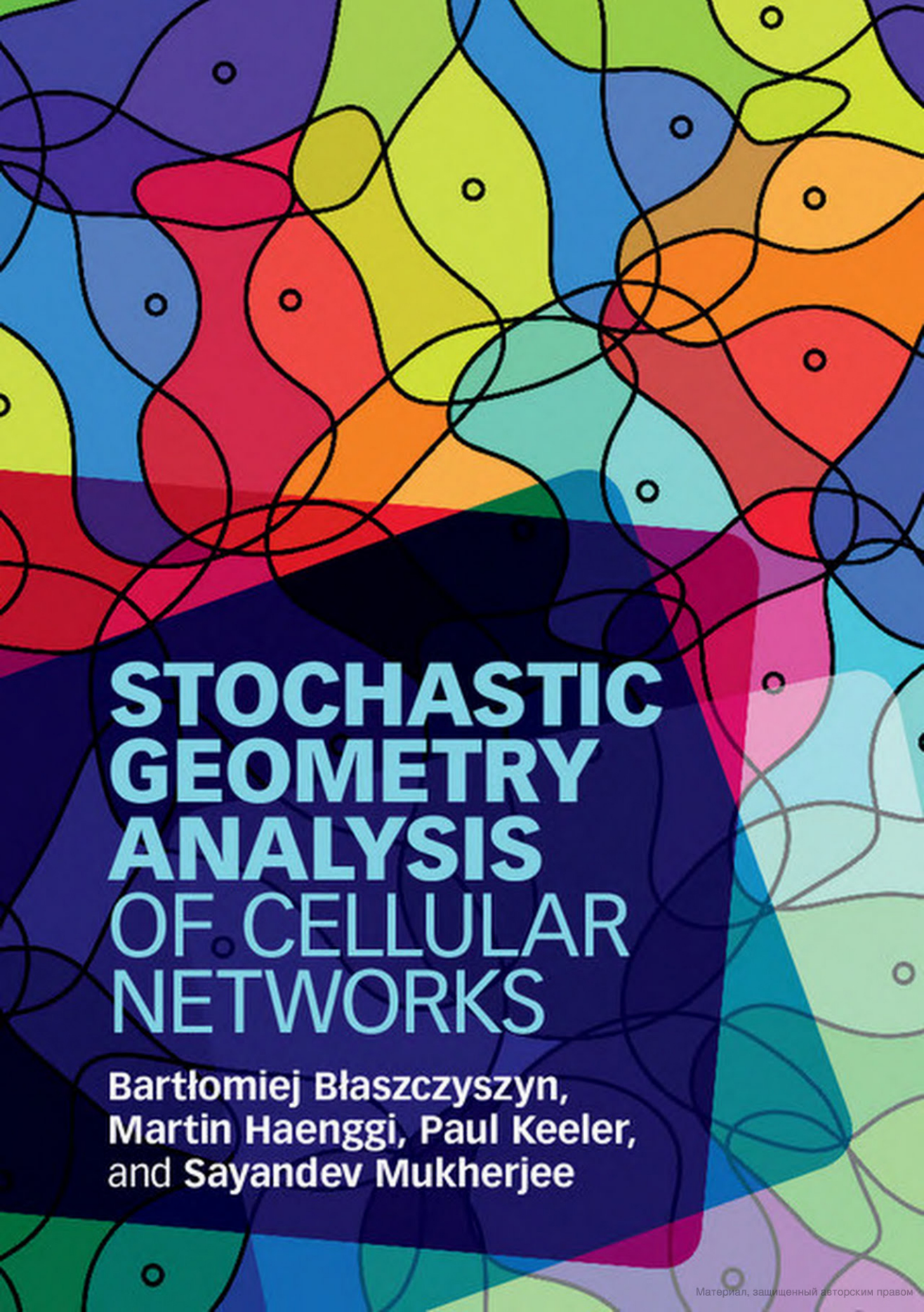

STOCHASTIC GEOMETRY ANALYSIS OF CELLULAR NETWORKS

Bartłomiej Błaszczyszyn,
Martin Haenggi, Paul Keeler,
and **Sayandev Mukherjee**

Stochastic Geometry Analysis of Cellular Networks

Achieve faster and more efficient network design and optimization with this comprehensive guide. Some of the most prominent researchers in the field explain the very latest analytic techniques and results from stochastic geometry for modeling the signal-to-interference-plus-noise ratio (SINR) distribution in heterogeneous cellular networks. This book will help you understand the effects of combining different system deployment parameters on such key performance indicators as coverage and capacity, enabling efficient allocation of simulation resources. In addition to covering results for network models based on the Poisson point process, this book presents recent results for when non-Poisson base station configurations appear Poisson due to random propagation effects such as fading and shadowing, as well as non-Poisson models for base station configurations, with a focus on determinantal point processes and tractable approximation methods. Theoretical results are illustrated with practical long-term evolution (LTE) applications and compared with real-world deployment results.

BARTŁOMIEJ BŁASZCZYSZYN is a Research Director at Inria in Paris, France, and a faculty member of the joint Inria–ENS research group, DYOGENE.

MARTIN HAENGGI is the Frank M. Freimann Professor of Electrical Engineering and a concurrent Professor of Applied and Computational Mathematics at the University of Notre Dame in Indiana, USA, and the author of *Stochastic Geometry for Wireless Networks* (Cambridge University Press, 2012).

PAUL KEELER is a researcher at the Weierstrass Institute for Applied Analysis and Stochastics in Berlin, Germany, and an Associate Fellow of the School of Mathematics and Statistics at the University of Melbourne in Victoria, Australia.

SAYANDEV MUKHERJEE is a senior research engineer at DOCOMO Innovations, Inc. in California, USA, and the author of *Analytical Modeling of Heterogeneous Cellular Networks* (Cambridge University Press, 2014).

"These four renowned experts deliver a comprehensive yet curated treatment on the modeling and analysis of modern cellular networks using stochastic geometry, which has been one of the most important recent lines of wireless research. Highly recommended for interested researchers and engineers. Can serve as a useful companion to Haenggi's landmark stochastic geometry textbook, which had fairly minimal treatment of cellular networks."

Jeff Andrews,
University of Texas at Austin

Stochastic Geometry Analysis of Cellular Networks

BARTŁOMIEJ BŁASZCZYSZYN
Inria and ENS

MARTIN HAENGGI
University of Notre Dame

PAUL KEELER
Weierstrass Institute for Applied Analysis and Stochastics

SAYANDEV MUKHERJEE
DOCOMO Innovations, Inc.

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107162587
DOI: 10.1017/9781316677339

First published 2018

Printed in the United Kingdom by Clays, St Ives plc

A catalogue record for this publication is available from the British Library.

Library of Congress Cataloging-in-Publication Data
Names: Błaszczyszyn, Bartłomiej, 1967– author. | Haenggi, Martin, 1969– author. | Keeler, Paul, 1981– author. | Mukherjee, Sayandev, 1970– author.
Title: Stochastic geometry analysis of cellular networks / Bartłomiej Błaszczyszyn, Inria and ENS, Martin Haenggi, University of Notre Dame, Paul Keeler, Weierstrass Institute for Applied Analysis and Stochastics, Sayandev Mukherjee, DOCOMO Innovations, Inc.
Description: Cambridge, United Kingdom ; New York, NY, USA : Cambridge University Press, 2018. | Includes bibliographical references and index.
Identifiers: LCCN 2017033066 | ISBN 9781107162587 (hardback : alk. paper)
Subjects: LCSH: Wireless communication systems–Mathematics. | Stochastic models. | Stochastic geometry.
Classification: LCC TK5102.83 .B54 2018 | DDC 621.3845/60151922–dc23
LC record available at https://lccn.loc.gov/2017033066

ISBN 978-1-107-16258-7 Hardback

Contents

Preface

Since 2010, our knowledge of coverage and capacity in heterogeneous cellular networks (HetNets) has expanded rapidly, primarily through analytical results using stochastic geometry. Most of these results assume that the locations of the base stations in a given tier of the HetNet are the points of a homogeneous Poisson point process (PPP). This modeling assumption was made for mathematical tractability, and the coverage results for a single tier were determined to be about as pessimistic (relative to the real coverage in real-world deployments) as the regular hexagonal lattice location model was optimistic (Andrews, Baccelli, & Ganti 2011).

However, since 2013 we have obtained results showing the fundamental importance of the PPP model to the analysis of real-world deployments in two ways: (a) the set of propagation losses to the typical location in an arbitrary network deployment converge asymptotically to that from a PPP deployment of base stations; (b) long before this convergence is reached, the coverage results for a PPP deployment can be employed to obtain very accurate approximations to the coverage results for various regular deployments.

This book provides the first detailed expository treatment of these results, and includes additional exact analytical results on coverage for certain special non-PPP deployment models. We expect the book to be of interest to researchers in academia and industry, and anyone interested in the application of stochastic geometry to problems in communication. HetNets are an important component of future cellular network standards (LTE Release 12 and later),[1] and the theoretical results in the book are illustrated with examples of their application to transmission scenarios specified in the LTE standard.

Although we have made every effort to make the book self-contained, the reader will benefit from having had some prior exposure both to the theory and techniques of stochastic geometry, and to their application to the derivation of coverage results for PPP base station deployments. For an introduction to stochastic geometry with emphasis on wireless communications, we recommend Haenggi's *Stochastic Geometry for Wireless Networks* (Haenggi 2012) at the introductory level (a condensed version of which is available in Chapter 3) and the two volumes of Baccelli and Błaszczyszyn's *Stochastic Geometry and Wireless Networks* (Baccelli & Błaszczyszyn 2009a, 2009b) for a comprehensive treatment at the advanced level. Note that both references rely on

[1] See Appendix B.

mathematical results from a measure-theoretic treatment of stochastic geometry, such as that found in Chiu et al. 2013. For a simplified treatment of some of the results on coverage and capacity in PPP network deployments in Chapter 5, the interested reader is referred to Mukherjee (2014), which does not require knowledge of measure theory.

Knowledge of 3GPP-LTE is not required, but familiarity with the LTE standard will help the reader in understanding the applications of the results presented in the book. An outline of the different releases in the LTE standard and the features relevant to HetNets in each one is provided in Appendix B.

Acknowledgments

It is a remarkable testament to modern communications that four authors on three continents were able to collaborate so amiably and productively on this book, having met collectively but once, and we are grateful to Professors Jeffrey G. Andrews, François Baccelli, and Gustavo de Veciana of the University of Texas at Austin, and the Simons Foundation for making that singular occasion possible, at the Simons Conference on Networks and Stochastic Geometry in Austin in May 2015 that they organized.

We are grateful to Naoto Miyoshi for his careful reading of the manuscript and his precious remarks, in particular in Chapter 8. We are also grateful to Julie Lancashire and Heather Brolly, our editor and content manager, respectively, at Cambridge University Press for their patience as deadlines for the finished manuscript came and went as the authors juggled their work and teaching schedules with writing.

B. B. would like to express his gratitude to François Baccelli for introducing him to the art and science of modeling of communication networks, as well as to the Inria–ENS team where, since 1999, B. B. had the pleasure and honor of collaborating with François Baccelli on the foundations of stochastic geometric modeling of wireless networks. Without these early, visionary scientific choices and the leadership of François Baccelli, as well as the working environment provided by Inria and ENS, the author's more recent contribution to this book would not have been possible. The author is also grateful to Mohamed Kadhem Karray for a long, close collaboration and very many stimulating discussions. Finally, but equally importantly, B. B. would like to thank his lovely family – Mira, Klara, Karol, and Antoni – for their understanding and support while he worked on his share of the manuscript.

M. H. thanks Béla Bollobás for sparking his interest in random geometric graphs and stochastic geometry in the early 2000s, when the application of stochastic geometry to wireless networks modeling and analysis was still in its infancy. Since then, the annual number of papers published on the topic has grown by almost 30 dB! He is grateful to his PhD students, postdoc, and other collaborators; most of the credit for the results and insights obtained are due to their tireless efforts and stimulating discussions. And he would like to express his deep gratitude to his family – Roxana and little Kevin – for their love, support, and understanding that this book project would take time away from them.

P. K. thanks his PhD supervisor, Peter G. Taylor, at the University of Melbourne, who set him on the path of using stochastic geometry models, which ultimately led to doing interesting projects with B. B. in the Inria–ENS team, then headed by François Baccelli.

This journey continued by working at the Weierstrass Institute in a team headed by Wolfgang König and focused on applying stochastic methods to wireless networks. P. K. is grateful to the aforementioned researchers and research institutes. None of this would have started without his family who long ago inspired in him interests in mathematics, science, and engineering.

S. M. would like to express his gratitude once more to Professor Jeffrey G. Andrews of the University of Texas at Austin, for inviting him to speak at the 2015 Simons Conference, an occasion on which he met for the first time not only two of his co-authors, but also nearly all the top researchers in stochastic geometry, most of whom he knew until then only from their publications. S. M. also thanks the management at DOCOMO Innovations, Inc., for their support, and the corporate general counsel, Jeremy Tucker, Esq., for negotiating the book contract with Cambridge University Press. Last but certainly not least, he would like to thank his family – Chandreyee, Rik, Inika, and the four-legged tail-wagging distraction Johnny – for their patience while he worked on his share of the manuscript.

Notations

$\sim$	distributed as
${}_2F_1(a, b; c; z)$	Gaussian hypergeometric function
α	slope of path loss model (path loss exponent)
δ	$= 2/\alpha$
$\mathcal{B}_{[0,1]}$	the Borel σ-algebra on $[0, 1]$
τ	SINR threshold
$\Gamma(z)$	gamma function
$\gamma(z, a)$	lower incomplete gamma function
Φ, Ψ	PPP of base station locations
$\tilde{\Phi}, \tilde{\Psi}$	PPP of received powers from the base stations of the PPP Φ or Ψ
x	scalar
$\boldsymbol{x}$	vector
x	a point of a point process
$\mathbf{A}$	matrix
X	random variable
$f_X(x)$	probability density function of X evaluated at x
$\boldsymbol{X}$	random vector
$f_{\boldsymbol{X}}(\boldsymbol{x})$	joint probability density function of $\boldsymbol{X}$ evaluated at $\boldsymbol{x}$
$\Lambda(\cdot)$	intensity measure
I	interference power
K	intercept of path loss model
$\ell(\cdot)$	path loss function
L_{x}	propagation loss between the user at the origin and the base station x
$\mathcal{L}_X(s)$	Laplace transform of X evaluated at s: $\mathbb{E}[\exp(-sX)]$
$\mathcal{N}$	the space of all sequences
N_φ	counting measure of φ
ν	Lebesgue measure
φ	a point pattern
B	set or event
$\mathbf{1}_B(\cdot)$	indicator function of the set B
$\mathbb{E}[X]$	expectation of the random variable X
$\mathbb{P}(B)$	probability of the event B, also equal to $\mathbb{E}[\mathbf{1}_B]$
$\mathbb{R}$	the set of real numbers, also written $(-\infty, \infty)$
$\mathbb{R}_+$	the set of nonnegative reals, also written $[0, \infty)$
$\mathbb{R}_{++}$	the set of positive reals, also written $(0, \infty)$

Acronyms and Abbreviations

3GPP	Third Generation Partnership Project
ASAPPP	approximate SIR analysis based on the Poisson point process
AWGN	additive white Gaussian noise
BLER	block error rate
CapEx	capital expenses
CCDF	complementary cumulative distribution function
CDF	cumulative distribution function
CoMP	coordinated multipoint
EFIR	expected fading-to-interference ratio
GPP	Ginibre point process
HetNet	heterogeneous cellular network
IA	interference alignment
ICIC	intercell interference coordination
ICSC	interference cancellation and signal combination
IIC	independent interference cancellation
iid	independent and identically distributed
ISR	interference-to-signal ratio
LTE	long-term evolution
LTE-A	long-term evolution-advanced
MIMO	multiple-input multiple-output
MISR	mean interference-to-signal ratio
MRC	maximal ratio combining
pgfl	probability generating functional
PPP	Poisson point process
RDP	relative distance process
REB	range expansion bias
SC	signal combination
SER	symbol error rate
SIC	successive interference cancelation
SINR	signal-to-interference-plus-noise ratio
SIR	signal-to-interference ratio
SISO	single-input single-output
SNR	signal-to-noise ratio
STINR	signal-to-total-interference-plus-noise ratio
STIR	signal-to-total-interference ratio
UMTS	universal mobile telecommunications system
ZFBF	zero-forcing beamforming

Part I

Stochastic Geometry

1 Introduction

This chapter discusses the reasons for the importance of heterogeneous network (HetNet) architectures in future wireless network deployments. We begin with a brief survey of the state of wireless networks and the chief challenges faced by designers and operators of such networks. We then discuss possible solutions to these problems and single out small-cell networks as probably the best engineering solution. The deployment of small-cell networks then leads directly to the design and configuration of the resulting heterogeneous networks, which is the focus of this book. We make the case for new approaches to complement the traditional simulation-based schemes for design and performance evaluation when the wireless networks in question are heterogeneous, and provide an overview of the principal results that are now known (and that continue to be added to at a rapid pace), mostly as a result of research done since 2010.

1.1 The Demand for Ubiquitous Connectivity

The Information Age and the Computer Age have converged in the form of portable wireless devices that are full-fledged computers and that can, with a connection to the internet, provide instant always-on access to information anywhere. In other words, we are moving toward the "nomadic computing" vision of Kleinrock (Kleinrock 1996), where we can perform information processing "anytime, anywhere," but instead of doing so in Kleinrock's "disconnected" world, we do so by ensuring that the world is fully "connected."

The main source of such ubiquitous connectivity, especially for mobile and outdoor users, is the *cellular network* as it has evolved over three decades and four generations of air interfaces (see Appendix B for details). Indoors, the cellular network is often complemented or even supplanted by a mix of wired and wireless (WiFi) networks, but since the indoor scenario is the simplest one in which to maintain connectedness, we shall focus on outdoor and mobile scenarios in the following discussion of ubiquitous connectivity. At any rate, it is a fait accompli that the *cellular* model, where *base stations* serve as exclusive access points to the internet within defined geographical areas called *cells*, is the foundation for any network that offers such ubiquitous connectivity.[1]

[1] We note here that recent trends in the design of future wireless network architectures envision a weakening of the "exclusive" role of base stations within geographic regions, diluting the whole concept of the cell itself. We discuss these topics later in the book.

The value of such ubiquitous connectivity to business and industry is obvious. What was not obvious before, but is now clear from the popularity of social networking applications on these wireless portable devices, is that such ubiquitous connectivity seems to address a fundamental human desire to connect, and stay connected, to other humans. It is not an exaggeration to say that the availability of such "anytime, anywhere" connectivity helps humans live more fulfilled lives. It is therefore imperative to serve more users than ever before, at higher data rates than ever before.

Given the enormous economic and social forces in favor of increasing the availability and quality of such connectivity, the providers of such connectivity (hereafter called *network operators*) are in the happy situation of facing little to no social challenge to the expansion of their business,[2] and it remains only for them to address the technical challenges involved in designing, deploying, and maintaining networks that offer such ubiquitous connectivity at an affordable price. In Section 1.2, we articulate some of these challenges and discuss suggested solutions.

1.2 Technical Challenges for a Network Operator

A network operator designing or deploying a wireless communication network needs to consider, at a minimum, the following three criteria:

1. *Affordability*: When we talk of affordability from the point of view of the network operator, we mean the costs of deploying (capital expenses, or CapEx) and maintaining (operating expenses) such a network. Note that CapEx includes not only the cost of the equipment on the network, but also the cost of acquiring the spectrum license(s), which can often be substantial for the low (<2 GHz) frequency bands. These bands are in higher demand because of the reduced propagation losses compared to higher-frequency bands, but clearly if higher-frequency bands could be used efficiently by the network operator, they could save on the acquisition costs of this spectrum. Further, large cellular base station towers are expensive to deploy and operate, and efficiencies therein directly translate to the bottom line of the network operator.
2. *Availability*: The challenge of mere access to the internet has been largely addressed nearly worldwide – even the all but obsolete 2G (second generation; see Appendix B) cellular networks offer low-speed (around 300 kbps) access to the internet, and such networks are available worldwide and cover most locations with a settled population. The main technical challenge is the next criterion, namely the quality of the data connections to the users of this operator's network.
3. *Quality*: In Section 7.3 we will provide a more detailed and technical definition of the quality of a data connection, or a collection of data connections (say, one per user of the network), but for the moment, let us say that from the point of view of the network operator, the quality of the network is indicated by the average total

[2] Except from some people who do not want unsightly cellular towers in their towns; as we shall see, even their concerns can be addressed by technology.

throughput per unit area in the network (that is, from all base stations to all users per unit area).[3] This average total throughput per unit area, which we loosely refer to as the *capacity* of the network,[4] is given by the product of the average number of active (actually transmitting and/or receiving data) connections per unit area with the average rate on such a data connection, which we know is in turn the product of the *bandwidth* and the average *spectral efficiency* on this data connection, where the spectral efficiency of a link is closely related to the signal-to-interference-plus-noise ratio (SINR) on that link due to Shannon's formula. In other words, the network capacity is the product of the area-density of active connections, the average spectral efficiency on each such connection, and the bandwidth available to such a connection.

