless networks modelled by an MHCPP. Future research work can be foreseen such that the piecewise PDF model can be applied for the analysis of more general performance metrics such as the average probability of coverage in MHCPP-modelled wireless networks. For example, the distribution of the received signal power can be determined from the distribution of the closest-point distance R. A derivation of the SINR distribution would also be desired. However, this would require a tractable form of the PGFL for an MHCPP, which is not yet available. Ideally, if some closed-form expression or a tight numerical/empirical approximation of the PGFL for an MHCPP can be found, then a complete and accurate analysis for the probability of coverage can be conducted, similarly to the analysis in [21].

Appendix A

A.1 Proof of Proposition 4.1

We begin with

$$f_R(r) = \begin{cases} 2\pi\lambda_{\text{MH}} r, & 0 \leq r \leq \delta/2 \\ \alpha r^{\beta-1} \exp\left(-\varphi r^{\beta}\right), & r > \delta/2 \end{cases} \tag{A.1}$$

Being a PDF, the area under $f_R(r)$ should equal to 1, i.e. :

$$\begin{aligned} \int_0^{\infty} f_R(r)\mathrm{d}r = & \int_0^{\delta/2} 2\pi\lambda_{\text{MH}} r \mathrm{d}r + \int_{\delta/2}^{\infty} \alpha r^{\beta-1} \exp\left(-\varphi r^{\beta}\right)\mathrm{d}r \\ = & \frac{\pi\lambda_{\text{MH}}\delta^2}{4} + \frac{\alpha}{\varphi\beta} \exp\left\{-\varphi\left(\frac{\delta}{2}\right)^{\beta}\right\} = 1 \end{aligned} \tag{A.2}$$

After some routine manipulations, we obtain

$$\alpha = \frac{\beta\varphi}{4}(4 - \pi\lambda_{\text{MH}}\delta^2) \exp\left\{\varphi\left(\frac{\delta}{2}\right)^{\beta}\right\} \tag{A.3}$$

We also have the constraint on $f_R(r)$ that the function should be continuous at $r = \delta/2$, which gives

$$f_R\left(\frac{\delta}{2}\right) = 2\pi\lambda_{\text{MH}}\frac{\delta}{2} = \alpha\left(\frac{\delta}{2}\right)^{\beta-1} \exp\left\{-\varphi\left(\frac{\delta}{2}\right)^{\beta}\right\} \tag{A.4}$$

Substituting (A.3) into (A.4), φ is therefore given in terms of β as

$$\varphi = \frac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4-\pi\lambda_{\text{MH}}\delta^2)}\left(\frac{2}{\delta}\right)^{\beta} \tag{A.5}$$

Substituting (A.5) back into (A.3), α can be expressed in terms of β as

$$\alpha = 2\pi\lambda_{\text{MH}}\left(\frac{\delta}{2}\right)^{2-\beta}\exp\left\{\frac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4-\pi\lambda_{\text{MH}}\delta^2)}\right\} \tag{A.6}$$

Finally, substituting the expressions for α and φ into $f_R(r)$, we obtain

$$f_R(r) = \begin{cases} 2\pi\lambda_{\text{MH}}r, & 0 \le r \le \delta/2 \\ 2\pi\lambda_{\text{MH}}r\left(\dfrac{2r}{\delta}\right)^{\beta-2} \times & r > \delta/2 \\ \quad \exp\left\{\dfrac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4-\pi\lambda_{\text{MH}}\delta^2)}\left(1-\left(\dfrac{2r}{\delta}\right)^{\beta}\right)\right\}, & \end{cases} \tag{A.7}$$

By direct comparison of (A.7) to (4.16), we obtain the result in (4.18) for $\theta(r)$.