Note that the goal of serving more users at higher data rates than before translates into increasing both the average number of active connections per unit area and the average data rate per active connection. Of the three quantities that together determine the network capacity (the "quality" of the network), the bandwidth is the least easy to increase. Typically, new spectrum is freed up for auction by a national regulator at infrequent intervals, often years apart, and the winners of such spectrum auctions pay large sums for the bandwidth they get to use. Moreover, there is simply never enough spectrum released in a single tranche to increase the bandwidth by a factor of 10, say. Thus, the average data rate cannot increase by much due to increased bandwidth, but relies on improvements in spectral efficiency.

Multiple-antenna transmission technologies (collectively called *multiple-input multiple-output or MIMO*) do yield improvements in spectral efficiency, but the physical sizes of wireless devices preclude the use of more than, say, four antennas (at least for <5 GHz bands), and this in turn limits the achievable spectral efficiency gains. Further, the transmission schemes that seek to exploit these gains are complex and require sophisticated signaling between base stations and user devices.

This leaves us with the final factor determining the network capacity, namely the average activity (number of active data connections per unit area that can be supported with acceptable data rate by the network). This number could be increased by, for example, dividing the available total bandwidth into two parts (bands), and distributing the transmissions over these two bands, provided that each of the two bands is adequate to support an acceptable data rate. In light of the scarcity of available spectrum, it is clear that the above division of spectrum cannot occur too many times. An alternative to spectrum division is spectrum reuse, where we attempt to support all transmissions on the same band while mitigating their mutual interference. It has been estimated that between 1955 and 2010, there was an approximately million-fold increase in network capacity, of which only about a five-fold increase can be attributed to spectral efficiency gains, but a 25,000-fold increase is attributable to spectrum reuse (Andrews 2011).

[3] As we shall see, distribution of the total throughput per unit area is more descriptive of network performance than the average, but for now let us focus on the average.

[4] There are several definitions of *capacity* but for the purposes of this section, this is the definition we shall use.

Highly sophisticated means of interference avoidance and/or cancellation exist, but for a robust and scalable network architecture, it would be most desirable to have interference between cells be naturally small. This can indeed be obtained by moving to *small cells*, as we shall see in Section 1.3.

1.3 The Case for Small-Cell Architectures

In the wireless industry, the term *small cell* refers to any cell that is smaller in area than the typical cell of a cellular deployment today, called a *macrocell*, which has an area of about 0.5–2 km^2. There is a loose classification of small cells by size, with the names *microcell* and *picocell* frequently used in the literature. The analysis in this book applies to all small cells of any size, and indeed to macrocells too.

At any rate, even if a base station transmits with power comparable with that of a macrocell base station, its area of coverage will be a small cell if the transmission occurs in a high-frequency band, because propagation losses increase with frequency. In other words, a simple and elegant way to get small cells, and thereby obtain the benefits of spectrum reuse, is to employ high-frequency spectrum. Of course, even without switching to a different spectrum from macrocells, the transmission powers of small-cell base stations can be scaled down compared with macrocells, to reduce the size of the cell.

In short, a smallcell (as opposed to a macrocell) architecture brings the benefits of spectrum reuse while also offering the possibility of increased spectral efficiency (simply because the users are now nearer to the base stations than they would be in macrocells) and reduced equipment costs.[5] These are the primary reasons why small-cell architectures play such an important role in future wireless network deployments.

Consider the following rough calculation (Andrews 2010): take an urban area with a medium-to-high user density of $10,000\,\text{km}^2$, and suppose each of these users is to be served at an average rate of 1 Mbps. Then the required network capacity is $10\,\text{Gbps/km}^2$. If we tried to support this with $1\,\text{km}^2$-area macrocells, then, even with the unrealistic assumption of a spectral efficiency of 2 bps/Hz, we would require a bandwidth of 5 GHz! On the other hand, if we used small cells with an average radius of 100 m (instead of the roughly 600-m radius of a macrocell with area $1\,\text{km}^2$), then the density of base stations rises by a factor of approximately 30, and the bandwidth requirement may be satisfied with as little as 500 MHz.

1.4 Future Wireless Networks Will Be Heterogeneous

With all the promise of small-cell architectures, it should not be forgotten that nearly all present-day wireless cellular networks use macrocells. Network operators have spent

[5] Although each small-cell base station may cost less than a macrocell base station, the density of small-cell base stations is greater than that of macrocell base stations, so whether or not CapEx is reduced depends on the actual prices and densities.

enormous amounts of money and time in deploying and optimizing them, and they expect to amortize these investments over several decades. In other words, the existing macrocellular networks are not going to go away. It follows that the wireless networks of the future will be *heterogeneous* in nature, comprising a macrocell *tier* overlaid with one or more tiers of small cells. Thus the design of wireless networks that meet the desired criteria of affordability, availability, and quality must take into account the fact that such networks are likely to be HetNets.

Notwithstanding the benefits from operating small-cell networks in a different (higher) frequency band from the existing macrocells, it is crucial to study scenarios where the small-cell and macrocell tiers use the same frequency band. This is because nearly all deployments today are macrocellular, so wireless devices support only the bands used in macrocells, hence upcoming small-cell deployments are also more likely than not to use the same bands. However, there are some interesting proposals for future wireless networks that envision entirely different classes of devices that operate solely on dedicated frequency bands in small cells. In this book, we analyze both classes of architectures (with overlapping and nonoverlapping spectrum between the tiers).

1.5 Approaches to the Design of Future Wireless Networks

A key part of the standards development process for future wireless networks is the evaluation of network performance for proposed network architectures and transmission schemes. This evaluation is mostly done according to the following two criteria (Novlan et al. 2015):

1. The plot of block error rate (BLER) vs. SINR on an arbitrarily selected link for the proposed modulation and coding scheme.
2. The cumulative distribution function of the SINR at an arbitrarily selected user in the network, with the details of the modulation and coding scheme abstracted for the sake of simplicity.

The symbol error rate (SER) (or BLER) vs. SINR plot on a single link has traditionally always been analyzed via detailed (down to the bits going into transmitted packets) simulation, because there are no exact analytical results for almost any modulation and coding schemes (although bounds on the SER or BLER can often be derived). These simulators are called link-level simulators.

In this book, we are interested in the aggregate network performance involving multiple users in multiple cells. As described in detail in Section 1.8 until about 2010 there were no analytical results on SINR distribution at an arbitrary user location in such a network. Thus, the SINR distribution at an arbitrary location had to be found via simulation. For the sake of computational feasibility, such multicell simulations usually abstracted the individual link-level packets and their transmission and reception, replacing the link-level behavior with look-up tables modeling what was separately obtained via link-level simulation. These multicell simulations with abstracted physical layer

coding/decoding are called *system-level simulations*. In this chapter, when we talk of simulations to evaluate network performance, we mean these system-level simulations.

As stated in the preceding paragraph, until about 2010 the simulation approach was the only one available to evaluate the performance of proposed network architectures. To generate results in a consistent format to compare standards proposals by different companies, the standards bodies defined detailed scenarios for simulation together with detailed lists of results that had to be obtained from such simulation. For example, there were several different possible base station layouts, for each of which several different antenna models had to be simulated, and for each of which in turn several different propagation models had to be simulated. Even though nearly all simulation scenarios are for single-tier macrocell layouts, the number of total scenarios to evaluate via simulation is often very large, requiring several person-weeks of effort to code, debug, and run in each company before every major standards meeting. For example, a typical exhaustive simulation-based investigation for a macrocellular network might include all combinations from a set of two layouts, three or four antenna heights, three cell densities, four or five base station transmit powers, and ten user densities (Novlan et al. 2015).

1.6 The Case against Pure Simulation-Based Investigation

The above set of scenarios to simulate is daunting enough, but even if they are all simulated in full, the total set of all possible scenarios is too large to be explored comprehensively with only a handful of selections each for antenna heights, cell density, base station transmit power, and user density. In other words, the comparison of simulation results across companies for these numerous, yet still limited, set of scenarios serves mainly to calibrate the simulators developed independently by the different companies. It is not even obvious that a significant fraction of practically useful usage situations are captured and/or modeled by the handful of scenarios defined for simulation. It follows that the results for the simulated scenarios may not necessarily yield useful insights as to network performance in other scenarios that have not been simulated.

These comments apply to a single-tier macrocell network. Note that the number of scenarios to be simulated will be squared if there is a second, small-cell, tier overlaid on the macrocell tier, and even if this resulting enormous number of simulations were done, they would cover an even smaller fraction of the space of all possible scenarios than was true for a single-tier macrocell network.[6]

To summarize, exclusive reliance on simulation to evaluate multitier HetNet performance may not only be computationally infeasible, it may even be misleading if used as a basis for HetNet design.

[6] This is a consequence of the *curse of dimensionality*, which in this case is simply the fact that if $f < 1$ is the fraction of scenarios simulated for the single-tier network, then even if we simulated the same set of scenarios for each tier of a two-tier network, the resulting fraction of scenarios covered is $f^2 < f$.

1.7 The Case for an Analytical Approach to HetNet Design

Suppose that, by some means, we could sift through the space of scenarios to definitively eliminate those that are *not* going to yield the performance we seek (as measured by the distribution of the SINR at an arbitrarily located user, say). Then we could attempt careful, thorough, and detailed simulation of the remaining scenarios of practical interest with the assurance that we did not overlook any potentially useful or important scenario.

We would have a quick way to eliminate network architecture proposals that do not result in the desired performance if, for example, there existed analytical or semi-analytical results relating the user SINR distribution to the HetNet deployment parameters such as base station transmit powers and densities in the tiers. These analytical results would still be useful even if they were obtained for a slightly simplified (but still realistic) HetNet model, so long as we had reason to believe that the relative performances of two sets of deployment parameters, as obtained via analysis, would mirror what we would obtain via simulation.

The message of this book is that such analytical and semianalytical results do exist, covering a wide range of HetNet models. Nearly all of these results were derived since 2010. In subsequent chapters, we develop the analytical techniques whereby such results can be derived, but an overview of the principal results obtained from this approach is given in Section 1.8.

1.8 The Stochastic Geometric Approach to HetNet Analysis

The set of analytical techniques used to study HetNets that is described in this book comes from a field of mathematics called *stochastic geometry*. The key idea in the application of stochastic geometry to wireless cellular network analysis is to model the locations of base stations and user terminals as realizations of a class of random mathematical objects called *point process*.

1.8.1 A preview of the main results in the book

The remainder of the book is devoted to developing such models and deriving analytical expressions for relevant network performance criteria such as user SINR distribution, beginning with a self-contained overview of the basic theory of point processes in the next chapter. Some of the major results that have been derived are listed below.

1. Consider an arbitrary layout of a tier of base stations in the plane with independent and identically distributed (iid) log-normal[7] fading on all links to the user. Then, under mild conditions on the density of the base stations, asymptotically in the standard deviation of the log-normal fading, the point process of link losses at the user terminal converges in distribution to the point process of link losses from a tier of base stations whose locations are the points of a Poisson point process,

[7] See Appendix C for more details on the probability distributions that appear in this book.

which is often abbreviated PPP or called just the Poisson process (Błaszczyszyn, Karray, & Keeler 2013).

2. This result holds irrespective of the propagation loss model.
3. This result also holds if the iid log-normal fading is replaced by iid Suzuki fading, which combines log-normal *shadow fading* or *shadowing*, with Rayleigh fading (Błaszczyszyn, Karray, & Keeler 2015).
4. For an arbitrarily located user in a tier of base stations whose locations are the points of a homogeneous (in other words, constant-density) Poisson process, with arbitrary iid fading to the user, the point process of link losses at this user terminal from all base stations in the tier has the same intensity function as the path-loss (with no fading) point process from a different homogeneous Poisson-process-located tier of base stations whose density is that of the original tier of base stations times a moment of the fading coefficient. In other words, a network with iid fading can be replaced for analysis purposes by an equivalent network with no fading (Błaszczyszyn, Karray, & Klepper 2010; Haenggi 2008; Madhusudhanan et al. 2014).
5. The signal-to-signal-plus-interference ratio, called *signal-to-total-interference ratio* or STIR, at the typical location from all base stations in a given tier (modeled as a homogeneous Poisson process) is a type of Poisson-Dirichlet point process (Błaszczyszyn & Keeler 2014), and results on this process also hold for the process formed by the STIR and, in some instances, the *signal-to-total-interference-plus-noise ratio* values.
6. A consequence of the this is that the joint distribution of the n strongest STIRs from each tier is available in exact analytic form (Handa 2009).
7. Similarly, the joint distribution of the SINR at the typical location from the *nearest* base stations (one per tier) in the accessible tiers of the HetNet (with the locations of the base stations in the tiers modeled by independent homogeneous Poisson processes) can be obtained in exact analytic form (Mukherjee 2012).

The above selection of results concerns a HetNet where the locations of base stations in the tiers are modeled as independent homogeneous Poisson processes. Although the Poisson model has great advantages for analytical tractability, and is also the asymptotic limit of an arbitrary base station layout, it is not necessarily the best model for a finite layout of base stations.

Fortunately, we note that although the model of independent Poisson point process tiers for the HetNet does not fit real-world SINR distributions exactly, in several practically important scenarios it is possible to "shift" the Poisson process curves by a known amount to obtain excellent agreement with real-world measurements. This shift depends only on a quantity called the *mean interference-to-signal ratio* (MISR). The resulting method, called *approximate signal-to-interference ratio analysis based on the Poisson point process*, provides a quick and computationally efficient scheme that is eminently suitable for engineering applications (Guo & Haenggi 2015). As one example, we also study the use of the MISR in approximating the ergodic spectral efficiency of links under advanced transmission techniques such as interference alignment.

We shall also see from our study of two-tier HetNets that the design of such HetNets naturally leads to a deployment where the macrocellular base stations are used for coverage as the primary goal and the small-cell tiers are used for capacity as the primary goal. This also conforms with network design choices to support user handoff between base stations, namely that the macrotier is designed to support such handoff but the small-cell tier does not support it very well (principally because of the control signaling overheads and latencies involved, and the frequency of needing to do handoffs if the cells are small). This coverage vs. capacity choice in deployments has implications for the best choice of stochastic model for these deployments, as discussed in the next section.

1.8.2 Extension to non-Poisson point processes

Although the Poisson point process is the best-understood model, it is not the only analytically tractable model. In fact, the Poisson model is a special case that may be seen as belonging to two distinct, important, classes of models: the *determinantal point processes* and the *permanental point processes*. The key distinction between these two classes of models is that the determinantal point processes tend to be *repulsive* (meaning the points tend to be farther from one another than in a Poisson process), whereas the permanental point processes tend to be *attractive* (meaning the points tend to be closer to one another than in a Poisson process). Determinantal processes have been proposed as tractable network models, but there is very little or no work involving permanental network models. This imbalance is perhaps partly due to determinantal processes being historically considered more tractable than permanental processes. But permanental processes form a subclass of point processes known as Cox point processes, which are tractable in some cases. Cox processes are a generalization of Poisson processes and are also attractive, resulting in (nonpermanental) Cox network models being proposed.

For repulsive and attractive point processes, we can consider two deployment scenarios.

1. *Coverage-oriented deployments*: In coverage-oriented deployments, the designers try to maintain a certain minimum separation between neighboring base stations. Thus, determinantal processes would seem to model such coverage-oriented deployments better than Poisson processes. Indeed, it has been seen that for certain measures on real-world macrotier deployments, principally coverage probability, determinantal processes give a better fit than Poisson point processes (Li et al. 2015). We therefore provide here a discussion of determinantal models for such base station layouts.

   As with the Poisson models, the expressions for SINR distributions obtained from the determinantal models can be derived exactly, and the book shows how to perform such calculations. However, the resulting expressions are complicated and involve multidimensional integrals that are challenging to evaluate numerically. Thus, we include a discussion of recent (since 2008) methods to evaluate such high-dimensional integrals numerically using a variant of the Monte Carlo method called quasi-random Monte Carlo.

2. *Capacity-oriented deployments*: On the other hand, the point processes whose points tend to be closer to one another than in a Poisson point process would seem to model base stations in hot spots, that is, capacity-oriented deployments, better than Poisson models. We discuss the known results for certain Cox processes.

1.8.3 Applications to link-level analysis

Note that this discussion has primarily been about applying techniques of stochastic geometry to determine analytically the distribution of SINR at an arbitrarily located user in the network. That is the focus of this book. As discussed in Section 1.5, the results derived in this book enable us to avoid many tedious and time-consuming *system-level* simulations. Recall that system-level simulators are built using results obtained from *link-level* simulators, which yield curves for SER (or BLER) vs. SINR. The results in the present book on SINR distribution are not directly applicable to the analysis of SER (or BLER) vs. SINR, but it turns out that stochastic geometry can also be applied with some success to derive the relationship between SER and SINR (ElSawy et al. 2017), thereby reducing our reliance on link-level simulations as well.

2 The Role of Stochastic Geometry in HetNet Analysis

In Chapter 1, Section 1.8, we list several important and general analytical results for single-tier networks and multitier HetNets that were derived using results from stochastic geometry. Of course, the value of analytical results in yielding insights to help network design has long been recognized by the industry. Stochastic geometric approaches are only the latest, and by far the most successful, of a long line of analytical approaches that began almost immediately after the first "cellular" models for wireless communication were proposed. Before commencing a detailed study of stochastic geometry, it is useful to take a step back and examine competing analytical approaches to get a better idea of why stochastic geometric approaches achieved greater success. To do so, let us revisit the origins of the now-canonical hexagonal cell model.

2.1 The Hexagonal Cellular Concept

The earliest proposals for wireless coverage via hexagonal "cells" date back to 1947, even before the principles of information theory were published by Shannon in 1948. A more up-to-date overview of the cellular concept may be found in MacDonald's 1979 paper (MacDonald 1979), applied to what is now considered the "first-generation" or "1G" cellular standard, advanced mobile phone service. MacDonald's carefully reasoned rationale for the hexagonal cellular model bears reading even today[1]:

1. He acknowledged that a practical cellular deployment would contain cells of different shapes and sizes (owing to topographic obstacles, restricted rights-of-way, etc.), but made the reasonable argument that systematization of the design and layout of a cellular system is greatly aided by modeling all cells of the system as having the same shape.
2. At the time MacDonald wrote his paper, cellular base stations transmitted using omnidirectional antennas. This in turn means that the boundary of the coverage region of a base station is circular. In other words, omnidirectional transmitting antennas at the base stations mean circular disk-shaped cells.

[1] As an aside, although the paper is almost entirely about a single-tier macrocell layout, MacDonald even discussed a scenario with a tier of (also hexagonal) small cells overlaid on the macrocell tier, but only to propose a dedicated allocation of spectrum to the small-cell tier.

3. Unfortunately, circular disks do not *tile* (meaning, cover completely, without overlaps or gaps) the two-dimensional plane.
4. Three shapes that do tile the plane are equilateral triangles, squares, and (regular) hexagons.
5. For each of these three shapes, and with the base station located at the center of a cell with that shape, the farthest point within the cell is a vertex of that shape.
6. To assure satisfactory received signal strength at this farthest point (that is, "worst-case" or "cell-edge" coverage), the distance between the center of the cell and the vertex must be kept below a maximum.
7. For a fixed center-to-vertex distance, the hexagonal shape has a larger area than an equilateral triangle or a square.
8. This implies that a cellular deployment with hexagonal cells requires fewer base stations to cover the same region than either of the other shapes.

Notwithstanding its prescience on many matters, MacDonald's paper contains almost no analytical results, except for some basic calculations of distances between points located on a hexagonal lattice. In particular, there is no quantitative analysis of spectral efficiency or signal-to-interference ratio (SIR) distributions, because closed-form analytical results for hexagonal lattices do not exist.

In fact, these arguments in favor of the hexagonal cell shape implicitly assume that the criterion for network design is the distribution of received signal power at a user location, and not the SIR distribution at the user terminal. However, the relationship between the data rate on an individual link to a user terminal (or, indeed, the network average throughput per unit area, which we identified as one of the key criteria for evaluating the performance of a network) and the distribution of SIR at this user's location, means that the design of the network must be driven by the distribution of SIR rather than the distribution of signal power. Before we proceed further, let us take a moment to provide a clear definition of what we mean by the received signal power and the SIR (and its close relative, the SINR [signal-to-interference-plus-noise ratio]) of a link, and in the process introduce some basic notation that is used throughout the book.

2.2 Propagation, Fading, and SINR

When a transmitter (say, a base station) transmits with power P_{tx}, the received power P_{rx} at a receiver (say, a user terminal) at a distance of d from the transmitter is given by

$$P_{\text{rx}} = P_{\text{tx}} \times G_{\text{tx}} \times G_{\text{rx}} \times H \times \ell(d). \tag{2.1}$$

Note that in Eq. (2.1), G_{tx}, G_{rx}, H, and $\ell(d)$ are all dimensionless.

The quantities G_{tx} and G_{rx} are the gains of the transmit and receive antennas, respectively.

The random variable H is the *fading coefficient* on the link. Note that in industry terminology, the fading is described by the probability distribution of the amplitude of the equivalent baseband complex waveform that is actually transmitted over the wireless

channel. Thus, if the link is said to be subject to Rayleigh fading, then the distribution of H, which applies to the received power and not the received signal amplitude, is that of the square of a Rayleigh-distributed random variable, and is therefore exponential. Without loss of generality, the expected value of H can be set to unity, and it is customary to do so. Thus, in the case of Rayleigh fading on this link, H is exponentially distributed with unit mean, and its probability density function is given by

$$f_H(x) = \exp(-x), \quad x > 0. \tag{2.2}$$

In Eq. (2.1), $\ell(d)$ is called the *path loss* over the link where d is the transmitter-receiver distance.[2] A widely used model has the form

$$\ell(d) = K(d/d_{\text{unit}})^{-\alpha}, \tag{2.3}$$

where $d_{\text{unit}} = 1\,\text{m}$ and α is called the *path loss exponent*. The free-space path loss exponent is $\alpha = 2$, so $\alpha > 2$ for nearly all links, except when the link is in the form of a waveguide (for example, a tunnel). In addition, α is dependent upon the frequency band. In Eq. (2.3), the dimensionless quantity K is a function of the geometry of the link (mostly, the relative heights of the transmitter and receiver, with some dependence upon the frequency band of the transmission).

It is customary in the industry to work with powers and gains in decibels. Rewriting Eq. (2.3), we have

$$10\,\log_{10}[\ell(d)] = -10\alpha \times \log_{10}(d/d_{\text{unit}}) + 10\,\log_{10} K, \tag{2.4}$$

or equivalently in decibels,

$$(\ell(d))_{\text{dB}} = -\alpha \times (d/d_{\text{unit}})_{\text{dB}} + (K)_{\text{dB}}, \tag{2.5}$$

which is the equation of a straight line with slope $-\alpha$ and intercept $(K)_{\text{dB}}$. This is why Eq. (2.3) is often called the *slope-intercept* model. For notational convenience in the future, we shall define the quantity

$$\kappa = \frac{K^{-1/\alpha}}{d_{\text{unit}}}, \tag{2.6}$$

so that Eq. (2.3) may now be written as

$$\ell(d) = (\kappa d)^{-\alpha}. \tag{2.7}$$

Note that unlike K, κ is not dimensionless.

The receivers at the user terminals always have a certain level of background noise power, called *thermal noise power*, which we denote by N.

Suppose a user terminal is receiving from its serving base station at a distance of d_0 while n other base stations at distances $d_1, \ldots, d_n$ from this user location are simultaneously transmitting (to other users served by them). Assuming all base stations are identical and transmit using omnidirectional antennas with the same power P_{tx}, and the fading coefficients on the links from the n interfering base stations to the user are iid

[2] Note that, following the convention in the wireless literature, we call $\ell(d)$ a *loss* even though it multiplies the transmit power, and is therefore actually a *gain*, albeit of magnitude less than unity.

random variables $H_1, \ldots, H_n$, assumed independent of the fading coefficient H_0 on the link between the user and its serving base station, then the received power at the user from its serving base station, called the *received signal power*, is given by

$$S = PH_0\ell(d_0), \tag{2.8}$$

where

$$P = P_{\text{tx}}G_{\text{tx}}G_{\text{rx}}.$$

Note that the antenna gains G_{tx} and G_{rx} have no units, so P may be treated as the effective transmit power of each base station. The interference power at the user location is similarly given by

$$I = \sum_{i=1}^{n} PH_i\ell(d_i). \tag{2.9}$$

Finally, the SIR and SINR are defined by

$$\text{SIR} = \frac{S}{I} \quad \text{and} \quad \text{SINR} = \frac{S}{I+N}. \tag{2.10}$$

The distribution of SI(N)R is the single most important quantity in the design and analysis of wireless communication networks because it simply determines the quality of communication channels, their sustainable bit rates, outage probabilities, etc. Specific relations between SI(N)R and channel quality metrics depend on particular communication technology, such as the coding schemes, type of multiplexing, and so forth. Information theory provides theoretical upper bounds on these metrics as functions of SI(N)R for several mathematical models of communication channel. For example, Shannon's celebrated result (Cover & Thomas 1991) states that in the additive white Gaussian noise (AWGN) channel a given transmission bit rate is sustainable in the long term (that is, such transmissions over such an AWGN channel can be decoded without error) if and only if this transmission bit rate is smaller than $W \log_2(1 + \text{SNR})$, where W is the channel bandwidth.

The distribution of SI(N)R at the user location for the above scenario has an elegant form when the link to the serving base station at distance d_0 has Rayleigh fading. Then H_0 has the probability density (function) (Eq. [2.2]), and from Eqs. (2.8) and (2.10), the complementary cumulative distribution function (CCDF) of SIR can be derived as follows:

$$\mathbb{P}\left\{\frac{S}{I} > x\right\} = \mathbb{P}\left\{H_0 > \frac{x}{P\ell(d_0)}I\right\} = \mathbb{E}\exp\left(-\frac{x}{P\ell(d_0)}I\right) = \mathcal{L}_I\left(\frac{x}{P\ell(d_0)}\right), \quad x > 0, \tag{2.11}$$

where, for any random variable X,

$$\mathcal{L}_X(s) = \mathbb{E}\exp(-sX), \quad s \geq 0 \tag{2.12}$$

is the Laplace transform of X. Note that from Eq. (2.11), the distribution of SIR is given in terms of the Laplace transform of interference power I at the user. We next discuss the circumstances under which $\mathcal{L}_I(\cdot)$ can be derived analytically.

2.3 Base Station Locations Modeled by Point Processes

When we re-examine MacDonald's arguments from the perspective of SIR distribution, it is obvious that we no longer require cells to have the same shape (that is, there is no need for a regular tiling of the plane). Instead, it is apparent that "irregular" layouts of base stations can still lead to the same cell-edge distribution of SIR. Moreover, such irregular base station layouts are a better match to real-world network deployments than a hexagonal lattice.

It is appealing to propose a "realistic" model of base station locations by "perturbing" the points of a hexagonal lattice in some way (see Banani, Eckford, & Adve 2014, for a modern example of this approach, including a comparison to the regular hexagonal lattice and the Poisson point process). Unfortunately, it turns out that perturbed hexagonal lattices are no easier to analyze exactly than the regular hexagonal lattice. However, the modeling of these "perturbations" by random variables means that the resulting base station locations are now random. This idea leads naturally to modeling base station locations as realizations of the class of stochastic processes called *point processes*.[3]

Once we accept the idea of modeling base station locations by point processes, we can focus on the reasons for analytical tractability using results from stochastic geometry. As we discuss in Chapter 4, the main reasons are the following:

1. assumption of iid fading on all links to a user terminal from all points of the point process;
2. existence of Campbell theorems yielding analytical expressions for the *probability generating functional* of a point process, resulting in an analytical expression for the Laplace transform of the interference power at the user terminal; and
3. assumption of Rayleigh fading on the link between the user terminal and its serving base station, which allows us to write an expression for the CCDF of the SIR at the user terminal in terms of the Laplace transform of the interference power *without needing to invert this Laplace transform*.

It is not surprising that the first successful examples of such analysis were for base station locations modeled as points of homogeneous Poisson point processes, which are the most analytically tractable point processes. However, recent advances have extended the set of analytic results on user SIR distribution to encompass other point processes such as Ginibre and Cox point processes.

To proceed further, we require the contents of the next chapter, Chapter 3, which provides a rigorous and detailed introduction to the theory of point processes.

[3] A point process may be loosely described as a random pattern of points, and indeed a realization of a point process is called a *point pattern*.

3 A Brief Course in Stochastic Geometry

3.1 Purpose

In this chapter, we give a brief introduction to stochastic geometry, in particular point process theory, with a focus on the models and results that are used in the other chapters of the book. It is assumed that the reader has some familiarity with the concepts of stochastic geometry. Most results are presented without proofs. More detailed introductions to the theory and proofs can be found in Chiu et al. (2013), Daley & Vere-Jones (2008), Haenggi (2012), and Last & Penrose (2017).

3.2 Fundamental Definitions and Notation

In this section, we formally define point processes, also called random point fields, and introduce the notation that is used throughout the book. We restrict ourselves to point processes on the Euclidean space $\mathbb{R}^d$.

3.2.1 Definition

We call $\varphi = (x_k)_{k\in\mathcal{K}}$ where $x_k \in \mathbb{R}^d$ and $\mathcal{K} = \mathbb{N}$ or $\mathcal{K} = [n]$, a (deterministic) point pattern or point sequence. The space of all sequences is denoted by $\mathcal{N}$. We can view $\mathcal{N}$ as the space of all n-dimensional vectors with elements in $\mathbb{R}^d$, where $n = \infty$ is allowed. Usually the ordering of the points is irrelevant and thus φ can also be considered as a multiset and is sometimes referred to as a collection of points. If $\varphi_k \neq \varphi_i$ for $k \neq i$, the point pattern is called *simple*, in which case φ is just a countable subset of $\mathbb{R}^d$.

The *counting measure* $N_\varphi \colon \mathcal{B} \mapsto \mathbb{N}_0$ of φ is defined as

$$N_\varphi(B) \stackrel{\triangle}{=} \sum_{k\in\mathcal{K}} \mathbf{1}_B(x_k).$$

As the name suggests, it counts the number of elements of φ that fall in B.

If $\mathcal{N}$ is the outcome space of a random process, then this random process is called a *point process*, defined as follows.

DEFINITION 3.2.1 (Point process) *A point process Φ is a measurable map $\Phi \colon \Omega \mapsto \mathcal{N}$, where $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.*

A more mathematically formal definition requires a σ-algebra of measurable subsets, which we introduce in Section 3.2.3.

The associated counting measure N_Φ then is a *random measure*, and $N_\Phi(B)$ is an $\mathbb{N}_0$-valued random variable. Measurability requires that $N_\Phi(B)$ be defined for a sufficiently rich class of subsets of $\mathbb{R}^d$, and the natural choice is the Borel σ-algebra denoted as $\mathcal{B}$.

A point process can be equivalently defined and interpreted as a random point pattern and as a random counting measure. If the realizations of Φ are simple *almost surely* (or with probability one) the point process can be viewed as a random set $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$. The equivalence between (countable) random sets and random measures is discussed in detail in the next subsection.

3.2.2 Equivalence of random sets and random measures

We first introduce the concept of a measurable decomposition.

DEFINITION 3.2.2 (Measurable decomposition) *The* measurable decomposition *of a simple point process Φ is the decomposition of the counting measure into Dirac measures:*

$$N_\Phi \equiv \sum_{\mathbf{x}\in\Phi} \delta_{\mathbf{x}}\,. \tag{3.1}$$

$\delta_x(B) = \mathbf{1}_B(x)$ is the Dirac measure at x for $B \in \mathcal{B}$. It is 1 if $x \in B$ and 0 otherwise. The Dirac measure and thus N_Φ are *atomic measures*, that is, they are concentrated on a countable collection of locations.[1] Hence the random measure is obtained easily from the random set. Conversely, we can retrieve the random set from the counting measure by setting

$$\Phi \equiv \{x \in \mathbb{R}^d \colon N_\Phi(\{x\}) = 1\}.$$

As a result, henceforth we use Φ to denote both the random set and the counting measure; that is, we write $\Phi(B)$ instead of $N_\Phi(B)$ to denote the number of points in B. This notation is simple and widely used. Similarly, we may view φ as a counting measure and write, for example, $\varphi = \delta_x + \delta_y$ instead of $\varphi = \{x, y\}$.

If $\mathbb{P}(\Phi(\{x\}) > 0) > 0$ for some $x \in \mathbb{R}^d$, the point process is said to have a *fixed atom* at location x. The set of all fixed atoms is a deterministic set, as it does not depend on the realization of a point process.

DEFINITION 3.2.3 (Intensity [mean] measure) *The intensity measure*[2] Λ *is defined as*

$$\Lambda(B) \stackrel{\triangle}{=} \mathbb{E}\Phi(B)\,, \quad \forall B \in \mathcal{B}\,.$$

At a fixed atom x, $\Lambda(\{x\}) > 0$, whereas for all other locations x, $\Lambda(\{x\}) = 0$. Other basic properties of point processes that are based on the counting or intensity measures are defined as follows.

[1] We use the term *location* instead of point when talking about a general element in $\mathbb{R}^d$, to avoid confusion with the point of a point process.

[2] Some authors call it *mean measure* and reserve the use of *intensity measure* for Poisson processes only.

DEFINITION 3.2.4 (Diffuse, atomic, and discrete point process) *We call a point process* diffuse *if its intensity measure is a diffuse measure. If it has one or more atoms, it is called* atomic. *If there exists a countable set X of locations such that* $\Phi(\mathbb{R}^d \setminus X) = 0$ *the process is called* discrete *or* purely atomic.

Hence in a discrete process, the entire mass of the intensity measure is concentrated on atoms. Because any locally finite measure is uniquely decomposable into a purely atomic and a diffuse measure, any point process is the superposition (or union) of a diffuse and a discrete point process.

3.2.3 Distribution of a point process

To define the distribution of a point process, we need to specify which subsets of $\mathcal{N}$ should be measurable. A basic set of events that need to be measurable is

$$E_{B,k} \stackrel{\triangle}{=} \{\Phi(B) = k\} = \{\varphi \in \mathcal{N} \colon \varphi(B) = k\}. \tag{3.2}$$

The σ-algebra induced by $E_{B,k}$, $B \in \mathcal{B}$, $k \in \mathbb{N}_0$ is sufficiently rich to make all events of interest measurable; it is denoted by $\mathfrak{N}$. This way, we have defined a measurable space $(\mathcal{N}, \mathfrak{N})$. By equipping it with the point process distribution P, we obtain the probability space $(\mathcal{N}, \mathfrak{N}, \mathsf{P})$. It is called the *canonical probability space* for point processes.

DEFINITION 3.2.5 (Point process distribution) *The distribution of* Φ *is the probability measure pertaining to the outcome measure space* $(\mathcal{N}, \mathfrak{N}, \mathsf{P})$*:*

$$\mathsf{P}(E) \stackrel{\triangle}{=} \mathbb{P} \circ \Phi^{-1}(E) \equiv \mathbb{P}(\Phi \in E) \qquad \forall E \in \mathfrak{N}\,.$$

Here $\Phi^{-1}(E)$ is the pre-image of E, given by

$$\Phi^{-1}(E) = \{\omega \in \Omega \colon \Phi^{\omega} \in E\},$$

and the symbol $\circ$ stands for the concatenation of functions, for example, $(f \circ g)(t) = f(g(t))$. Measurability requires that the pre-image of E is measurable with $\mathbb{P}$, that is, $\Phi^{-1}(E) \in \mathcal{F}$, which is guaranteed if $\mathcal{N}$ is defined as above and $(\Omega, \mathbb{P}, \mathcal{F})$ is an abstract probability space such as $([0,1], \mathcal{B}_{[0,1]}, \nu)$, where $\mathcal{B}_{[0,1]}$ is the Borel σ-algebra on $[0,1]$ and ν is the Lebesgue measure.

Because the outcome space has probability 1 we have

$$\mathsf{P}(\mathcal{N}) = \int_{\mathcal{N}} \mathsf{P}(\mathrm{d}\varphi) = 1,$$

and a general expectation over the point process is given by

$$\mathbb{E}f(\Phi) = \int_{\mathcal{N}} f(\varphi)\mathsf{P}(\mathrm{d}\varphi)$$

for $f \colon \mathcal{N} \mapsto \mathbb{R}$. As an example, the intensity measure can be expressed as

$$\Lambda(B) = \mathbb{E}\Phi(B) = \mathbb{E}\left(\int_B \Phi(\mathrm{d}x)\right) = \int_{\Omega} \Phi^{\omega}(B)\mathbb{P}(\mathrm{d}\omega) = \int_{\mathcal{N}} \varphi(B)\mathsf{P}(\mathrm{d}\varphi)\,.$$

A point process for which $\Lambda(B) < \infty$ if $|B| < \infty$ is called a point process with locally finite mean measure.

An important class of point processes are those whose distribution is translation-invariant. They are called *stationary*.

DEFINITION 3.2.6 (Stationary point process) *A point process* Φ *is* stationary *if its distribution is invariant to translations, that is, if* $\mathsf{P}(E) = \mathsf{P}(E + x)$ *for all* $E \in \mathfrak{N}$ *and all* $x \in \mathbb{R}^d$*, where* $E + x$ *is the event* E *translated by* $x \in \mathbb{R}^d$.

For an event $E \in \mathfrak{N}$, the translated event consists of the translated elements of E, defined as $E + x \triangleq \{\varphi \in E \colon \varphi + x\}$.

It follows from the definition that stationary point processes cover the entire space $\mathbb{R}^d$ and have an infinite number of points almost surely, and have an intensity measure that is proportional to the Lebesgue measure: $\Lambda(B) = \lambda|B|$. The proportionality constant λ is called the *intensity* of the stationary point process.

If the point process distribution is also invariant to rotations about the origin, the point process is called *motion-invariant*.

Another important point process property is *ergodicity*. It implies that ensemble averages equal spatial averages, that is, all statistical properties can be extracted from a single realization of the point process. Conversely, it means that probabilities of certain events, say, that a fixed location is covered at an SIR of θ, can be interpreted as the fraction of the area that is covered at θ in each realization of the point process. Another important consequence of ergodicity is that translation-invariant events are independent of themselves, meaning they have probability 0 or 1. This fact is often exploited in percolation theory.

3.2.4 Palm measures

It is often desirable to make statistical statements about a randomly chosen (or "typical") point of a point process. We cannot simply pick a point uniformly from an infinite number of points, and if we define a rule how to pick such a point, we introduce biasing because the rule will have to depend on the point's surroundings (or some other property of the point process realization). To resolve this issue, we instead define a conditional distribution of the point process given that it has a point at a certain location, which leads to the *Palm distribution* or *Palm measure* of the point process. Informally, it is the conditional distribution

$$\mathsf{P}_x(E) \triangleq \mathsf{P}(E \mid x \in \Phi) = \mathbb{P}(\Phi \in E \mid x \in \Phi), \quad E \in \mathfrak{N}.$$

This is not a formal definition because usually $\mathbb{P}(x \in \Phi) = 0$, and it is not possible to use the simple quotient of the probabilities $\mathbb{P}(\Phi \in E, x \in \Phi)$ and $\mathbb{P}(x \in \Phi)$. A formal way of defining the Palm measure is via an integral relationship as follows.

DEFINITION 3.2.7 (Palm measure and expectation) *For a point process* Φ *of locally finite mean measure* Λ*, the family of probability measures* P_x*,* $x \in \mathbb{R}^d$*, defined on* $(\mathcal{N}, \mathfrak{N})$ *and satisfying*

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi}\mathbf{1}(\mathsf{x}\in B,\Phi\in E)=\int_B \mathsf{P}_x(E)\Lambda(\mathrm{d}x) \tag{3.3}$$

for all Borel $B\subset\mathbb{R}^d$ *and* $E\in\mathfrak{N}$*, is called the family of Palm measures or Palm distributions of* Φ*. Accordingly, the expectation with respect to the Palm measure* P_x *is defined as*

$$\mathbb{E}_x f(\Phi)\stackrel{\triangle}{=}\int_{\mathcal{N}} f(\varphi)\mathsf{P}_x(\mathrm{d}\varphi). \tag{3.4}$$

The left side of Eq. (3.3) is called the *Campbell measure*, often denoted as $\mathcal{C}(B,E)$, and from the equation we can see that the Palm measure is the Radon–Nikodym derivative of the Campbell measure with respect to the mean measure, that is, $\mathsf{P}_x(E)=\mathrm{d}\mathcal{C}(\cdot\times E)/\mathrm{d}\Lambda$.

Note that the definition of P_x is meaningful only up to a subset of locations x of null measure Λ. For example, it is not meaningful for $x=2$ for a point process on $\mathbb{R}$ with intensity measure $\Lambda(B)=|B\cap[0,1]|$. One can prove that $\mathsf{P}_x(\varphi(\{x\})\geq 1)=1$ for Λ-almost all $x\in\mathbb{R}^d$, which is consistent with the interpretation of P_x as the conditional probability of Φ given it has an atom at x. The point at location x under the Palm measure P_x becomes the *typical point* at that location.