A.2 Proof of Remark 4.1

The second piece of $f_R(r)$ in (4.17), as $\beta \to 2$, becomes

$$\begin{aligned} f_R(r) =& 2\pi\lambda_{\text{MH}}r\exp\left\{\frac{2\pi\lambda_{\text{MH}}\delta^2}{\beta(4-\pi\lambda_{\text{MH}}\delta^2)}\left(1-\left(\frac{2r}{\delta}\right)^{\beta}\right)\right\}\left(\frac{2r}{\delta}\right)^{\beta-2} \\ =& 2\pi\lambda_{\text{MH}}r\exp\left(\frac{\pi\lambda_{\text{MH}}\delta^2}{4-\pi\lambda_{\text{MH}}\delta^2}\right)\exp\left(\frac{-4\pi\lambda_{\text{MH}}r^2}{4-\pi\lambda_{\text{MH}}\delta^2}\right) \end{aligned} \tag{A.8}$$

As λ_p and δ tend to zero, and more specifically when the product $\pi\lambda_p\delta^2 \ll 1$, then λ_{MH} becomes

$$\lambda_{\text{MH}} = \frac{1-\exp\{-\lambda_p\pi\delta^2\}}{\pi\delta^2} \approx \frac{1-(1-\lambda_p\pi\delta^2)}{\pi\delta^2} = \frac{\lambda_p\pi\delta^2}{\pi\delta^2} = \lambda_p \tag{A.9}$$

Thus, $\lambda_{\text{MH}} \approx \lambda_p$ and $\pi\lambda_{\text{MH}}\delta^2 \approx \pi\lambda_p\delta^2 \ll 1 < 4$. Hence:

$$\begin{aligned} f_R(r) =& 2\pi\lambda_{\text{MH}} r \exp\left(\frac{\pi\lambda_{\text{MH}}\delta^2}{4-\pi\lambda_{\text{MH}}\delta^2}\right)\exp\left(\frac{-4\pi\lambda_{\text{MH}}r^2}{4-\pi\lambda_{\text{MH}}\delta^2}\right) \\ \approx& 2\pi\lambda_p r \exp\left(\frac{\pi\lambda_p\delta^2}{4}\right)\exp\left(\frac{-4\pi\lambda_p r^2}{4}\right) \\ \approx& 2\pi\lambda_p r\left(1+\frac{\pi\lambda_p\delta^2}{4}\right)\exp\left(-\pi\lambda_p r^2\right) \\ \approx& 2\pi\lambda_p r \exp\left(-\pi\lambda_p r^2\right) \end{aligned} \tag{A.10}$$

Note that this is the same PDF as for a PPP as seen in (4.3). Moreover, if $r \leq \delta/2$ as in the first piece of $f_R(r)$ in (4.17), then $\pi\lambda_p r^2 \leq \pi\lambda_p\delta^2/4 \ll 1$, and so

$$2\pi\lambda_p r \exp\left\{-\pi\lambda_p r^2\right\} \approx 2\pi\lambda_p r(1-\pi\lambda_p r^2) \approx 2\pi\lambda_p r \tag{A.11}$$

Thus, as λ_p and δ tend to zero and $\beta \to 2$, the MHCPP distribution in (4.17) reduces to that of a PPP where $\lambda_{\text{MH}} \approx \lambda_p$, for all values of r.

Bibliography

[1] G. L. Stüber, *Principles of Mobile Communication*, 4th ed. Springer, Cham, 2017.

[2] D. López-Pórez, İ. Güvenç, G. de la Roche, M. Kountouris, T. Q. S. Quek, and J. Zhang, "Enhanced intercell interference coordination challenges in heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 22–30, Jun. 2011.

[3] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, "Interference management in OFDMA femtocell networks: Issues and approaches," *IEEE Wireless Commun.*, vol. 19, no. 3, pp. 86–95, Jun. 2012.

[4] J. G. Andrews, H. Claussen, M. Dohler, S. Rangan, and M. C. Reed, "Femtocells: Past, present, and future," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 497–508, Apr. 2012.

[5] S.-M. Cheng, S.-Y. Lien, F.-S. Hu, and K.-C. Chen, "On exploiting cognitive radio to mitigate interference in macro/femto heterogeneous networks," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 40–47, Jun. 2011.