In the stationary case, only one Palm measure is needed, because due to the translation invariance, $\mathsf{P}_x(E)\equiv\mathsf{P}_o(E-x)$ for all $E\in\mathfrak{N}$ and $x\in\mathbb{R}^d$. Hence, only P_o is needed, and the origin o is the typical point.

In general, the Palm measure of a stationary point process is not stationary, given that conditioning on $o\in\Phi$ may affect the distribution of points in the neighborhood of the origin, or even in the entire space. For example, if Φ is a randomly translated (and thus stationary) lattice $\mathbb{Z}^d+\mathbf{u}$, where $\mathbf{u}$ is uniformly chosen from $[0,1]^d$, conditioning on $o\in\Phi$ fixes all the points, hence the Palm measure is just the distribution of $\mathbb{Z}^d$ (which is degenerate because all mass is concentrated on the single point pattern $\varphi=\mathbb{Z}^d$).

The Palm measure of a stationary point process can be defined as follows. For an arbitrary $B\in\mathcal{B}$ with $|B|>0$,

$$\mathsf{P}_o(E)=\frac{1}{\lambda|B|}\mathbb{E}\left(\sum_{\mathsf{x}\in\Phi\cap B}\mathbf{1}_E(\Phi-\mathsf{x})\right). \tag{3.5}$$

$\Phi-x$ is the translated point process[3] $\{\mathsf{x}\in\Phi\colon \mathsf{x}-x\}$. In this expression, for each point of $\Phi\cap B$, the point process gets translated so that this point is at the origin, and the indicator is one if this translated point process belongs to the event E. The factor $1/(\lambda|B|)$ normalizes the sum by the expected number of points in $\Phi\cap B$.

Using the Palm measure, we can define quantities that are otherwise hard to define. An an example, consider the nearest-neighbor distance distribution, denoted by $G(r)$. Heuristically, we set $G(r)=\mathbb{P}(\mathsf{NND}\leq r)$ for a random variable NND representing the nearest-neighbor distance, but without the concept of the Palm measure, it would be cumbersome to rigorously define what NND is. Letting $E_r=\{\varphi\in\mathcal{N}\colon \varphi(b(o,r))>1\}$

[3] Not to be confused with $\Phi\setminus\{\mathsf{x}\}$, which may also be expressed as $\Phi-\delta_{\mathsf{x}}$ (counting measure).

be the event that the point process has more than one point in the disk of radius r centered at o, we can simply write

$$G(r) = \mathsf{P}_o(E_r) = \mathbb{P}(\Phi(b(o,r)) > 1 \mid o \in \Phi). \tag{3.6}$$

We denote by $\Phi_o \stackrel{\triangle}{=} (\Phi \mid o \in \Phi)$ the conditional point process corresponding to P_o, that is, the conditional point process that has a point at the origin. This allows us to define $\mathsf{NND} \stackrel{\triangle}{=} \min\{\mathsf{x} \in \Phi_o \setminus \{o\} \colon \|x\|\}$.

As this example shows, it is often useful to remove the atom at o from Φ_o and obtain more compact expressions. Also, in wireless networks, the point at o we condition on may be the desired transmitter while the other points are interfering transmitters. In this case, we may be interested in the interference at the receiver, which is a sum over $\Phi_o \setminus \{o\}$.

Disregarding the point o from Φ_o leads to a point process $\Phi_o^! = \Phi_o \setminus \{o\}$ whose distribution is governed by the *reduced Palm measure* $\mathsf{P}_o^!$, informally defined as

$$\mathsf{P}_o^!(E) \stackrel{\triangle}{=} \mathbb{E}\mathbf{1}_E(\Phi \setminus \{o\} \mid o \in \Phi) = \mathbb{E}\mathbf{1}_E(\Phi_o^!) = \mathbb{P}(\Phi_o^! \in E), \quad E \in \mathfrak{N}.$$

Analogously to the Palm measure, it is formally defined as follows.

DEFINITION 3.2.8 (Reduced Palm measure) *For a point process Φ of locally finite mean measure Λ, the family of probability measures $\mathsf{P}_x^!$, $x \in \mathbb{R}^d$, defined on $(\mathcal{N}, \mathfrak{N})$ and satisfying*

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} \mathbf{1}(\mathsf{x} \in B, \Phi - \delta_{\mathsf{x}} \in E) = \int_B \mathsf{P}_x^!(E)\Lambda(\mathrm{d}x) \tag{3.7}$$

for all Borel $B \subset \mathbb{R}^d$ and $E \in \mathfrak{N}$, is called the family of reduced Palm measures *(*distributions*) of Φ. Accordingly, the expectation with respect to the reduced Palm measure $\mathsf{P}_x^!$ is defined as*

$$\mathbb{E}_x^! f(\Phi) \stackrel{\triangle}{=} \int_{\mathcal{N}} f(\varphi)\mathsf{P}_x^!(\mathrm{d}\varphi). \tag{3.8}$$

Using the reduced Palm measure, we can write $G(r) = \mathsf{P}_o^!(\{\varphi \in \mathcal{N} \colon \varphi(b(o,r)) > 0\})$ and $\mathsf{NND} \stackrel{\triangle}{=} \min\{\mathsf{x} \in \Phi_o^! \colon \|x\|\}$.

3.2.5 Functions of point processes and the Campbell-Mecke theorem

All performance metrics of a wireless network depend on the node locations and thus on the point process modeling the transmitters and receivers. To derive these metrics, we need to evaluate the statistical properties of such functions of point processes. Stochastic geometry provides several powerful tools that help accomplish this.

Sums over point processes

The expected sum over a point process is given by *Campbell's formula*[4]

[4] We call this result Campbell's formula to distinguish it from a result on the moment-generating function of a Poisson point process that is often called Campbell's theorem.

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} f(\mathsf{x}) = \mathbb{E}\int_{\mathbb{R}^d} f(x)\Phi(\mathrm{d}x) = \int_{\mathbb{R}^d} f(x)\Lambda(\mathrm{d}x). \tag{3.9}$$

This is essentially a corollary of Fubini's theorem. In the stationary case, the expected sum is just the integral of f over the plane multiplied by the intensity λ:

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} f(\mathsf{x}) = \lambda\int_{\mathbb{R}^d} f(x)\mathrm{d}x$$

The function f is often taken to be an indicator function. For example, for $f(x) = \mathbf{1}_B(x)$, we obtain $\mathbb{E}\sum_{\mathsf{x}\in\Phi\cap B} 1 = \lambda|B|$ for the mean number of points in B, as expected.

The Campbell-Mecke theorem is an extension of Campbell's formula to functions on $\mathbb{R}^d \times \mathcal{N}$.

THEOREM 3.2.9 (Campbell-Mecke theorem) *For a point process Φ with intensity measure Λ and a nonnegative measurable function $g\colon \mathbb{R}^d \times \mathcal{N}$,*

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} g(\mathsf{x},\Phi) = \int_{\mathbb{R}^d}\int_{\mathcal{N}} g(x,\varphi)\mathsf{P}_x(\mathrm{d}\varphi)\Lambda(\mathrm{d}x) \tag{3.10}$$

and

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} g(\mathsf{x},\Phi\setminus\{\mathsf{x}\}) = \int_{\mathbb{R}^d}\int_{\mathcal{N}} g(x,\varphi)\mathsf{P}_x^!(\mathrm{d}\varphi)\Lambda(\mathrm{d}x). \tag{3.11}$$

The second result is the reduced form.

In the Campbell-Mecke theorem, the translation of φ and the use of the Palm measure is essential, because without them, the connection between x and Φ would be lost. On the left side of the equations, g is a function of x and a point process Φ containing x, not a function of a generic location and a point process. This needs to be reflected on the right side also.

For stationary point processes, the result can be stated using the Palm measure at o only.

COROLLARY 3.2.10 (Campbell-Mecke theorem for stationary point processes) *For a stationary point process and a nonnegative measurable function $g\colon \mathbb{R}^d \times \mathcal{N}$,*

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} g(\mathsf{x},\Phi) = \lambda\int_{\mathbb{R}^d}\int_{\mathcal{N}} g(x,\varphi+x)\mathsf{P}_o(\mathrm{d}\varphi)\mathrm{d}x \tag{3.12}$$

and

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi} g(\mathsf{x},\Phi\setminus\{\mathsf{x}\}) = \lambda\int_{\mathbb{R}^d}\int_{\mathcal{N}} g(x,\varphi+x)\mathsf{P}_o^!(\mathrm{d}\varphi)\mathrm{d}x. \tag{3.13}$$

Here, the Palm measures P_o and $\mathsf{P}_o^!$ put a point at the origin of φ and thus $\varphi + x$ has a point at x.

The result, Eq. (3.12), can also be stated as

$$\int_{\mathcal{N}}\sum_{x\in\varphi} g(x,\varphi)\mathsf{P}(\mathrm{d}\varphi) = \lambda\int_{\mathbb{R}^d}\mathbb{E}_o(g(x,\Phi+x))\mathrm{d}x.$$

For example, the mean interference at the origin in a stationary cellular network can be determined by setting

$$g(x, \Phi) = \ell(x)\mathbf{1}(\Phi(b(o, \|x\|)) > 0),$$

where $b(o, r)$ is the open disk of radius r centered at o. For $\mathsf{x} \in \Phi$, the indicator function is zero if and only if x is the point closest to the origin, hence this point is excluded from the sum.

Products over point processes

The expected product over a point process is called *probability generating functional* (pgfl), defined as follows.

DEFINITION 3.2.11 (Probability generating functional) *For a measurable function $v\colon \mathbb{R}^d \mapsto [0, 1]$ such that $1 - v$ has bounded support, the pgfl $G[v]$ of the point process Φ is defined as*

$$G[v] \stackrel{\triangle}{=} \mathbb{E} \prod_{\mathsf{x} \in \Phi} v(\mathsf{x}). \tag{3.14}$$

The above condition on v is sufficient and necessary for the pgfl to exist for all locally finite point processes. If the intensity measure of a point process is known, the less restrictive condition

$$\int_{\mathbb{R}^d} |\log v(x)| \Lambda(\mathrm{d}x) < \infty \tag{3.15}$$

is sufficient. This condition follows from the fact that the pgfl can be written as

$$G[v] = \mathbb{E} \exp\left(\int_{\mathbb{R}^d} \log v(x) \Phi(\mathrm{d}x) \right).$$

The pgfl is often used in the evaluation of Laplace transforms of sums over the point process. Letting $Z = \sum_{\mathsf{x} \in \Phi} f(\mathsf{x})$, we have

$$\mathcal{L}_Z(s) \stackrel{\triangle}{=} \mathbb{E}\mathrm{e}^{-sZ} = \mathbb{E} \prod_{\mathsf{x} \in \Phi} \mathrm{e}^{-sf(\mathsf{x})} = G[\mathrm{e}^{-sf(x)}].$$

From Eq. (3.15), the existence for stationary Φ is guaranteed if $\int_{\mathbb{R}^d} |f(x)| \mathrm{d}x < \infty$.

3.2.6 Moment measures and factorial moment measures and their densities

DEFINITION 3.2.12 (Moment measures) *The nth moment measure of a point process Φ is defined as the expected product of the number of points falling in regions $B_1, B_2, \ldots, B_n \in \mathcal{B}$:*

$$\mu^{(n)}(B_1 \times B_2 \times \cdots \times B_n) \stackrel{\triangle}{=} \mathbb{E}\big(\Phi(B_1)\Phi(B_2) \ldots \Phi(B_n)\big) \tag{3.16}$$

For $n = 1$, this reduces to the intensity measure. The nth moment measure can also be viewed as the intensity measure of the product point process $\Phi^{(n)} = \underbrace{\Phi \times \cdots \times \Phi}_{n \text{ times}}$. The elements of $\Phi^{(n)}$ are the (ordered) n-tuples $(\mathsf{x}_1, \mathsf{x}_2, \ldots, \mathsf{x}_n) \in \mathbb{R}^{nd}$, where $\mathsf{x}_k \in \Phi$.

For $n = 2$ and $B_1 = B_2 = B$, we have $\mu^{(2)}(B^2) = \mathbb{E}(\Phi(B)^2)$, and thus

$$\operatorname{var}(\Phi(B)) = \mu^{(2)}(B^2) - \Lambda(B)^2.$$

The disadvantage of the moment measures for $n \geq 2$ is that they are not absolutely continuous to the Lebesgue measure and hence do not have a density (with respect to the Lebesgue measure). The reason is that $\Phi^{(n)}$ has points that are concentrated on a subset of (nd-dimensional) Lebesgue measure 0. For example, for $d = 1$ and $n = 2$, let $\Phi = \{\mathsf{x}_1, \mathsf{x}_2, \ldots\}$. Then for all $k \in \mathbb{N}$, $(\mathsf{x}_k, \mathsf{x}_k) \in \Phi^{(2)}$, that is, $\Phi^{(2)}$ has points on the diagonal line $(x, x) \in \mathbb{R}^2$.

This issue can be resolved by prohibiting that the same point of Φ be taken twice (or more times) in $\Phi^{(n)}$. Accordingly, for a simple point process, we define the product point process of n-tuples consisting of *distinct* points

$$\Phi^{[n]} = \{(\mathsf{x}_1, \mathsf{x}_2, \ldots, \mathsf{x}_n) \in \Phi^{(n)} : \mathsf{x}_i \neq \mathsf{x}_k \ \forall i \neq k\}.$$

The intensity measure of $\Phi^{[n]}$ is called *factorial moment measure*.

DEFINITION 3.2.13 (Factorial moment measure) *The factorial moment measure of a point process Φ is defined as*

$$\alpha^{(n)}(B_1 \times B_2 \times \cdots \times B_n) \stackrel{\triangle}{=} \mathbb{E}\left(\sum_{\mathsf{x}_1, \ldots, \mathsf{x}_n \in \Phi}^{\neq} \mathbf{1}_{B_1}(\mathsf{x}_1)\mathbf{1}_{B_2}(\mathsf{x}_2) \ldots \mathbf{1}_{B_n}(\mathsf{x}_n) \right), \tag{3.17}$$

where the symbol $\neq$ indicates that the multisum is taken over distinct elements of Φ only.

If the regions B_k are all disjoint, then $\mu^{(n)} = \alpha^{(n)}$, because no point can fall in two disjoint regions. If $B_1 = B_2 = \cdots = B_n = B$, then

$$\alpha^{(n)}(B^n) = \mathbb{E}\big(\Phi(B)(\Phi(B) - 1)(\Phi(B) - 2) \cdots (\Phi(B) - n + 1)\big),$$

hence the name factorial moment measure.

If the factorial moment measure is absolutely continuous with respect to the Lebesgue measure,[5] we can express it using a density, which is called the n*th-order product density* or the n*th* moment density[6] and denoted by $\varrho^{(n)}$.

$$\alpha^{(n)}(B_1 \times \cdots \times B_n) = \int_{B_1} \ldots \int_{B_n} \varrho^{(n)}(x_1, \ldots, x_n) \mathrm{d}x_n \ldots \mathrm{d}x_1 \tag{3.18}$$

$\varrho^{(n)}$ is the intensity function of $\Phi^{[n]}$, hence we can use Campbell's formula to write

$$\mathbb{E}\left(\sum_{\mathsf{x}_1, \ldots, \mathsf{x}_n \in \Phi}^{\neq} f(\mathsf{x}_1, \ldots, \mathsf{x}_n) \right) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} f(x_1, \ldots, x_n)\varrho^{(n)}(x_1, \ldots, x_n) \mathrm{d}x_1 \ldots \mathrm{d}x_n.$$

[5] A necessary but not sufficient condition is that the point process be simple.

[6] Sometimes it is also called the factorial moment density, but because the higher moment measures do not have densities, there is no ambiguity in simply calling it the moment density.

Loosely speaking, because

$$\alpha^{(n)}(\mathrm{d}x_1 \times \cdots \times \mathrm{d}x_n) \equiv \varrho^{(n)}(x_1, \ldots, x_n)\mathrm{d}x_1 \ldots \mathrm{d}x_n,$$

the moment densities correspond to the joint probabilities of finding n points in infinitesimal balls around the locations x_1, x_2,..., x_n.

For a stationary point process, the moment densities are invariant to translations, that is, we have

$$\varrho^{(n)}(x_1, \ldots, x_n) \equiv \varrho^{(n)}(x_1 - y, \ldots, x_n - y), \quad y \in \mathbb{R}^d,$$

and we can set $y = x_1$ and eliminate one of the coordinates and define

$$\varrho^{(n)}_{\mathrm{st}}(x_1, \ldots, x_{n-1}) \stackrel{\triangle}{=} \varrho^{(n)}(0, x_2 - x_1, \ldots, x_n - x_1).$$

This is particularly useful if $n = 2$. If the point process is motion-invariant, the second moment density only depends on the distance $\|x_2 - x_1\|$ between the two arguments of $\varrho^{(2)}$, and we may define $\varrho^{(2)}_{\mathrm{mi}}(r)$ such that

$$\varrho^{(2)}_{\mathrm{st}}(x) \equiv \varrho^{(2)}_{\mathrm{mi}}(\|x\|).$$

When normalized by the squared intensity, we obtain the *pair correlation function*

$$g(x) \stackrel{\triangle}{=} \frac{\varrho^{(2)}_{\mathrm{st}}(x)}{\lambda^2}.$$

Reduced second moment measure

An analogous elimination of one dimension is also possible in the second factorial moment measure.

DEFINITION 3.2.14 (Reduced second moment measure) *Let Φ be a stationary point process on $\mathbb{R}^d$ with intensity λ. Then there is a measure $\mathcal{K}$ on $\mathbb{R}^d$ such that for measurable f,*

$$\mathbb{E}\left(\sum_{x,y\in\Phi}^{\neq} f(x, y)\right) = \lambda \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} f(x, x + u)\mathrm{d}x\, \mathcal{K}(\mathrm{d}u)\,.$$

$\mathcal{K}$ is called the reduced second moment measure[7] *of* Φ.

Hence $\varrho^{(2)}_{\mathrm{st}}/\lambda$ is the density of $\mathcal{K}$ with respect to the Lebesgue measure:

$$\mathcal{K}(B) = \frac{1}{\lambda} \int_B \varrho^{(2)}_{\mathrm{st}}(u)\mathrm{d}u\,. \tag{3.19}$$

Alternatively, $\mathcal{K}$ can be defined using the Palm measure:

$$\mathcal{K}(B) \stackrel{\triangle}{=} \int_{\mathcal{N}} \varphi(B \setminus \{o\})\mathsf{P}_o(\mathrm{d}\varphi) = \int_{\mathcal{N}} \varphi(B)\mathsf{P}^!_o(\mathrm{d}\varphi) = \mathbb{E}^!_o\Phi(B)\,.$$

[7] Note that there are some variations in the definition of $\mathcal{K}$ in the literature. Other versions used are scaled by λ or $1/\lambda$ relative to our definition here.

Hence, $\mathcal{K}(B)$ is the mean number of points in $\Phi \cap B \setminus \{o\}$ under the condition that $o \in \Phi$, that is, it is the intensity measure of the reduced Palm distribution.

For motion-invariant processes, a simpler function is often useful, namely *Ripley's K function*, also called *reduced second moment function* (Ripley 1976).

DEFINITION 3.2.15 (Ripley's K function) *The K function is defined as*

$$K_{\mathrm{R}}(r) \stackrel{\triangle}{=} \frac{1}{\lambda}\mathcal{K}(b(o,r)), \quad r \geq 0\,.$$

So $\lambda K_{\mathrm{R}}(r)$ is the mean number of points y of the process that satisfy $0 < \|\mathsf{y}-\mathsf{x}\| \leq r$ for a given point x of the process.

The K function is related to the *Fry plot*[8] that visualizes all the difference vectors $\mathsf{x}-\mathsf{y}$, $(\mathsf{x},\mathsf{y}) \in \Phi^{[2]}$. The expected number of points in the Fry plot lying within $b(o,r)$ is an unnormalized and uncorrected (for edge effects) version of $K(r)$.

3.3 Marked Point Processes

A *marked point process* is obtained if each point of a point process is equipped with a random variable or random element. For a simple point process, we may write $\hat{\Phi} = \{(\mathsf{x}_1,\mathsf{m}_1),(\mathsf{x}_2,\mathsf{m}_2),\ldots\}$, where $\Phi = \{\mathsf{x}_1,\mathsf{x}_2,\ldots\}$ is the *ground process* and $\mathsf{m}_1,\mathsf{m}_2,\ldots$ are the marks associated with each point of Φ. Often the (ground) point $\mathsf{x} \in \Phi$ is used as an index of the mark, that is, the mark associated with x is denoted as m_{x}. The marks belong to a mark space $\mathbb{M}$ with σ-algebra $\mathcal{M}$.

The marks can be used to describe properties of the point. For example, if x is the location of a base station, m_{x} may denote its transmit power, its antenna height, the number of antennas, etc. Marks can also have a more complicated structure. We may define

$$\mathsf{m}_{\mathsf{x}} = \{\mathsf{y} \in \Phi \colon \mathsf{y}-\mathsf{x}\},$$

in which case m_{x} is a point process itself, namely the translated ground process such that x is moved to the origin. This mark is sometimes referred to as the *universal mark*, since it comprises the entire point process, viewed from point x. For stationary Φ, it permits an interpretation of the Palm measure P_o as the distribution of m_{x}.

If the marks pertaining to different ground points are independent, the marked point process is called *independently marked*. For an independently marked process, the mark of a point at location x follows the *mark distribution* M_x, which is a probability measure on $(\mathbb{M},\mathcal{M})$. If, moreover, the marks do not depend on the location of the ground point, the marked point process is iid marked.

The marked point process may be viewed as a standard point process on the product space $\mathbb{R}^d \times \mathbb{M}$. While the general theory of point process of course applies to points in such more abstract spaces, it is often more convenient to separate the location of the points from their marks, that is, to define a marked point process. Specifically, stationarity is an important property for point processes on $\mathbb{R}^d$ that would be lost if translations were applied to both locations and marks.

[8] The Fry plot is implemented in the `spatstat` package in R as `fryplot`.

DEFINITION 3.3.1 (Marked point process) *A marked point process is a measurable map* $(\Omega, \mathcal{F}, \mathbb{P}) \mapsto (\hat{\mathcal{N}}, \hat{\mathfrak{N}})$*, where* $\hat{\mathcal{N}}$ *is the space of counting measures on* $\mathbb{R}^d \times \mathbb{M}$ *and* $\hat{\mathfrak{N}}$ *is the smallest* σ*-algebra that makes* $\{\hat{\Phi}(B \times L) = k\}$ *measurable for all* $B \in \mathcal{B}$*,* $L \in \mathcal{M}$*, and* $k \in \mathbb{N}$*. Its distribution is denoted by* $\hat{\mathsf{P}}$*.*

For $L \in \mathcal{M}$, the process

$$\Phi_{[L]} = \{x \in \mathbb{R}^d \colon (x, m_x) \in \hat{\Phi}, m_x \in L\}$$

is an unmarked point process consisting of the points with marks in L, with the marks removed. $\Phi_{[\mathbb{M}]}$ is the ground process. For stationary $\hat{\Phi}$, we let $\lambda_{[L]}$ denote the intensity of $\Phi_{[L]}$. Then we can write the intensity measure of $\hat{\Phi}$ as

$$\hat{\Lambda}(B \times L) \stackrel{\triangle}{=} \mathbb{E}\hat{\Phi}(B \times L) = \lambda_{[L]}|B|.$$

The quotients $\lambda_{[L]}/\lambda$ define a distribution

$$M(L) = \frac{\lambda_{[L]}}{\lambda},$$

which is the mark distribution of the typical point. The intensity measure can also be expressed as $\hat{\Lambda} = \lambda \nu_d \otimes M$, where ν_d is the Lebesgue measure in d dimensions.

Campbell's formula (see Eq. (3.9)) can be stated for marked processes in a straightforward manner as

$$\mathbb{E} \sum_{(\mathsf{x},\mathsf{m}) \in \hat{\Phi}} f(\mathsf{x}, \mathsf{m}) = \int_{\mathbb{R}^d \times \mathbb{M}} f(x, m) \hat{\Lambda}(\mathrm{d}(x, m)).$$

The intensity measure $\hat{\Lambda}(\mathrm{d}(x, m))$ disintegrates as

$$\hat{\Lambda}(\mathrm{d}(x, m)) = M_x(\mathrm{d}m)\Lambda(\mathrm{d}x)$$

given that

$$\begin{aligned}
\hat{\Lambda}(B \times L) &\stackrel{\triangle}{=} \mathbb{E}\left(\int_{\mathbb{R}^d}\int_{\mathbb{M}} \mathbf{1}(x \in B)\mathbf{1}(m \in L)\hat{\Phi}(\mathrm{d}(x, m))\right) \\
&= \mathbb{E}\left(\int_{\mathbb{R}^d} \mathbf{1}(x \in B) M_x(L)\Phi(\mathrm{d}x)\right) \\
&= \int_B M_x(L)\Lambda(\mathrm{d}x)\,.
\end{aligned}$$

3.4 The Poisson Point Process and Its Properties

The most important and widely used point process is the Poisson point process, often abbreviated to PPP or called just the Poisson process. Lacking any information on the dependence of the points, it is the "model of choice" due to its superb analytical tractability.

3.4.1 Definition

DEFINITION 3.4.1 (Poisson point process) *A point process with intensity measure Λ is a Poisson point process (PPP) if for every compact $B \in \mathcal{B}$, $\Phi(B)$ has a Poisson distribution with mean $\Lambda(B)$, that is,*

$$\mathbb{P}(\Phi(B) = k) = \mathrm{e}^{-\Lambda(B)} \frac{\Lambda(B)^k}{k!}. \tag{3.20}$$

If Λ admits a density λ, the Poisson distribution can be expressed as

$$\mathbb{P}(\Phi(B) = k) = \exp\left(-\int_B \lambda(x)\mathrm{d}x\right) \cdot \frac{\left(\int_B \lambda(x)\mathrm{d}x\right)^k}{k!}.$$

A consequence of this definition is the independence property: If $B_1, B_2, \ldots, B_n$ are disjoint compact sets, then $\Phi(B_1), \Phi(B_2), \ldots, \Phi(B_n)$ are independent (Rényi 1967). The stationary PPP where $\Lambda(B) = \lambda|B|$ is also called *uniform* or *homogeneous*. A realization of such Poisson process in presented in Figure 3.1.

The general PPP is not restricted to diffuse measures Λ. If the intensity measure Λ has a discrete component of mass λ_0 at location x_0, then $\Phi(\{x_0\})$ is Poisson distributed with mean λ_0. In this case it cannot be a simple point process because there is a positive probability that two (or more) points are located at the position of the atom. Conversely, the PPP is simple if and only if it has no fixed atoms or, equivalently, if the intensity measure is diffuse (has no discrete component).

The PPP can also be characterized in the following ways:

- By its void probabilities only. If Λ is a diffuse measure, a point process is a PPP if and only if $\mathbb{P}(\Phi(B) = 0) = \mathrm{e}^{-\Lambda(B)}$ for all compact B.
- Applying the independence property to the Palm distribution. Because conditioning on a point at x does not affect the rest of the point process, we have $\Phi \cup \{x\} \overset{\mathrm{d}}{=} (\Phi \mid x \in \Phi)$ or, in the stationary case,

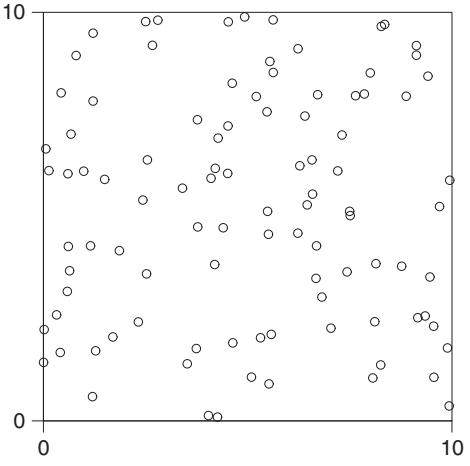

Figure 3.1 A realization of a homogeneous Poisson point process observed in a finite window.

$$\mathsf{P}_o^! \equiv \mathsf{P},$$

or, equivalently

$$\mathbb{P}(\Phi_o \in E) \equiv \mathbb{P}(\Phi + \delta_o \in E), \tag{3.21}$$

where $\Phi + \delta_o$ is the sum of Φ (viewed as a random counting measure) and the Dirac measure at o.

The second property (Eq. [3.21]) is known as *Slivnyak's theorem*. It states that for a homogeneous PPP, conditioning on a point at o is equivalent to adding a point at o.

For the PPP, higher moment measures and product densities are given in a simple form:

$$\alpha^{(n)}(B_1 \times B_2 \cdots \times B_n) = \Lambda(B_1)\Lambda(B_2) \ldots \Lambda(B_n). \tag{3.22}$$

In the uniform or homogeneous case,

$$\alpha^{(n)}(B_1 \times B_2 \times \cdots \times B_n) = \lambda^n \prod_{k\in[n]} |B_k|. \tag{3.23}$$

Accordingly, the nth moment density for a stationary PPP is simply $\varrho^{(n)} = \lambda^n$, and the pair correlation function is $g(r) \equiv 1$. This indicates that all interpoint distances $\|\mathsf{x}_k - \mathsf{x}_j\|$, $j \neq k$, are equally likely to occur. If $g(r) > 1$, this means that interpoint distances r occur more likely than in the PPP, indicating *clustering* at that distance. Conversely, if $g(r) < 1$, then interpoint distances r are less likely to occur, hence the point process exhibits repulsion at that distance.

For stationary PPPs on the plane, the reduced second factorial moment measure is given by $\mathcal{K}(B) = \lambda|B|$, and Ripley's K function (see Def. 3.2.15) is $K_{\mathrm{R}}(r) = \mathcal{K}(b(o,r))/\lambda = \pi r^2$. In fact, for all stationary point processes on the plane, $K(r) \sim \pi r^2$, $r \to \infty$, given that the expected number of points in a disk of radius r becomes independent of whether there is a point at the origin or not as $r \to \infty$.

3.4.2 Properties

The PPP has many useful properties that make it tractable.

Superposition

The superposition $\Phi_1 + \Phi_2$ (expressed as counting measure) of two point processes with intensity measures Λ_1 and Λ_2 yields, in general, a point process with intensity measure $\Lambda_1 + \Lambda_2$. If Φ_1 and Φ_2 are independent PPPs, their superposition is also a PPP.

Random independent coloring

To each point x of a point process Φ, add a randomly chosen color $c_{\mathsf{x}} \in [n]$ such that the colors (c_{x}) are iid with distribution $\mathbb{P}(c = k)$, $k \in [n]$. This yields an independently marked PPP $\hat{\Phi} = \{(\mathsf{x}, c_{\mathsf{x}})\}$, where c_{x} is the color of point x.

If this coloring is applied to a PPP Φ with intensity measure Λ and intensity function $\lambda(x)$, the point process

$$\Phi_{[k]} \stackrel{\triangle}{=} \{\mathsf{x} \in \Phi : c_{\mathsf{x}} = k\}, \quad k \in [n],$$

is a PPP with intensity function $\lambda(x)\mathbb{P}(c = k)$. Moreover, $\Phi_{[k]}$ is independent from $\Phi_{[i]}$ for $k \neq i$.

A special case of random coloring is independent thinning, where each point of a PPP is removed independently with probability $1-p$ and thus retained with probability p. The remaining points form a PPP with intensity measure $p\Lambda$ or intensity function $p\lambda(x)$. The thinning probability may be location dependent. If a point at location x is removed with probability $1 - p(x)$, the remaining points form a PPP with intensity function $p(x)\lambda(x)$.

Displacement

Next we consider the displacement of the points of a PPP by a random translation vector,

$$\Phi' = \{x \in \Phi : x + V_x\}, \tag{3.24}$$

where the random variables V_x are independent but their distribution may depend on x. When applied to a PPP, the resulting process is again a PPP whose intensity function is given in the following theorem.

THEOREM 3.4.2 (Displacement theorem) *Let Φ be a general PPP with intensity function $\lambda(x)$. If all points are independently displaced such that the new location of a point at x has probability density $\rho(x, \cdot)$, the displaced points form a PPP Φ' with intensity function*

$$\lambda'(y) = \int_{\mathbb{R}^d} \lambda(x)\rho(x, y)\mathrm{d}x. \tag{3.25}$$

In particular, if $\lambda(x) \equiv \lambda$ and $\rho(x, y)$ is a function only of $y - x$, then $\lambda'(y) \equiv \lambda$.

In this formulation, $\rho(x, \cdot)$ is the probability density of the new location $x + V_x$ in Eq. (3.24), also called the *displacement kernel*, and the second statement in the theorem asserts that if V_x does not depend on x and Φ is uniform with intensity λ, then Φ' is also uniform with the same intensity. If ρ is a function of $y - x$ only, Eq. (3.25) is a convolution integral. In the context of wireless networks, random displacements can be used to model the impact of fading or mobility.

Mapping

Point processes may be transformed by mapping each point of the process deterministically to another point, possibly in a space of different dimension. When applied to Poisson point processes, the resulting process is still Poisson in many cases.

THEOREM 3.4.3 (Mapping theorem) *Let Φ be a PPP on $\mathbb{R}^d$ with intensity Λ and intensity function λ, and let $f : \mathbb{R}^d \mapsto \mathbb{R}^s$ be a measurable function with the property that $\Lambda(f^{-1}\{y\}) = 0, \forall y \in \mathbb{R}^s$, such that f does not shrink a (nonsingleton) compact set to a singleton. Then*

$$\Phi' = f(\Phi) \stackrel{\triangle}{=} \bigcup_{\mathsf{x} \in \Phi} \{f(\mathsf{x})\}$$

is a PPP with intensity measure

$$\Lambda'(B') = \Lambda(f^{-1}(B')) = \int\limits_{f^{-1}(B')} \lambda(x)\mathrm{d}x \quad \text{for all compact } B' \subset \mathbb{R}^s.$$

In wireless networks, the mapping theorem is often applied to map the d-dimensional PPP to a one-dimensional one that represents the distances of the points from the origin. This is useful because the signal and interference powers only depend on the distances between points, not their location in $\mathbb{R}^d$.

Marking theorem

Another important result for PPPs is the marking theorem.

THEOREM 3.4.4 (Marking theorem for Poisson point processes) *Let $\hat{\Phi}$ be a marked point process on $\mathbb{R}^d \times \mathbb{M}$, and let Φ be the ground process. Then the following two statements are equivalent:*

1. *Φ is a Poisson process on $\mathbb{R}^d$ with intensity measure Λ, and, given Φ, the marks (m_x) are independent with distribution M_x on $\mathbb{M}$.*
2. *$\hat{\Phi}$ is a Poisson process on $\mathbb{R}^d \times \mathbb{M}$ with intensity measure $\hat{\Lambda} = \Lambda \otimes M_{(\cdot)}$, that is, for $A \subset \mathbb{R}^d \times \mathbb{M}$,*

$$\hat{\Lambda}(A) = \iint\limits_{(x,m)\in A} \Lambda(\mathrm{d}x) M_x(\mathrm{d}m)\,.$$

Distances

The distribution of the interpoint distances are important in the performance evaluation of wireless networks. In Poisson networks, the probability densities of the distances from a point to its nth-nearest neighbor are given in simple form. From Eq. (3.6) and Slivnyak's theorem, the nearest-neighbor distance distribution and the empty space function, which is the distance distribution of the point nearest to an arbitrary location, are identical for the PPP. Hence we can focus on a ball of radius r centered at the origin, and conditioning on having a point at o is unnecessary. The main observation is that the nth-nearest node is at a distance larger than r if there are at most $n-1$ nodes in $b(o,r)$. Denoting the distribution of the distance to the nth-nearest node by G_n, we have

$$\begin{aligned} G_n(r) &= 1 - \mathbb{P}(\Phi(b(o,r)) < n) \\ &= 1 - \exp(-\lambda c_d r^d) \sum_{k=0}^{n-1} \frac{(\lambda c_d r^d)^k}{k!}\,, \end{aligned}$$

where $c_d = |b(o,1)| = \frac{\pi^{d/2}}{\Gamma(d/2+1)}$ is the volume of the unit ball in d dimensions. When taking the derivative, all terms in the sum but the one for $n-1$ cancel out, so the probability density is the generalized gamma distribution

$$g_n(r) = \exp(-\lambda c_d r^d) \frac{d(\lambda c_d r^d)^n}{r\Gamma(n)}\,.$$

In two dimensions, it is

$$g_n(r) = \frac{2}{\Gamma(n)} (\lambda\pi)^n r^{2n-1} \exp(-\lambda\pi r^2)\,, \tag{3.26}$$

which, for $n = 1$, reduces to the Rayleigh distribution with mean $1/(2\sqrt{\lambda})$.

Due to Slivnyak's result, the contact distribution function and the nearest-neighbor distance distribution function are the same.

Since the signal-to-interference ratios are determined by relative distances,[9] it is sometimes convenient to work directly with the point process of *relative distances*, introduced in Ganti & Haenggi (2016a). For the uniform PPP, the distances relative to the nearest point are tractable also and given next.

DEFINITION 3.4.5 (Relative distance process [RDP]) *For a point process* Φ*, let* $\mathsf{x}_0 = \arg\min\{\mathsf{x} \in \Phi : \|x\|\}$ *be the point closest to the origin. The* relative distance process *is defined as*

$$\mathcal{R} \stackrel{\triangle}{=} \left\{ \mathsf{x} \in \Phi \setminus \{\mathsf{x}_0\} : \frac{\|\mathsf{x}_0\|}{\|\mathsf{x}\|} \right\}.$$

By definition $\mathcal{R} \subset (0, 1]$, and if Φ is simple, $\mathcal{R} \subset (0, 1)$. The RDP is useful in interference-limited cellular networks, where only relative distances matter for the SIR. For the PPP, we have the following result.

LEMMA 3.4.6 *If* Φ *is a homogeneous PPP, the intensity measure of the RDP is*

$$\Lambda(\mathrm{d}r) = 2r^{-3}\mathrm{d}r. \tag{3.27}$$

Proof For a PPP $\Phi = \{\mathsf{x}_k\}_{k=0}^{\infty}$ of intensity λ on the plane, assume the points are ordered according to their distance to the origin, and let $\mathsf{y}_k = \|\mathsf{x}_k\|^2$. Because $\{\mathsf{y}_k\}_{k=0}^{\infty}$ forms a PPP with constant intensity $\lambda\pi$ on $\mathbb{R}^+$, the gap $E_k \stackrel{\triangle}{=} \mathsf{y}_k - \mathsf{y}_0$ is Erlang distributed with parameter $\lambda\pi k$. The CDF of the relative distance $\|\mathsf{x}_0\|/\|\mathsf{x}_k\|$ is, for $k > 0$ and $r \in [0, 1]$,

$$\begin{aligned}
\mathbb{P}(\|\mathsf{x}_0\| \le r\|\mathsf{x}_k\|) &= \mathbb{P}\left(\frac{\mathsf{y}_0}{\mathsf{y}_0 + E_k} \le r^2\right) \\
&= \mathbb{P}\left(\mathsf{y}_0 \le \frac{r^2}{1 - r^2} E_k\right) \\
&= \mathbb{E}\left(1 - \exp\left(-\lambda \frac{r^2}{1 - r^2} E_k\right)\right) \\
&= 1 - (1 - r^2)^k.
\end{aligned} \tag{3.28}$$

The corresponding probability density is $2kr(1 - r^2)^{k-1}$, and the intensity function follows as

$$\lambda(r) = \sum_{k=1}^{\infty} 2kr(1 - r^2)^{k-1} = 2r^{-3}.$$

The CDF Eq. (3.28) can be obtained more directly by noting that given y_k, the points y_0 through y_{k-1} correspond to the order statistic of a uniform distribution on $[0, \mathsf{y}_k]$ and thus $\mathbb{P}(\mathsf{y}_0/\mathsf{y}_k > x) = (1 - x)^k$ because for $\mathsf{y}_0 > x\mathsf{y}_k$ to happen, k iid uniform random variables on $[0, 1]$ need to exceed x. □

[9] This holds for the standard path loss law $\ell(r) = r^{-\alpha}$.

The RDP of a stationary point process is not locally finite on intervals $[0, \epsilon)$, $\epsilon > 0$.

Alternatively, the *squared (inverted) RDP* may be defined (Ganti & Haenggi 2016b), which is

$$\mathcal{S} \triangleq \mathcal{R}^{-2} = \left\{ \mathsf{x} \in \Phi \setminus \{\mathsf{x}_0\} \colon \frac{\|\mathsf{x}\|^2}{\|\mathsf{x}_0\|^2} \right\} \subset [1, \infty).$$

This point process has constant intensity 1 on $[1, \infty)$ if Φ is a stationary PPP. Its pgfl is given by the simple expression

$$G_{\mathcal{S}}[v] = \frac{1}{1 + \int_1^\infty (1 - v(r))\mathrm{d}r} \tag{3.29}$$

whenever the integral in the denominator converges.

3.4.3 The pgfl and the Campbell-Mecke theorem

Another significant advantage of the PPP is that its pgfl has a simple closed-form expression. It is given by

$$G[v] = \exp\left(-\int_{\mathbb{R}^d} (1 - v(x))\Lambda(\mathrm{d}x) \right). \tag{3.30}$$

For the PPP, the Campbell-Mecke theorem can be written in the following form.