[6] P. Lin, J. Zhang, Y. Chen, and Q. Zhang, "Macro-femto heterogeneous network deployment and management: From business models to technical solutions," *IEEE Wireless Commun.*, vol. 18, no. 3, pp. 64–70, Jun. 2011.

[7] N. Abramson, "The Alohanet - surfing for wireless data," *IEEE Commun. Mag.*, vol. 47, no. 12, pp. 21–25, Dec. 2009.

[8] T.-C. Hou and V. Li, "Transmission range control in multihop packet radio networks," *IEEE Trans. Commun.*, vol. 34, no. 1, pp. 38–44, Jan. 1986.

[9] S. Weber, J. G. Andrews, and N. Jindal, "An overview of the transmission capacity of wireless networks," *IEEE Trans. Commun.*, vol. 58, no. 12, pp. 3593–3604, Dec. 2010.

[10] H. Q. Nguyen, F. Baccelli, and D. Kofman, "A stochastic geometry analysis of dense IEEE 802.11 networks," in *Proc. 26th Int. Conf. Comput. Commun. (INFOCOM 2007)*, Anchorage, AK, USA, May 2007, pp. 1199–1207.

[11] G. Alfano, M. Garetto, and E. Leonardi, "New insights into the stochastic geometry analysis of dense CSMA networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 2, pp. 324–336, Feb. 2014.

[12] H. ElSawy, E. Hossain, and S. Camorlinga, "Characterizing random CSMA wireless networks: A stochastic geometry approach," in *Proc. IEEE Int. Conf. Communications (ICC 2012)*, Ottawa, Canada, Jun. 2012, pp. 5000–5004.

[13] H. ElSawy and E. Hossain, "A modified hard core point process for analysis of random CSMA wireless networks in general fading environments," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1520–1534, Apr. 2013.

[14] Y. Kim, F. Baccelli, and G. de Veciana, "Spatial reuse and fairness of ad hoc networks with channel-aware CSMA protocols," *IEEE Trans. Inf. Theory*, vol. 60, no. 7, pp. 4139–4157, Jul. 2014.

[15] A. Hasan and J. G. Andrews, "The guard zone in wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 3, pp. 897–906, Mar. 2007.

[16] M. Kaynia, N. Jindal, and G. E. Øien, "Improving the performance of wireless ad hoc networks through MAC layer design," *IEEE Trans. Wireless Commun.*, vol. 10, no. 1, pp. 240–252, Jan. 2011.

[17] H. Sindhwal, M. Dasari, and N. Vattikuti, "Slot conflict resolution in TDMA based mobile ad hoc networks," in *2015 Annu. IEEE India Conf. (INDICON)*, New Delhi, India, Dec. 2015, pp. 1–6.

[18] Y. Khan, M. Derakhshani, S. Parsaeefard, and T. Le-Ngoc, "Self-organizing TDMA MAC protocol for effective capacity improvement in IEEE 802.11 WLANs," in *2015 IEEE Globecom Workshops*, San Diego, CA, USA, Dec. 2015, pp. 1–6.

[19] T. Kawakami and K. Kamakura, "Modified TDMA-based MAC protocol for vehicular ad hoc networks," in *2015 IEEE Int. Conf. on Pervasive Comput. Commun. Workshops (PerCom Workshops)*, St. Louis, MO, USA, Mar. 2015, pp. 93–98.

[20] H. Inaltekin, M. Chiang, H. V. Poor, and S. B. Wicker, "On unbounded path-loss models: Effects of singularity on wireless network performance," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1078–1092, Sep. 2009.

[21] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Trans. Commun.*, vol. 59, no. 11, pp. 3122–3134, Nov. 2011.

[22] H. S. Dhillon, R. K. Ganti, F. Baccelli, and J. G. Andrews, "Modeling and analysis of K-tier downlink heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 550–560, Apr. 2012.