COROLLARY 3.4.7 (Mecke's formula) *For a PPP with intensity measure Λ and a measurable function $g \colon \mathbb{R}^d \times \mathcal{N} \mapsto \mathbb{R}^+$,*

$$\mathbb{E}\left(\sum_{\mathsf{x} \in \Phi} g(\mathsf{x}, \Phi) \right) = \int_{\mathbb{R}^d} \mathbb{E}(g(x, \Phi \cup \{x\}))\Lambda(\mathrm{d}x) \tag{3.31}$$

and

$$\mathbb{E}\left(\sum_{\mathsf{x} \in \Phi} g(\mathsf{x}, \Phi \setminus \{\mathsf{x}\}) \right) = \int_{\mathbb{R}^d} \mathbb{E}(g(x, \Phi))\Lambda(\mathrm{d}x). \tag{3.32}$$

The first form is obtained from Eq. (3.10) because conditioning on a point at x is the same as adding that point, and the second (reduced) form follows from Eq. (3.11) because $\mathsf{P}_x^! \equiv \mathsf{P}$.

Applying Eq. (3.31) to $g(x, \varphi) = f(x) \prod_{y \in \varphi} h(y)$, we obtain

$$\begin{aligned} \mathbb{E}\left(\sum_{\mathsf{x} \in \Phi} f(\mathsf{x}) \prod_{\mathsf{y} \in \Phi} h(\mathsf{y}) \right) &= \int_{\mathbb{R}^d} \mathbb{E}\left(f(x) \prod_{\mathsf{y} \in \Phi \cup \{x\}} h(\mathsf{y}) \right) \Lambda(\mathrm{d}x) \\ &= \int_{\mathbb{R}^d} \mathbb{E}\left(f(x)h(x) \prod_{\mathsf{y} \in \Phi} h(\mathsf{y}) \right) \Lambda(\mathrm{d}x) \end{aligned}$$

$$= \mathbb{E}\left(\prod_{\mathsf{y}\in\Phi} h(\mathsf{y})\right) \int_{\mathbb{R}^d} f(x)h(x)\Lambda(\mathrm{d}x)$$

$$= \exp\left(-\int_{\mathbb{R}^d} (1-h(x))\Lambda(\mathrm{d}x)\right) \int_{\mathbb{R}^d} f(x)h(x)\Lambda(\mathrm{d}x). \quad (3.33)$$

This is a useful expression to calculate the expectation of interference functionals of the form $\mathbb{E}(Ie^I)$, where I is the interference power. Such functionals appear in the analysis of networks with Nakagami fading (Schilcher et al. 2016).

3.5 Alternative Models

In the PPP, there is no interaction between the points. Although such a model is a good starting point in many cases, a refined model is preferred for those cases where nodes exhibit clustering or repulsion. For example, the base stations in a coverage-oriented deployment usually exhibit repulsion, that is, they form a more regular point process than the PPP. Repulsion or regularity is reflected in the second-order statistics such as the pair correlation function. If $g(r) < 1$, at least for $r < r_0$ for some r_0, we speak of a *soft-core process*, whereas if $g(r) = 0$ for $r < r_0$, we speak of a *hard-core process* with hard-core distance r_0. In a hard-core process, it is not possible that two points are within distance smaller than r_0.

3.5.1 Determinantal point processes

The product densities of determinantal point process are given by the determinant of a matrix that is defined using a kernel function. We focus on the two-dimensional case and view $\mathbb{R}^2$ as isomorphic to the complex plane $\mathbb{C}$

DEFINITION 3.5.1 (Determinantal point process) *If the product densities of a point process Φ are given by*

$$\varrho^{(n)}(x_1,\ldots,x_n) = \det A, \quad x_1,\ldots,x_n \in \mathbb{C}, \quad (3.34)$$

where $A_{ij} = K(x_i,x_j)$, $i,j \in [n]$, for a kernel $K\colon \mathbb{C}^2 \mapsto \mathbb{C}$, then Φ is a determinantal point process.

Determinantal point processes are repulsive because the determinant approaches 0 when two rows or columns in A become close to being linearly dependent. If $x_i = x_j$ for some $i \neq j$, $\varrho^{(n)} = 0$, indicating that it is not possible for two points to be colocated.

An important special case is the Ginibre point process, often abbreivated as GPP, where the kernel is given by

$$K(x,y) = \pi^{-1}\mathrm{e}^{-(|x|^2+|y|^2)/2}\mathrm{e}^{x\bar{y}}.$$

The intensity of the Ginibre process is $\lambda = \pi^{-1}$, and the second moment density is

$$\varrho^{(2)}(x,y) = \pi^{-2}(1-\mathrm{e}^{-|x-y|^2}),$$

which only depends on $|x - y|$ because the Ginibre process is motion-invariant. Hence we can write $\varrho^{(2)}(x, y) = \varrho^{(2)}_{\mathrm{mi}}(|x - y|)$. The pair correlation function follows as

$$g(r) \stackrel{\triangle}{=} \frac{\varrho^{(2)}_{\mathrm{mi}}(r)}{\pi^{-2}} = 1 - \mathrm{e}^{-r^2},$$

which is small for $r \ll 1$, as expected. Because $g(r) < 1$ for all r, the Ginibre process is repulsive at all distances. The Ginibre process has the remarkable property that

$$\{\|\mathsf{x}_1\|^2, \|\mathsf{x}_2\|^2, \ldots\} \stackrel{\mathrm{d}}{=} \{Q_1, Q_2, \ldots\},$$

where Q_k are independent gamma-distributed random variables with probability density

$$f_{Q_k}(x) = \frac{x^{k-1}\mathrm{e}^{-x}}{\Gamma(k)}.$$

The mean of these random variables equals their index, $\mathbb{E}Q_k = k$.

In Definition 3.5.1, replacing the determinant with a permanent, where all the minus signs in the determinant become positive signs, gives the definition of a *permanental process*. Permanental processes exhibit clustering and form a subclass of Cox processes. Historically, they were not considered as tractable as determinantal processes because a permanent, unlike a determinant, lacks an intuitive geometric interpretation. Recent results, however, based on more mathematical techniques from functional analysis, have resulted in a uniform approach, treating determinantal and permanental processes together (Shirai & Takahashi 2003).

Determinantal and permanental point processes are also called fermion and boson point processes, respectively, due to their origins in physics. Poisson processes are determinantal *and* permanental. But the Poisson process is often excluded from this family of processes, as key results of determinantal processes need a particular condition on the kernel, which the kernels of Poisson processes lack (Lavancier, Møller, & Rubak 2015).

3.5.2 Matérn hard-core processes

Two such models are the Matérn hard-core processes types I and II, introduced in Matérn (1986).

DEFINITION 3.5.2 *Matérn hard-core process of type I: Starting with a uniform PPP Φ_{b} with intensity λ_{b}, first flag for removal all points that have a neighbor within distance r. Then remove all flagged points.*

DEFINITION 3.5.3 (Matérn hard-core process of type II) *Starting with a uniform PPP Φ_{b} with intensity λ_{b}, add to each point x an independent random variable $m(x)$, called a mark, uniformly distributed on* $[0, 1]$. *Flag for removal all points that have a neighbor within distance r that has a smaller mark. Then remove all flagged points. Formally,*

$$\Phi \stackrel{\triangle}{=} \{x \in \Phi_{\mathrm{b}} \colon m(x) < m(y) \ \text{ for all } y \in \Phi_{\mathrm{b}} \cap b(x, r) \setminus \{x\}\}.$$

In both types of hard-core processes, all points are removed simultaneously, so even points that are thinned out can eliminate other points. A denser packing could be achieved for both types if points were removed one by one and the conditions rechecked

after each removal. Such a process is sometimes referred to as a Matérn process of type III, but it is significantly less tractable than the processes of type I and II.

The intensity of the type I process is $\lambda = \lambda_b e^{-\lambda_b c_d r^d}$. Given r (and d), it is maximized at $\lambda_b = 1/(c_d r^d)$, so the highest density that can be achieved is $\lambda_{\max} = (e c_d r^d)^{-1}$.

To determine the intensity of the type II process, we first condition on a point having a given mark t. This point is retained with probability $e^{-t\lambda_b c_d r^d}$, because $t\lambda_b$ is the density of points with marks smaller than t. Unconditioning on t, we obtain

$$\lambda = \lambda_b \int_0^1 e^{-t\lambda_b c_d r^d} dt = \frac{1 - e^{-\lambda_b c_d r^d}}{c_d r^d}.$$

In this case, the maximum density is achieved as $\lambda_b \to \infty$, yielding $\lambda_{\max} = (c_d r^d)^{-1}$, which is a factor e higher than for the type I process. For $d = 2$ and $r = 1$, the maximum density is $1/\pi \approx 0.318$. In comparison, the densest deterministic arrangement, the triangular lattice, achieves a density of $2/\sqrt{3} \approx 1.15$. Examples of type I and type II processes are shown in Fig. 3.2. Each realization in the figure has a number of points in $[0, 10]^2$ that is slightly above its mean.

3.5.3 Strauss processes

Strauss processes belong to the family of Gibbs point processes, which are defined using a density function with respect to the distribution of a PPP. We first give a definition of Gibbs processes.

A function $f : \mathcal{N} \mapsto [0, 1]$ is called *hereditary* if $f(\varphi) > 0$ implies that $f(\varphi') > 0$ whenever $\varphi' \subseteq \varphi$. So hereditary means that if a realization φ occurs with positive probability, so do all realizations with one or more points removed. Conversely, if $f(\varphi) = 0$, then all supersets of φ must also have zero probability.

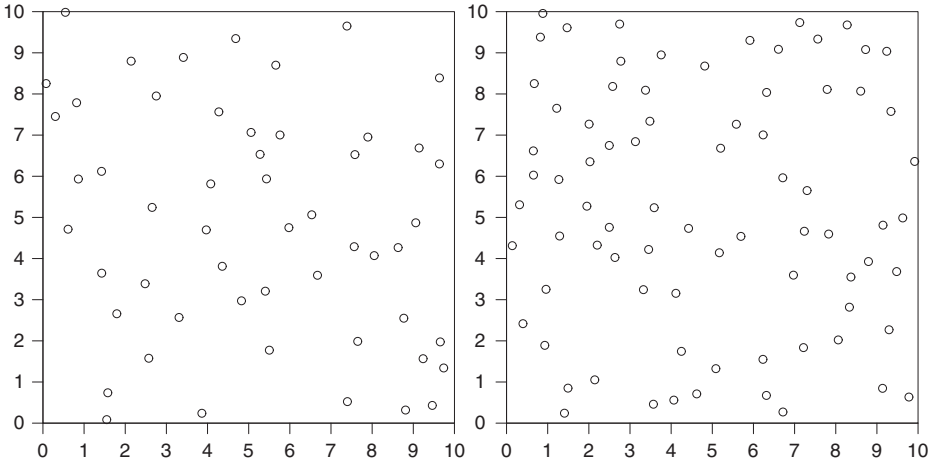

Figure 3.2 Matérn hard-core processes with $\lambda_b = 1$ and $r = 1/2$. (*Left*) Realization of type I process with 53 points on $[0, 10]^2$. The density of this process is $\lambda = e^{-\pi/4} \approx 0.456$. (*Right*) Realization of type II process with 78 points in $[0, 10]^2$. The density is $\lambda = (1 - e^{-\pi/4})/(\pi/4) \approx 0.693$.

DEFINITION 3.5.4 (Gibbs point process) *Let* Q *be the distribution of a Poisson process with intensity* Λ *and* $f\colon \mathcal{N} \mapsto \mathbb{R}^+$ *be a hereditary function such that*

$$\int_{\mathcal{N}} f(\varphi)\mathsf{Q}(\mathrm{d}\varphi) = 1\,. \tag{3.35}$$

Then

$$\mathsf{P}(E) = \int_{E} f(\varphi)\mathsf{Q}(\mathrm{d}\varphi)\,, \quad \forall E \in \mathfrak{N}\,, \tag{3.36}$$

is a probability measure on $(\mathcal{N}, \mathfrak{N})$ *and thus the distribution of a point process. This point process is called a* Gibbs point process.

Expression (3.36) shows that P is absolutely continuous with respect to Q, so, formally, f is the Radon-Nikodým derivative $f = \mathrm{d}\mathsf{P}/\mathrm{d}\mathsf{Q}$. So f is the density of P with respect to the probability measure of the PPP.

The density f is used to make certain classes of realizations of the PPP more likely and others less likely. In the Strauss process, realizations where points have many nearby neighbors are "discouraged."

DEFINITION 3.5.5 (Strauss process) *Let* $a > 0$, $R > 0$, *and* $b \in \mathbb{R}^+ \cup \{\infty\}$. *A Gibbs process is called Strauss process on* W *if the density* f *has the form*

$$f(\varphi) = ca^{\varphi(W)} \exp(-bt_R(\varphi))\,, \quad \forall \varphi \in \mathcal{N}\,,$$

where

$$t_R(\varphi) = \frac{1}{2} \sum_{x,y\in\varphi}^{\neq} \mathbf{1}_{W\times W}(x,y)\mathbf{1}_{(0,R)}(\|x-y\|)$$

denotes the number of point pairs $\{x, y\}$ *of* φ *whose distance is smaller than* R. c *is a normalizing constant so that Eq. (3.35) holds, that is,*

$$c^{-1} = \int_{\mathcal{N}} a^{\varphi(W)} \exp(-bt_R(\varphi))\mathsf{Q}(\mathrm{d}\varphi)\,.$$

The parameters a, b, and R are called *activity*, *interaction parameter*, and *interaction radius*, respectively. The interaction parameter b determines the strength of repulsion between the points. If $b = \infty$, in particular,

$$f(\varphi) = \begin{cases} 0 & \text{if } t_R(\varphi) > 0 \\ ca^{\varphi(W)} & \text{if } t_R(\varphi) = 0\,. \end{cases} \tag{3.37}$$

This likelihood function describes a hard-core process with hard-core radius R, given that f assigns probability 0 to all point sets that have two points closer than R. This process is often called a *Poisson hard-core process*. At the other extreme, if $b = 0$, then $f(\varphi) = ca^{\varphi(W)}$, and it is easily verified that the resulting Strauss process is a uniform (on W) PPP of intensity a. So in this case there is no repulsion at all between points. Figure 3.3 shows two Strauss processes with $a = 2$ for $b = 1$ and $b = 5$. For $b = 0$, the process would be a PPP of intensity 2 and contain 200 points on the $[0, 10]^2$ square on average.

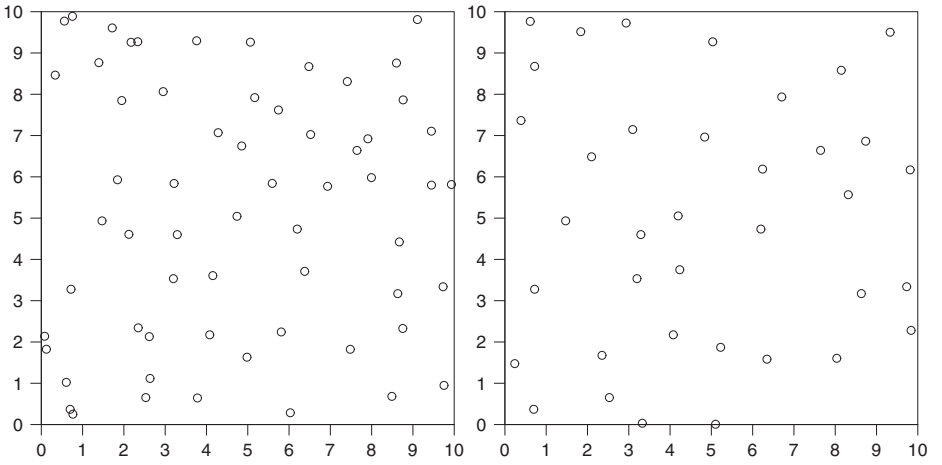

Figure 3.3 Realizations of Strauss processes with $a = 2$ and $R = 1$. (*Left*) $b = 1$. This realization has 61 points. (*Right*) $b = 5$. In this case, patterns with point pairs within distance 1 are penalized more severely; as a result, this realization has only 37 points.

3.5.4 Shot noise Cox processes

A Cox process is a doubly random Poisson process. It is defined using a random intensity measure $\mathcal{Z}$ or a random intensity function (random field) Z. Conditioned on $\mathcal{Z}$ or Z, the Cox process is a Poisson process with the given intensity measure or function. This implies that the intensity function of the Cox process is $\lambda(x) = \mathbb{E}Z(x)$. It is understood that $\mathcal{Z}(B) = \int_B Z(x)\mathrm{d}x < \infty$ almost surely for all bounded Borel sets $B \subset \mathbb{R}^d$.

An important subclass are *shot noise Cox processes* (Møller 2003) defined as follows.

DEFINITION 3.5.6 (Shot noise Cox process) *A Cox process is a* shot noise Cox process *if its intensity function* $Z\colon \mathbb{R}^d \to \mathbb{R}^+$ *is of the form*

$$Z(x) = \sum_{i\in\mathbb{N}} \gamma_i \kappa(\xi_i, x),$$

where the kernel $\kappa(\xi_j, \cdot)$ *is a probability density on* $\mathbb{R}^d$, $\Xi = \{\xi_i\}$ *is a PPP on* $\mathbb{R}^d$, *and* $\Psi = \{\gamma_i\}$ *is a PPP on* $\mathbb{R}^+$.

It follows that the intensity function of the shot noise Cox process is given by

$$\lambda(x) = \int_{\mathbb{R}^+}\int_{\mathbb{R}^d} r\kappa(y, x)\Lambda_\Xi(\mathrm{d}y)\Lambda_\Psi(\mathrm{d}r),$$

where Λ_ξ and Λ_γ are the intensity measures of the PPPs $\{\xi_i\}$ and $\{\gamma_i\}$, respectively.

If κ is translation-invariant, that is, $\kappa(y, x) = \kappa(y - x)$ (in a slight abuse of notation), and $\Xi = \{\xi_i\}$ is a stationary PPP, then the shot noise Cox process is stationary also. If, moreover, $\kappa(x) = \kappa(\|x\|)$, the Cox process is motion-invariant.

It is often useful to interpret the shot noise Cox process Φ as a cluster process

$$\Phi = \bigcup_{i \in \mathbb{N}} \Phi_i$$

of independent PPPs Φ_i with intensity function $\gamma_i \kappa(\xi_i, \cdot)$. The PPP Ξ is referred to as the *process of cluster centers* or *parent process*, and the clusters Φ_i as the *daughter points* or *offspring points*. γ_i denotes the mean number of points in cluster i. If the parent process has intensity λ_p and $\gamma_i = \bar{c}$ for all $i \in \mathbb{N}$, we obtain a Neyman-Scott process.

3.5.5 The Poisson hole process

DEFINITION 3.5.7 (Poisson hole process) *Let Φ_1, Φ_2 be independent uniform PPPs of intensities λ_1 and λ_2, respectively. Further let*

$$\Xi_r \stackrel{\triangle}{=} \bigcup \{x \in \Phi_1 : b(x, r)\}$$

be the union of all disks of radius r centered at a point of Φ_1. The Poisson hole process is defined as

$$\Phi \stackrel{\triangle}{=} \Phi_2 \setminus \Xi_r \,.$$

In the Poisson hole process, each point in Φ_1 carves out a hole of radius r from Φ_2. Hence it is a Cox process with driving measure

$$\mathcal{Z} = \lambda_2 \nu_d(\cdot \setminus \Xi_r) \,.$$

The union of disks centered at the points of a PPP Ξ_r is called a *Boolean model*.

This Cox process has applications in cognitive, cellular, and device-to-device (D2D) networks (Deng, Zhou, & Haenggi 2015b, Dhillon et al. 2016; Lee & Haenggi 2012). The process Φ can be interpreted as the secondary or D2D users that are allowed to transmit because they are at least at distance r from all primary users and therefore do not cause harmful interference.

Because the probability of a point being retained is the probability that it has no point of the PPP of intensity λ_1 within distance r, the intensity of the Poisson hole process is

$$\lambda = \lambda_2 \exp(-\lambda_1 \pi r^2) \,.$$

4 Statistics of Received Power at the Typical Location

In this chapter we use the Palm calculus results from Chapter 3 to examine a network from the perspective of a single user. This is achieved by introducing a general framework with path loss model and random propagation effects, such as random fading and shadowing. The rest of the chapter is focused on showing how large random propagation effects can make a non-Poisson network appear Poisson to a single user, due to results from probability theory.

4.1 Modeling Signal Propagation and Cells in Heterogeneus Networks

In Sections 3.4 and 3.5 we discussed several stationary point process models allowing us to represent base station locations. We develop this approach further by introducing random marks as models for *random propagation effects* of signals emitted by these stations, as discussed in Section 2.2. The random marking of base stations also lets us distinguish different *base station types* in the network, such as those with different transmission powers. We call such network models, with different station types, *heterogeneous network models*. This concept of heterogeneity does not rule out stationarity of the model, where distribution of the model is invariant under translation. For stationary, heterogeneous network models, we introduce concepts such as the typical network station of a given type and the typical location in the network by using tools from Palm calculus. A key result of this section is a general formula for relating characteristics of the typical location and the typical base station, which we then use to derive a number of useful results.

4.1.1 Stationary heterogeneous network with a propagation field

We consider a general stationary point process Φ as locations of base stations in the plane $\mathbb{R}^2$, where we will work, but all the results generalize to the $\mathbb{R}^d$ case. Recall stationarity implies that the average number of base stations in a given region B is equal to $\mathbb{E}[\Phi(B)] = \lambda|B|$, where $|B|$ is the Lebesgue measure, which is usually the area of $B \in \mathbb{R}^2$, and λ is assumed always to be $0 < \lambda < \infty$, is the intensity of the point process Φ or, in other words, the average *density of base stations*.

Signal propagation model

DEFINITION 4.1.1 (Path loss function) *A* path loss function $\ell(y)$ *is a general nonnegative function mapping from the plane* $\mathbb{R}^2$ *to the positive real line* $\mathbb{R}^+$.

DEFINITION 4.1.2 (Random propagation field) *For given base station locations* Φ *and a location* $y \in \mathbb{R}^2$, *a* random propagation field $\{H_{\mathsf{x}}(y)\}_{\mathsf{x}\in\Phi}$ *is a nonnegative random field.*

In our setting a random propagation field is simply a collection of random variables indexed by the space $\mathbb{R}^2$, on which the point process Φ is defined. For a given Φ, we further assume that the random fields $H_{\mathsf{x}}(\cdot)$ are independent across all $\mathsf{x} \in \Phi$ and are identically distributed. This is a great simplifying assumption, used almost universally. But as a side remark, we do not need to assume any particular distribution for $H_{\mathsf{x}}(\cdot)$, or even independence or identical distributions for $H_{\mathsf{x}}(y)$ across y. The results still hold if the assumptions of independence and identical distributions across $\mathsf{x} \in \Phi$ are removed, because stationarity of the marked point process is all that is needed. In addition to random fading and shadowing, $\{H_{\mathsf{x}}(y)\}_{\mathsf{x}\in\Phi}$ can be used to model other random factors such as randomly varying powers of the transmitting stations or antenna gains.

DEFINITION 4.1.3 (Propagation field) *For given base station locations* Φ *and a location* $y \in \mathbb{R}^2$, *the* propagation field *is the random field defined by*

$$\mathsf{L}_{\mathsf{x}}(y) = \frac{1}{\ell(y-\mathsf{x})\, H_{\mathsf{x}}(y-\mathsf{x})}, \tag{4.1}$$

which represents the strengths of signals between a given base station $\mathsf{x} \in \Phi$ *and all locations* $y \in \mathbb{R}^2$.

The propagation field $\{\mathsf{L}_{\mathsf{x}}(y)\}_{\mathsf{x}\in\Phi}$ is a mathematical description of the network under the standard statistical or stochastic propagation model.

Example 4.1 A very common simplifying assumption for the path loss function is (see Eq. [2.7]) the singular path loss function

$$\ell(x) \stackrel{\triangle}{=} (\kappa\|x\|)^{-\alpha}, \tag{4.2}$$

with path loss constant $\kappa > 0$ and path loss exponent $\alpha > 0$. For base stations in the plane $\mathbb{R}^2$ (or in $\mathbb{R}^d$), one often assumes $\alpha > 2$ (or $\alpha > d$) to ensure finiteness of the sum of all the signal powers received at a given location, but this assumption is not necessary to study the signal powers separately.

We look at some common assumptions for the marginal distributions of random propagation effects $H = H_{\mathsf{x}}(y)$.

Example 4.2 Under the log-normal shadowing model, H is a log-normal variable (see Section C.2.10), meaning

$$H = \mathrm{e}^{\mu+\sigma Z}, \tag{4.3}$$

where $\mu \in \mathbb{R}$, $\sigma > 0$, and Z is the standard Gaussian or normal random variable (see Section C.2.2) with zero mean and unit variance.

Example 4.3 Under the Rayleigh fading model (see Section C.2.4), H is an exponential random variable (see Section C.2.5) with mean $\mu > 0$.

Example 4.4 Under the Nakagami (k, μ) fading model (see Section C.2.9), H is a gamma variable, so it has the gamma distribution Gamma(k, μ) (see Section C.2.7) with mean $k\mu > 0$, where $\mu > 0$ and $k > 0$, which generalizes the Rayleigh fading model.

Station types in a heterogeneous network

Although we are using a stationary point process Φ, we can still model a heterogeneous network by assigning random, independent marks to each base station, where each mark represents the base station type with, for example, a different transmission power. For each $\mathsf{x} \in \Phi$, we let T_{x} be some parameter or a set of parameters of the station x with values in some state space $\mathcal{T}$. We will call T_{x} the *type* of base station x. We assume that T_{x} does not depend on the location of base station x but in general it may depend on the propagation effects H_{x} experienced by a signal coming from base station x. We often drop the subscript of marks when referring to a generic mark such as H or T. In summary, our heterogeneous network model with the propagation field is represented by a stationary, independently marked point process

$$\tilde{\Phi} \stackrel{\triangle}{=} \{\mathsf{x}, (H_{\mathsf{x}}(\cdot), T_{\mathsf{x}}) : \mathsf{x} \in \Phi\}. \tag{4.4}$$

We stress again that the concept of network heterogeneity here is not a removal of network stationarity.

4.1.2 Typical network station and typical location in the network

Two important notions in the analysis of stationary network models are the *typical network (base) station* and *typical location*.

Typical station in the network

When we talk about a typical station, we mean a mathematical formalization of a base station statistically representative in some way of all base stations. To put this on a solid theoretical footing, we use Palm theory, where the typical point is x_o located at the origin o under the Palm probability or Palm measure (see Section 3.2.4) $\mathbb{P}_o$ related to the stationary distribution $\mathbb{P}$ of a given process Φ. In the case of an iid-marked Poisson process, marks of all points under $\mathbb{P}_o$ are iid with the same distribution as under $\mathbb{P}$, and

the extra point $\mathsf{x}_o = o$ gets an independent copy $(H_o(\cdot), T_o)$ of the mark, with the same distribution as the marks of the other points.

Example 4.5 By Slivnyak's theorem (Eq. 3.21), the Palm distribution of the Poisson process corresponds to the homogeneous (stationary) Poisson process, with an additional point $\mathsf{x}_o = o$ placed at the origin. In addition to the Poisson process, the Palm distribution is known or calculable for some other point processes, such as Cox and determinantal processes, but Slivnyak's theorem definitely makes the Poisson case the most tractable.

Typical location in the network

The typical location in a network seeks to represent an arbitrary location in the plane, chosen without any relation to the marked point process $\tilde{\Phi}$. We can think of it as the position of a potential user of the network. When the model of network base stations is stationary, as $\tilde{\Phi}$ under $\mathbb{P}$, this location can be set to the origin without loss of generality. In a way, the typical location serves as an ersatz or substitution for the typical user, which can also be formalized with Palm theory, analogously to the typical station, by introducing a point process of users. We adopt a simpler approach, which does not require considering the second point process of users.

4.1.3 Exchange formula

To introduce a relation between the typical location and the typical base station, we consider a real valued function $\chi(\mathsf{x}, y)$ defined for $\mathsf{x} \in \Phi$ and $y \in \mathbb{R}^2$, which we interpret as the amount of some virtual service that the station located at x is achieving for the location $y \in \mathbb{R}^2$. This value may depend not only on the two respective locations, x of the serving station and y of the served location, but also on the mark $(H_{\mathsf{x}}(\cdot), T_{\mathsf{x}})$ of the serving station x, as well as locations and marks of all other stations in $\tilde{\Phi}$. In other words, for any $\mathsf{x} \in \Phi$ the quantity

$$\chi(\mathsf{x}, y) = \chi(\mathsf{x}, y; \tilde{\Phi}) \tag{4.5}$$

is the amount of the service that the station x is performing for the location $y \in \mathbb{R}^2$. We further assume that the function χ is translation invariant, so its value does not depend on the way we choose the origin in the plane, or more formally,

$$\chi(\mathsf{x}, y; \tilde{\Phi}) = \chi(o, y - \mathsf{x}; \tilde{\Phi} - \mathsf{x}), \tag{4.6}$$

where $\tilde{\Phi} - x$ denotes the translation of all points of Φ by x while still preserving their marks.

THEOREM 4.1.4 (Typical location-station exchange formula) *The total amount of service received at the typical location is equal to the total service supplied by the typical station multiplied by the network density*

$$\mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} \chi(\mathsf{x}, o; \tilde{\Phi})\right] = \lambda \int_{\mathbb{R}^2} \mathbb{E}_o[\chi(o, y; \tilde{\Phi})]\, \mathrm{d}y\,. \tag{4.7}$$

The left-hand side of Eq. (4.7) expresses the typical observer's point of view at the origin, whereas on the right-hand side there is Palm expectation $\mathbb{E}_o$ under which there is the typical station x located at the origin o, meaning $\mathsf{x} = o \in \Phi$.

Proof of Theorem 4.1.4 The result is an immediate consequence of the Campbell-Mecke theorem for stationary processes (Theorem 3.10). By the translation invariance of χ, we have

$$\mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} \chi(\mathsf{x}, o; \tilde{\Phi})\right] = \mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} \chi(o, -\mathsf{x}; \tilde{\Phi} - \mathsf{x})\right]$$

$$= \lambda \left[\int_{\mathbb{R}^2} \mathbb{E}_o[\chi(o, -x; \tilde{\Phi})\right] \mathrm{d}x\,,$$

and a simple substitution $y \stackrel{\triangle}{=} -x$ leads to the right-hand side of Eq. (4.7). □

Remark 4.1 The proof does not require any specific form on the sequence of marks. It is a spatial version of Little's law, a fundamental result in queuing theory. The proof of Theorem 4.1.4 can be further extended to account for a point process modeling user locations which is jointly stationary with the network model, and not just one observer at the origin. In the latter case, we speak about the typical user-station exchange formula. [1]

4.1.4 Shot noise model of all signal powers in the network

To illustrate an application of the exchange formula, we consider the *total received power* or *total interference* in a network at the typical location. For given base station locations Φ and the propagation field in Eq. (4.1), the total received power of the network is the random variable

$$I'(o) = \sum_{\mathsf{x}\in\Phi} (\mathsf{L}_{\mathsf{x}}(o))^{-1}\,. \tag{4.8}$$

A quantity such as $I'(o)$ that can be expressed as a sum of values of some (possibly independently randomized) function of the points of a point process is known as *shot noise*, which explains the name of the shot noise Cox process, introduced in Section 3.5.4.

Perhaps the simplest illustration of the exchange formula is an expression for the mean total received power. The exchange formula in Eq. (4.7) gives

$$\mathbb{E}[I'(o)] \stackrel{\triangle}{=} \mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} (\mathsf{L}_{\mathsf{x}}(o))^{-1}\right] \tag{4.9}$$

$$= \lambda \int_{\mathbb{R}^2} \mathbb{E}_o[H_o(y)]\ell(y)\,\mathrm{d}y$$

[1] This result can be compared to *Neveu's exchange formula* (Baccelli & Błaszczyszyn 2009a, Theorem 4.5) or, even more generally, the *mass-transport principle* (Gentner & Last 2011).

$$= \lambda \int_{\mathbb{R}^2} \mathbb{E}[H(y)]\ell(y)\,\mathrm{d}y\,, \tag{4.10}$$

where we have replaced the mean propagation field $\mathbb{E}_o[H_o(y)]$ from station o by the generic mean propagation field $\mathbb{E}[H(y)]$ due to the iid assumption on these fields. Expression (4.10) further simplifies in case of a propagation field identically distributed across $y \in \mathbb{R}^2$:

$$\mathbb{E}[I'(o)] = \lambda \mathbb{E}[H] \int_{\mathbb{R}^2} \ell(y)\,\mathrm{d}y\,. \tag{4.11}$$

We could have of course also directly used Campbell's formula for sums, Eq. (3.9), to derive the expression for $\mathbb{E}[I'(o)]$.

Example 4.6 For the simple path loss in Eq. (4.2) the integral $\int_{\mathbb{R}^2} \ell(y)\,\mathrm{d}y$ diverges because of the pole or singularity at the origin. This can be removed by assuming some modification of the path loss model at the origin, such as $\ell(x) = (\kappa \max(\mathrm{dist}_0, \|x\|))^{-\alpha}$, for some constant $\mathrm{dist}_0 > 0$. The mean received power $\mathbb{E}[I'(o)]$, however, then depends heavily on the value of the constant $\mathrm{dist}_0 > 0$. In general, all reasonable network characteristics should be insensitive to the modeling of the path loss at the origin. The expected value of the total received power does not satisfy this requirement. Results for more general path loss functions are presented in Chapter 7, Section 7.1.2.

Despite the problem with the mean received power, we can consider other distributional characteristics of I', for example, its Laplace transform $\mathcal{L}_{I'(o)}(s) \stackrel{\triangle}{=} \mathbb{E}[\mathrm{e}^{-sI'(o)}]$, which has an explicit expression under the Poisson network model:

$$\mathbb{E}[\mathcal{L}_{I'(o)}(s)] = \exp\left[-\lambda \int_{\mathbb{R}^2} \Big(1 - \mathcal{L}_{H_x(x)}(s\ell(x))\Big)\,\mathrm{d}x\right], \tag{4.12}$$

where $\mathcal{L}_{H_x(y)}(s) = \mathbb{E}[\mathrm{e}^{-sH_x(y)}]$ is the Laplace transform of the propagation field $H_x(y)$. This expression simplifies in case of a propagation field identically distributed across $y \in \mathbb{R}^2$ to give

$$\mathbb{E}[\mathcal{L}_{I'(o)}(s)] = \exp\left[-\lambda \int_{\mathbb{R}^2} \Big(1 - \mathcal{L}_{H}(s\ell(x))\Big)\,\mathrm{d}x\right], \tag{4.13}$$

and in case of the simple path loss function in Eq. (4.2) it simplifies further to

$$\mathbb{E}[\mathcal{L}_{I'(o)}(s)] = \exp\left[-\lambda\kappa^{-2}s^{\delta}D(\delta)\,\mathbb{E}[H^{\delta}]\right], \tag{4.14}$$

where $\delta \stackrel{\triangle}{=} 2/\beta$ and the $D(\delta) \stackrel{\triangle}{=} \pi\Gamma(1-\delta)$.

Remark 4.2 We refer to $\delta = 2/\beta$ and the $D(\delta) = \pi\Gamma(1-\delta)$ as the *characteristic exponent* and the *spatial contention factor*, respectively. More generally, the characteristic exponent is d/β for $\mathbb{R}^d$, and it and the contention factor usually appear whenever the power-law path loss model $\ell(x) \stackrel{\triangle}{=} (\kappa\|x\|)^{-\alpha}$ is used, simplifying notation.

Recalling the fading and shadowing models introduced in Examples 4.2–4.4, we give some examples of the key moment $\mathbb{E}(H^\delta)$.

Example 4.7 In the no-fading scenario $H \equiv 1$, the key moment is simply

$$\mathbb{E}[H^\delta] = 1\,, \tag{4.15}$$

which could also be replaced with some positive constant to give the constant fading scenario.

Example 4.8 For the log-normal shadowing model (Example 4.2),

$$\mathbb{E}(H^\delta) = \mathrm{e}^{\delta(\mu+\delta\sigma^2/2)}\,. \tag{4.16}$$

Example 4.9 For the Rayleigh fading model (Example 4.3),

$$\mathbb{E}[H^\delta] = \mu^\delta \Gamma(\delta+1)\,. \tag{4.17}$$

Example 4.10 The Nakagami fading model (Example 4.4) generalizes the Rayleigh model, and we have

$$\mathbb{E}[H^\delta] = \Big(\frac{\mu}{k}\Big)^\delta \frac{\Gamma(\delta+k)}{\Gamma(k)}\,. \tag{4.18}$$

4.1.5 Service zones or cells

The exchange formula in Eq. (4.7) does not assume any particular serving zones based on base stations. Several stations can serve a given point, but such zones can be incorporated into the model. One classical model assumes that each location $y \in \mathbb{R}^2$ in the plane is served by the station from which it receives the strongest signal or, equivalently, the smallest value of Θ. This assumption is equivalent to the following definition of service zones, which we refer to as cells.

DEFINITION 4.1.5 *For any base station* $\mathsf{x} \in \Phi$, *its* service zone *or* cell *is*

$$C_\mathsf{x} = \big\{y \in \mathbb{R}^2 : \mathsf{L}_\mathsf{x}(y) \le \min_{\mathsf{y}\in\Phi} \mathsf{L}_\mathsf{y}(y)\big\}\,. \tag{4.19}$$

Although C_x are well-defined random sets, some further assumptions are needed to ensure that they have well-behaving mathematical properties. But we skip such details, and assume implicitly that we can integrate over C_x with respect to the Lebesgue measure. We also assume that the Lebesgue measure of all the locations that belong to two or more stations is equal to zero with probability one.

Example 4.11 By replacing the $\mathsf{L}_\mathsf{x}(y)$ with the geometric distance $|y-\mathsf{x}|$ in Eq. (4.19), we obtain as a special case $C_\mathsf{x} = V_\mathsf{x}$, the *Voronoi cell* of each base station $\mathsf{x} \in \Phi$, which is defined as all the locations closer to x than any other point of Φ.

4.1.6 Typical cell vs. zero-cell

The cell C_o of the typical station $\mathsf{x}_o = o$ considered under $\mathbb{P}_o$ is called the *typical cell*. One might be tempted to compare this cell to the cell covering an "arbitrarily" selected location in the plane. The latter is represented by the origin o under stationary probability $\mathbb{P}$ and the cell that covers it is sometimes called the *zero-cell*, denoted here by C_*. Formally $C_* = C_{\mathsf{x}}$ for x such that the origin $o \in C_{\mathsf{x}}$, and we can safely ignore the event of the origin being located on the boundaries of several cells, which occurs with probability zero. The exchange formula (Section 4.1.4) with $\chi(\mathsf{x}, y) = \mathbf{1}(y \in C_{\mathsf{x}})\bar{\varphi}(C_{\mathsf{x}})$, where $\bar{\varphi}(A)$ is a translation invariant set function, and Eq. (4.21) gives the relation

$$\mathbb{E}[\bar{\varphi}(C_*)] = \lambda \mathbb{E}_o[\bar{\varphi}(C_o)|C_o|] = \frac{\mathbb{E}_o[\bar{\varphi}(C_o)|C_o|]}{\mathbb{E}_o[|C_o|]}, \tag{4.20}$$

where $|A|$ is the area of $A \subset \mathbb{R}^2$. The distribution of the zero-cell is equal to the area-biased distribution of the typical cell. The biasing corresponds to sets with stochastically or statistically larger areas. If we apply $\bar{\varphi}(A) = 1/|A|$ to Eq. (4.20) and use Jensen's inequality, we get $1/\mathbb{E}[|C_*|] \leq \mathbb{E}[1/|C_*|] = 1/\mathbb{E}_o[|C_o|]$ and consequently $\mathbb{E}[|C_*|] \geq \mathbb{E}_o[|C_o|]$. Unfortunately, neither the distribution of the typical cell nor that of the zero-cell is known explicitly, even for Voronoi tessellation, but some results on the moment characteristics do exist (Chiu et al. 2013, Section 9.7.2).

Example 4.12 By applying the exchange formula in Eq. (4.7) with $\chi(\mathsf{x}, y) = \mathbf{1}(y \in \mathsf{L}_{\mathsf{x}})$, the mean area of typical cell is equal to the inverse of the density of stations

$$\mathbb{E}_o[|C_o|] = \frac{1}{\lambda}. \tag{4.21}$$

This result holds regardless of the random propagation effects or the path loss function.

4.1.7 Rate coverage

By the term *peak bit rate* of a user, we mean the maximal transmission rate that a user can get at a given location from a given station under the theoretical assumption that there is only one user in the network at this location, and it does not share network resources with other users. Put another way, this is a potential characteristic of a location in the network, and it is perfectly amenable to our typical location analysis and the exchange formula. We now define the peak bit rate more precisely.