[23] A. Goldsmith, *Wireless Communications*. New York, NY: Cambridge University Press, 2005.

[24] R. K. Ganti and M. Haenggi, "Interference and outage in clustered wireless ad hoc networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4067–4086, Sep. 2009.

[25] F. Baccelli and B. Błaszczyszyn, "Stochastic geometry and wireless networks: Volume I - Theory," *Foundations and Trends in Networking*, vol. 3, no. 3-4, pp. 249–449, 2009.

[26] F. Baccelli and B. Błaszczyszyn, "Stochastic geometry and wireless networks: Volume II - Applications," *Foundations and Trends in Networking*, vol. 4, no. 1-2, pp. 1–312, 2009.

[27] K. S. Gilhousen, I. M. Jacobs, R. Padovani, A. J. Viterbi, L. A. Weaver Jr., and C. E. Wheatley III, "On the capacity of a cellular CDMA system," *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 303–312, May 1991.

[28] 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 Trans. Parallel Distrib. Syst.*, vol. 13, no. 9, pp. 963–971, Sep. 2002.

[29] V. P. Mhatre and C. P. Rosenberg, "Impact of network load on forward link inter-cell interference in cellular data networks," *IEEE Trans. Wireless Commun.*, vol. 5, no. 12, pp. 3651–3661, Dec. 2006.

[30] P. Charoen and T. Ohtsuki, "Codebook based interference mitigation with base station cooperation in multi-cell cellular network," in *Proc. 74th IEEE Veh. Technol. Conf. (VTC 2011-Fall)*, San Francisco, CA, USA, Sep. 2001, pp. 1–5.

[31] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications*, 3rd ed. Chichester, UK: Wiley, 2013.

[32] B. Matérn, *Spatial Variation*, 2nd ed., ser. Lecture Notes in Statistics, D. Brillinger *et al.*, Eds. Berlin/Heidelberg, Germany: Springer, 1986, vol. 36.

[33] D. Stoyan and H. Stoyan, *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. Chichester, UK: Wiley, 1994.

[34] H. ElSawy, E. Hossain, and M. Haenggi, "Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 996–1019, Third Quarter 2013.

[35] M. Haenggi, "Mean interference in hard-core wireless networks," *IEEE Commun. Lett.*, vol. 15, no. 8, pp. 792–794, Aug. 2011.

[36] N. Deng, W. Zhou, and M. Haenggi, "The Ginibre point process as a model for wireless networks with repulsion," *IEEE Trans. Wireless Commun.*, vol. 14, no. 1, pp. 107–121, Jan. 2015.

[37] Y. Li, F. Baccelli, H. S. Dhillon, and J. G. Andrews, "Statistical modeling and probabilistic analysis of cellular networks with determinantal point processes," *IEEE Trans. Commun.*, vol. 63, no. 9, pp. 3405–3422, Sep. 2015.

[38] Y. Zhou, R. Li, Z. Zhao, X. Zhou, and H. Zhang, "On the α-stable distribution of base stations in cellular networks," *IEEE Commun. Lett.*, vol. 19, no. 10, pp. 1750–1753, Oct. 2015.

[39] A. Guo and M. Haenggi, "Spatial stochastic models and metrics for the structure of base stations in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 11, pp. 5800–5812, Nov. 2013.

[40] J. Kibiłda, B. Galkin, and L. A. DaSilva, "Modelling multi-operator base station deployment patterns in cellular networks," *IEEE Trans. Mobile Comput.*, vol. 15, no. 12, pp. 3087–3099, Dec. 2016.

[41] D. J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Volume II*, 2nd ed. New York, NY: Springer, 2007.

[42] B. D. Ripley, "The second-order analysis of stationary point processes," *J. Appl. Probability,*, vol. 13, no. 2, pp. 255–266, Jun. 1976.

[43] H. S. Dhillon, R. K. Ganti, and J. G. Andrews, "Load-aware modeling and analysis of heterogeneous cellular networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 4, pp. 1666–1677, Apr. 2013.

[44] H.-S. Jo, Y. J. Sang, P. Xia, and J. G. Andrews, "Heterogeneous cellular networks with flexible cell association: A comprehensive downlink SINR analysis," *IEEE Trans. Wireless Commun.*, vol. 11, no. 10, pp. 3484–3495, Oct. 2012.

[45] T. D. Novlan, R. K. Ganti, A. Ghosh, and J. G. Andrews, "Analytical evaluation of fractional frequency reuse for OFDMA cellular networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 12, pp. 4294–4305, Dec. 2011.

[46] ——, "Analytical evaluation of fractional frequency reuse for heterogeneous cellular networks," *IEEE Trans. Commun.*, vol. 60, no. 7, pp. 2029–2039, Jul. 2012.

[47] S. Mukherjee, "Distribution of downlink SINR in heterogeneous cellular networks," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 3, pp. 575–585, Apr. 2012.

[48] V. Chandrasekhar and J. G. Andrews, "Spectrum allocation in tiered cellular networks," *IEEE Trans. Commun.*, vol. 57, no. 10, pp. 3059–3068, Oct. 2009.

[49] X. Zhang and M. Haenggi, "Random power control in Poisson networks," *IEEE Trans. Commun.*, vol. 60, no. 9, pp. 2602–2611, Sep. 2012.

[50] M. Haenggi, "On distances in uniformly random networks," *IEEE Trans. Inf. Theory*, vol. 51, no. 10, pp. 3584–3586, Oct. 2005.

[51] A. M. Hunter, J. G. Andrews, and S. Weber, "Transmission capacity of ad hoc networks with spatial diversity," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5058–5071, Dec. 2008.

[52] N. Jindal, S. Weber, and J. G. Andrews, "Fractional power control for decentralized wireless networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5482–5492, Dec. 2008.

[53] C. H. M. de Lima, M. Bennis, and M. Latva-aho, "Coordination mechanisms for self-organizing femtocells in two-tier coexistence scenarios," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2212–2223, Jun. 2012.

[54] M. G. Khoshkholgh, K. Navaie, and H. Yanikomeroglu, "Outage performance of the primary service in spectrum sharing networks," *IEEE Trans. Mobile Comput.*, vol. 12, no. 10, pp. 1955–1971, Oct. 2013.

[55] A. Ghasemi and E. S. Sousa, "Interference aggregation in spectrum-sensing cognitive wireless networks," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 41–56, Feb. 2008.

[56] A. Rabbachin, T. Q. S. Quek, H. Shin, and M. Z. Win, "Cognitive network interference," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 2, pp. 480–493, Feb. 2011.

[57] S. Srinivasa and M. Haenggi, "Modeling interference in finite uniformly random networks," in *Int. Workshop on Inf. Theory for Sensor Netw. (WITS 2007)*, Santa Fe, NM, USA, Jun. 2007, pp. 1–12.

[58] ——, "Distance distributions in finite uniformly random networks: Theory and applications," *IEEE Trans. Veh. Technol.*, vol. 59, no. 2, pp. 940–949, Feb. 2010.

[59] M. Haenggi and R. Ganti, "Interference in large wireless networks," *Foundations and Trends in Networking*, vol. 3, no. 2, pp. 127–248, NOW Publishers, 2008.

[60] R. Mathar and J. Mattfeldt, "On the distribution of cumulated interference power in Rayleigh fading channels," *Wireless Netw.*, vol. 1, no. 1, pp. 31–36, Mar. 1995.

[61] P. Cardieri, "Modeling interference in wireless ad hoc networks," *IEEE Commun. Surveys Tuts.*, vol. 12, no. 4, pp. 551–572, Fourth Quarter 2010.

[62] J. Venkataraman, M. Haenggi, and O. Collins, "Shot noise models for outage and throughput analyses in wireless ad hoc networks," in *Proc. IEEE Military Commun. Conf. (MILCOM06)*, Washington, DC, USA, Oct. 2006, pp. 1–7.

[63] E. S. Sousa and J. A. Silvester, "Optimum transmission range in a direct-sequence spread-spectrum multihop packet radio network," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 5, pp. 762–771, 1990.

[64] N. A. C. Cressie, *Statistics for Spatial Data*. (Revised Edition) New York, NY: John Wiley & Sons, 1993.

[65] B. D. Ripley, *Spatial Statistics*. New York, NY: John Wiley & Sons, 1981.

[66] A. J. Baddeley, J. Møller, and R. Waagepetersen, "Non- and semi-parametric estimation of interaction in inhomogeneous point patterns," *Statistica Neerlandica*, vol. 54, no. 3, pp. 329–350, Nov. 2000.

[67] A. Baddeley and R. Turner, "Spatstat: an R package for analyzing spatial point patterns," *J. Stat. Softw.*, vol. 12, no. 6, pp. 1–42, Jan. 2005.

[68] A. Baddeley, P. Gregori, J. Mateu, R. Stoica, and D. Stoyan, Eds., *Case Studies in Spatial Point Process Modeling*. Lecture Notes in Statistics. New York, NY: Springer, 2006, vol. 185.

[69] H. Akaike, "A new look at the statistical model identification," *IEEE Trans. Autom. Control*, vol. 19, no. 6, pp. 716–723, Dec. 1974.

[70] K. P. Burnham and D. R. Anderson, *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd ed. New York, NY: Springer, 2003.

[71] P. J. Diggle, *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns*, 3rd ed. Boca Raton, FL: CRC Press, 2013.

[72] M. N. M. van Lieshout, "A J-function for inhomogeneous point processes," *Statistica Neerlandica*, vol. 65, no. 2, pp. 183–201, May 2011.

[73] MountainMath Software. CensusMapper. [Online]. Available: https://censusmapper.ca/

[74] Statistics Canada. Census profile, 2016 census. [Online]. Available: http://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E

[75] M. Haenggi, *Stochastic Geometry for Wireless Networks*. Cambridge, UK: Cambridge Univ. Press, 2012.

[76] T. V. Nguyen and F. Baccelli, "A probabilistic model of carrier sensing based cognitive radio," in *Proc. 2010 IEEE Symp. New Frontiers in Dynamic Spectrum Access Netw. (DySPAN 2010)*, Singapore, Apr. 2010, pp. 1–12.

[77] A. Al-Hourani, R. J. Evans, and S. Kandeepan, "Nearest neighbor distance distribution in hard-core point processes," *IEEE Commun. Lett.*, vol. 20, no. 9, pp. 1872–1875, Sep. 2016.

[78] R. W. Heath, Jr., M. Kountouris, and T. Bai, "Modeling heterogeneous network interference using Poisson point processes," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4114–4126, Aug. 15, 2013.

[79] 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 J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.

[80] X. Yang and A. P. Petropulu, "Co-channel interference modeling and analysis in a Poisson field of interferers in wireless communications," *IEEE Trans. Signal Process.*, vol. 51, no. 1, pp. 64–76, Jan. 2003.

[81] K. Gulati, B. L. Evans, J. G. Andrews, and K. R. Tinsley, "Statistics of co-channel interference in a field of Poisson and Poisson-Poisson clustered interferers," *IEEE Trans. Signal Process.*, vol. 58, no. 12, pp. 6207–6222, Dec. 2010.

[82] R. K. Ganti and M. Haenggi, "Interference in ad hoc networks with general motion-invariant node distributions," in *Proc. IEEE Int. Symp. Inf. Theory (ISIT) 2008*, Toronto, ON, Canada, Jul. 2008, pp. 1–5.

[83] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, 6th ed., A. Jeffrey and D. Zwillinger, Eds. San Diego, CA: Academic Press, 2000.