DEFINITION 4.1.6 *We write $\boldsymbol{R}(\mathsf{x}, y)$ as the peak bit rate at location y with respect to the station $\mathsf{x} \in \Phi$, where the function $\boldsymbol{R}(\mathsf{x}, y)$ can be any translation invariant function of our network model $\tilde{\Phi}$.*

Example 4.13 An important choice for $\mathbf{R}(\mathsf{x}, y)$, inspired by Shannon's law, is taking $\mathbf{R}(\mathsf{x}, y) = \log(1 + \mathrm{SINR}(\mathsf{x}, y))$, where $\mathrm{SINR}(\mathsf{x}, y)$ is the SINR at y with respect to base station x, which has the form

$$\mathrm{SINR}(\mathsf{x}, \mathsf{y}) \stackrel{\triangle}{=} \frac{P_{rx}(\mathsf{x} - y)}{N + \sum\limits_{\mathsf{y} \in \Phi \setminus \{\mathsf{x}\}} P_{rx}(\mathsf{y} - y)}, \tag{4.22}$$

where N is a noise term and $P_{rx}(\mathsf{x} - y)$ is the received signal at y from base station x.

For the general case, we denote by $\mathbf{N}_{\mathbf{R}}(t)$ the average number of stations providing the typical location with the peak bit rate bigger than $t > 0$, that is

$$\mathbf{N}_{\mathbf{R}}(t) \stackrel{\triangle}{=} \mathbb{E}\left[\sum_{\mathsf{x} \in \Phi} \mathbf{1}(\mathbf{R}(\mathsf{x}, o) > t)\right], \tag{4.23}$$

we call $\mathbf{N}_{\mathbf{R}}(t)$ the *average coverage number at rate t*. We write $\mathcal{P}_{\mathbf{R}}(t, y)$ for the probability that the typical cell offers a peak bit rate greater than t at location $y \in \mathbb{R}^2$

$$\mathcal{P}_{\mathbf{R}}(t, y) \stackrel{\triangle}{=} \mathbb{P}_o\{\, \mathbf{R}(o, y) > t \,\}, \tag{4.24}$$

we call the function $\mathcal{P}_{\mathbf{R}}(t, y)$ *typical cell coverage probability*. Applying the exchange formula (Section 4.1.4) with $\chi(\mathsf{x}, y) = \mathbf{1}(R(\mathsf{x}, y) > t)$, we obtain the relation

$$\mathbf{N}_{\mathbf{R}}(t) = \lambda \int_{\mathbb{R}^2} \mathcal{P}_{\mathbf{R}}(t, y)\, \mathrm{d}y\,. \tag{4.25}$$

We focus on these concepts in Chapters 5–7 in the specific case of $\mathbf{R}(\mathsf{x}, y) = \log(1 + \mathrm{SINR}(\mathsf{x}, y))$, where explicit expressions are possible under certain model assumptions.

4.1.8 Cell loads

We return to the concept of service cells C_{x} by considering the general peak bit rate $\mathbf{R}(\mathsf{x}, y)$ at location y with respect to the station x serving y such that $y \in C_{\mathsf{x}}$, where again we ignore the null events when y is on cell boundaries.

Even if the peak bit rate $\mathbf{R}(\mathsf{x}, y)$ characterizes only a potential service that a user located at y could receive if it were the only user served by the network, we can still relate it to an important characteristic describing the performance of the network with a process of true users, which request service at different locations and time epochs, and share the network resources.[2]

The peak bit rate $\mathbf{R}(\mathsf{x}, y)$ represents the supremum of the traffic rate, the number of bits per second, that can be sustained at y in transmissions from (or to) station x. For the whole cell C_{x}, the corresponding *critical cell traffic rate* can be defined as the harmonic mean of $\mathbf{R}(\mathsf{x}, y)$ over $y \in C_{\mathsf{x}}$

[2] A further model extension is possible here that accounts for a space-time Poisson process of service requests coupled with a variant of the queueing-theoretic model for resource sharing (Błaszczyszyn, Jovanović, & Karray 2014).

$$\rho_{\mathbf{x}}^{c} \stackrel{\triangle}{=} \frac{|C_{\mathbf{x}}|}{\int_{C_{\mathbf{x}}} (\mathbf{R}(\mathbf{x}, y))^{-1} \, \mathrm{d}y} \tag{4.26}$$

where a harmonic mean is the inverse of an arithmetic mean of inverses, and such calculations usually appear when averaging rates. Assuming an appropriate Markovian space-time processor-sharing model for user service, the value of $\rho_{\mathbf{x}}^{c}$ corresponds to the value of the total requested traffic rate below which the station is able to transmit bits without delays statistically increasing in time, so we say that the cell is stable. When the total requested traffic rate is greater than or equal to $\rho_{\mathbf{x}}^{c}$, these delays increase to infinity and the cell is unstable.

We now assume that a uniform traffic demand ρ (bits per second per unit of surface) is requested at each location y, so the total requested traffic rate at cell $C_{\mathbf{x}}$ is $\rho \, |C_{\mathbf{x}}|$. The ratio between this total requested traffic rate and the critical one is called the *cell load*, which we denote by

$$\theta_{\mathbf{x}} \stackrel{\triangle}{=} \frac{\rho |C_{\mathbf{x}}|}{\rho_{\mathbf{x}}^{c}} = \rho \int_{C_{\mathbf{x}}} (\mathbf{R}(\mathbf{x}, y))^{-1} \, \mathrm{d}y \, , \tag{4.27}$$

where the second equality is due to Eq. (4.26). Cell loads are important performance metrics of cellular networks. A cell is stable if and only if its load is strictly smaller than one. In a Markovian processor-sharing queue with load $\theta < 1$, the distribution of the number of users (or calls) served at a given time is a geometric random variable with parameter θ, where θ corresponds to the fraction of time that the cell is idle.

Applying the exchange formula in Eq. (4.7) with $\chi(\mathbf{x}, y) = \rho / \mathbf{R}(\mathbf{x}, y) \mathbf{1}(y \in C_{\mathbf{x}})$ we obtain a relation between the *mean typical cell load* $\mathbb{E}_o[\theta_o]$ and the peak bit rate at the origin (that is, an arbitrary location) with respect to the serving station $\mathbf{x}_*$, namely

$$\frac{\rho}{\lambda} \, \mathbb{E}[(\mathbf{R}(\mathbf{x}_*, o)^{-1}] = \mathbb{E}_o[\theta_o] \, . \tag{4.28}$$

Given that $1/\lambda$ is the mean area of the typical cell, the quantity ρ/λ can be interpreted as the mean total requested traffic rate at the typical cell. Equation (4.28) says then that the mean typical cell load is equal to the mean total requested traffic rate at this cell multiplied by the mean inverse of the peak bit rate at the typical location in the network.

4.2 Heterogeneous Poisson Network Seen at the Typical Location

We now take the typical location approach to examine some interesting and useful characteristics of a heterogeneous Poisson network, where the base stations form a Poisson point process Φ with constant intensity λ, giving a stationary network model.

To clarify our terminology, the term *heterogeneous Poisson network* used in this book and other literature should be distinguished from inhomogeneous or nonhomogeneous Poisson process (or network model). The latter terms refer to a nonstationary Poisson process, whereas the former is a stationary Poisson process marked by different station types, such as transmission powers. By a *homogeneous Poisson network model*, we mean a stationary Poisson model where we do not distinguish station types.

4.2.1 Projection process and a propagation invariance

We are studying the network from one location, the origin, so we write H_{x} as the propagation effects between station $\mathsf{x} \in \Phi$ and the origin. Again, given Φ, the variables H_{x} are iid with some common distribution.

DEFINITION 4.2.1 (Projection process) *For a single observer located at the origin, the* projection process *is the point process*

$$\Theta \stackrel{\triangle}{=} \left\{ \frac{1}{H_{\mathsf{x}} \ell(\mathsf{x})} : \mathsf{x} \in \Phi \right\} \tag{4.29}$$

$$= \{ \mathsf{L}_{\mathsf{x}} : \mathsf{x} \in \Phi \},$$

which is defined on the positive half-line $\mathbb{R}^+$.

We call each L_{x} the *propagation loss*[3] between the user and station x. The point process Θ represents the inverse power values of the signals the user observes or experiences at the origin. We refer to this point process as the projection process because it is the projection of the underlying process Φ onto the positive real line via the propagation model. The process has been independently introduced a couple of times, and is also called the *propagation (loss) process* or the *path loss process with fading*. The strongest signals are the ones closest to the origin. In the next result, Lemma 4.2.2, we will see why the literature and we define the point process Θ to model the inverse received powers instead of the actual power values. Understanding the point process Θ, alongside its variations and extensions, is key for developing tractable models for the SINR and other quantities.

We often use the simple (that is, singular) path loss function in Eq. (4.2), which, coupled with a homogeneous Poisson network, gives tractable results.

LEMMA 4.2.2 *We assume that*

$$\mathbb{E}[H^{\delta}] < \infty, \tag{4.30}$$

where $\delta \stackrel{\triangle}{=} 2/\beta$. *Then the process* Θ *is an inhomogeneous Poisson point process with intensity measure* $Q([0,t]) = at^{\delta}$, *where the constant a is given by*

$$a \stackrel{\triangle}{=} \frac{\lambda \pi \mathbb{E}[H^{\delta}]}{\kappa^2}. \tag{4.31}$$

We skip the proof, which is an application of the Poisson mapping theorem, and instead prove an extension of Lemma 4.2.2 in Lemma 4.2.3. If we had chosen to define Θ as the process of received power values, then we would have an intensity measure with a singularity at $t = 0$, which would give an infinite measure for bounded sets of the form $[0, t]$, so the inverse power values are used instead.[4]

[3] Note that, unlike $\ell(\cdot)$, which is called a path loss but is strictly a path gain in the way it is defined, the propagation loss is really a loss by definition because it is the ratio of the power transmitted by the base station at x to the power received by the user at the origin.

[4] It is still possible to define such a point process of power values by using a more technical approach to point processes, where the intensity measure only needs to be bounded on closed, bounded subsets of $(0, \infty)$.

Example 4.14 We write $\{\mathsf{L}^{(i)}\}_{i\geq 1}$ to denote the points of Θ numbered in increasing order. Because $\Theta = \{\mathsf{L}_{\mathsf{x}} : \mathsf{x} \in \Phi\}$ is a Poisson process, the distribution of $\{\mathsf{L}^{(i)}\}_{i\geq 1}$ is

$$\begin{aligned}\mathbb{P}\{\,\mathsf{L}^{(n)} > y\,\} &= \mathbb{P}\{\,\Theta([0,y]) \leq k-1\,\} \\ &= \mathbb{P}\left\{\sum_{\mathsf{L}\in\Theta} \mathbf{1}(\mathsf{L}\in[0,y])) \leq k-1\right\} \\ &= \exp(-ay^{\delta})\sum_{k=0}^{n-1}\frac{(ay^{\delta})^{k}}{k!}\,,\end{aligned}$$

where a is given by Eq. (4.31). The kth smallest value of the process Θ does not necessarily correspond to the kth distant station, unless H is deterministic, that is, the model has no random propagation effects.

Example 4.15 By setting $\kappa = 1$, $\alpha = 1$ and $H \equiv 1$, we recover from Θ the distances to the successive base stations, which we recall has the tail distribution

$$\mathbb{P}\{\,\mathsf{R}^{(n)} > r\,\} = \exp(-\lambda\pi r^2)\sum_{k=0}^{n-1}\frac{(\lambda\pi r^2)^k}{k!}\,.$$

Remark 4.3 (Propagation invariance) The intensity measure of a Poisson process defines it. Based on our choice of ℓ, Eq. (4.31) reveals that the distribution of Θ and various functions of it, such as those in Examples 4.14 and 4.15 and the SINR, depend only on the propagation effects model H through one moment $\mathbb{E}[H^{\delta}]$, so the simple path loss function results in a kind of *propagation invariance* in the network model. Strikingly, different distributions of H that share common values of the moment $\mathbb{E}[H^{\delta}]$ stochastically or statistically generate the same propagation process, where we recall expressions for $\mathbb{E}[H^{\delta}]$ were given in Examples 4.7 to 4.10.

Example 4.16 Instead of having random H, we can, for mathematical convenience, represent the random propagation effects by setting $H \equiv 1$ and replacing λ with $\tilde{\lambda} = \lambda\mathbb{E}[\tilde{H}^{\delta}]$, where $\mathbb{E}[\tilde{H}^{\delta}]$ is the δ-order moment of the propagation effects that we originally wanted to introduce into the model.

Example 4.17 Another convenient choice is setting $H = E$, where E is an exponential random variable with unit mean, and replacing λ with $\tilde{\lambda} = \lambda\mathbb{E}[H^{\delta}]/\,\Gamma(1+\delta)$, where $\Gamma(1+\delta) = \mathbb{E}[E^{\delta}]$ is the δ-order moment of the normalized exponential variable. In the above example and this one, the value of a, given by Eq. (4.31), remains unchanged.

Another case where Lemma 4.2.2 simplifies calculations is when deriving the Laplace transform of the total received power, which we already gave in Eq. (4.14).

4.2.2 Heterogeneous Poisson network

We can extend the homogeneous Poisson network model by assigning random independent marks to each station in the network, as we did with the marked process in Eq. (4.4). For each $\mathsf{x} \in \Phi$, we recall that T_{x} is some parameter or vector of parameters of the base station x with values in some state space $\mathcal{T}$, and we assume that, given Φ, the T_{x} does not depend on the location of base station x but it can, in general, depend on the propagation effects H_{x}. In short, given the point process Φ, the marks $(H_{\mathsf{x}}, T_{\mathsf{x}})$ are iid across $\mathsf{x} \in \Phi$. The marking theorem (Theorem 3.4.4) says that jointly

$$\tilde{\Phi} \stackrel{\triangle}{=} \{\mathsf{x}, (H_{\mathsf{x}}, T_{\mathsf{x}}) : \mathsf{x} \in \Phi\}$$

forms an independently marked Poisson point process. We denote by (H, T) a generic random variable or mark. For $\tau \subset \mathcal{T}$, we write $G^T(\tau) = \mathbb{P}(T \in \tau)$ and $G^{T\,|\,H}(\tau \,|\, s) = \mathbb{P}(T \in \tau \,|\, H = s)$, respectively, for the marginal distribution on T and the conditional distribution of T given that $H = s$.

Example 4.18 A popular model for heterogeneous networks is known as the *multitier network model*, which consists of J types or tiers of base stations. Each tier of base station locations is usually modeled by an independent homogeneous Poisson point process Φ_j in the plane $\mathbb{R}^2$ with density λ_j. We write $\Phi = \bigcup_{j=1}^{J} \Phi_j$ for the superposition of $\Phi_1, \ldots, \Phi_J$, so capturing the locations of all the base stations of the network. We denote by $T_{\mathsf{x}} \in \mathcal{T} \stackrel{\triangle}{=} \{1, \ldots, J\}$ the type of station $\mathsf{x} \in \Phi$, meaning T_{x} is the index of the tier to which x belongs. By the Poisson superposition property of independent Poisson processes, covered in Section 3.4.2, the point process Φ is also a Poisson point process with density $\lambda = \sum_{j=1}^{J} \lambda_j$ and T_n form independent, identically distributed marks of Φ with the probability that an arbitrarily chosen station is of type j equal to $G^T(\{j\}) = \lambda_j/\lambda$.

We consider now two extensions of the projection process Θ of the network defined by Eq. (4.29). The first is the *projection process with station types*,

$$\widetilde{\Theta} \equiv \{(\mathsf{L}_{\mathsf{x}}, T_{\mathsf{x}}) : \mathsf{L}_{\mathsf{x}} \in \Theta\}, \tag{4.32}$$

and the second is *projection process with station types and distances*,

$$\widetilde{\Theta}' \equiv \{(\mathsf{L}_{\mathsf{x}}, (\mathsf{R}_{\mathsf{x}}, T_{\mathsf{x}})) : \mathsf{L}_{\mathsf{x}} \in \Theta\}, \tag{4.33}$$

where $\mathsf{R}_{\mathsf{x}} = \|\mathsf{x}\|$. Both $\widetilde{\Theta}$ and $\widetilde{\Theta}'$ are independently marked point processes with their points on $\mathbb{R}^+$, whereas their marks are in $\mathcal{T}$ and $\mathbb{R}^+ \times \mathcal{T}$, respectively. We can interpret $\widetilde{\Theta}$ as what an observer at the origin sees in the network in terms of signal strengths and station types, while the observer can also see base station distances in $\widetilde{\Theta}'$. The distribution of $\widetilde{\Theta}$ determines all the characteristics of a single location that can be expressed as a function of Θ such as the SINR, spectral and energy efficiency, and so on.

The marked point processes $\widetilde{\Theta}$ and $\widetilde{\Theta}'$ are also both Poisson processes by the next result, an extended version of Lemma 4.30, which treats jointly the points of Θ, the base station distances, and the randoms marks representing the base station types.

THEOREM 4.2.3 *Under the assumptions of Lemma 4.2.2, the process $\widetilde{\Theta}'$ is an independently marked inhomogeneous Poisson point process with intensity measure*

$$\tilde{Q}'([0,y]\times[0,\rho]\times\tau) \stackrel{\triangle}{=} \mathbb{E}[\#\{(\mathsf{L}_\mathsf{x},(\mathsf{R}_\mathsf{x},T_\mathsf{x}))\in\widetilde{\Theta}' : \mathsf{L}_\mathsf{x}\le y, \mathsf{R}_\mathsf{x}\le\rho, T_\mathsf{x}\in\tau\}]$$

$$= \int_0^y G_u(\rho,\tau)Q(\mathrm{d}u), \tag{4.34}$$

where $y,\rho\ge 0$, $\tau\subset\mathcal{T}$ and

$$G_u(\rho,\tau) = \frac{\mathbb{E}[H^\delta \mathbf{1}(H\le(\kappa\rho)^\alpha/u)\mathbf{1}(T\in\tau)]}{\mathbb{E}[H^\delta]}. \tag{4.35}$$

Before giving the proof, we note that G_u is the joint conditional distribution of the distance to a base station and its type, given that its propagation loss with respect to the single location is equal to u, which means that

$$G_u(\rho,\tau) = \mathbb{P}(\mathsf{R}\le\rho, T\in\tau \mid L=u). \tag{4.36}$$

The conditional distribution of R given that $\mathsf{L}=u$

$$G_u^{\mathsf{R}}(\rho) \stackrel{\triangle}{=} G_u(\rho,\mathcal{T}) = \frac{\mathbb{E}[H^\delta\,\mathbf{1}(H\le(\kappa\rho)^\alpha/u)]}{\mathbb{E}[H^\delta]} \tag{4.37}$$

depends on u. The dependence on u, however, is not the case for the conditional distribution of the base station type

$$\tilde{G}^T(\tau) \stackrel{\triangle}{=} G_u(\infty,\tau) = \frac{\mathbb{E}[H^\delta\,\mathbf{1}(T\in\tau)]}{\mathbb{E}[H^\delta]}. \tag{4.38}$$

The conditional distribution $\tilde{G}^T(\tau)$, as given by Eq. (4.38), is the distribution of the base station type T whose path loss is equal to $\mathsf{L}=u$. Interestingly, this conditional distribution does not depend on the value of u, which means that type T of the base station is independent of its propagation loss value L. This is the case despite the fact that this conditional distribution can, in general, depend on the respective propagation effects H, because we allow dependence between the components of the vector (H,T). In other words, $\widetilde{\Theta}$ is an independent and identically distributed marked version of process Θ.

In general, the distribution $\tilde{G}^T(\tau)$ is different from $G^T(\tau)=\mathbb{E}[\mathbf{1}(T\in\tau)]$, which is the distribution of the station type given the location of the base station. The distribution $G^T(\tau)$ does not depend on the station location as a consequence of the independent and identically distributed marking assumption of the model. Expression (4.38) reveals that $\tilde{G}^T(\tau)$ is a H^δ-biased version of $G^T(\tau)$, which is a surprising result.

Example 4.19 We assume again log-normal propagation effects $H=\hat{H}$ as defined by Eq. (4.3). Then

$$G_u(\rho,\tau) = \mathbb{E}\Big[\mathbf{1}(\hat{H}\mathrm{e}^{2\sigma^\delta}\le(\kappa\rho)^\alpha/u)G^{T|\hat{H}}(\tau|\hat{H}\mathrm{e}^{2\sigma^\delta})\Big] \tag{4.39}$$

for $\tau\subset\mathcal{T}$, where $G^{T|S}(\tau|s)$ is the conditional distribution of T given S. From this, we have the two conditional distributions

$$\tilde{G}^T(\tau) = \mathbb{E}[G^{T|\hat{H}}(\tau|\hat{H}\mathrm{e}^{2\sigma^\delta})], \tag{4.40}$$

and

$$G_u^{\mathsf{R}}(\rho) = \mathbb{E}[\mathbf{1}(\hat{H}\mathrm{e}^{2\sigma^\delta} \le (\kappa\rho)^\alpha/u)]. \tag{4.41}$$

The distribution of the distance to a random base station, given that its propagation loss is u, is the log-normal random variable

$$\mathsf{R}_u \stackrel{\triangle}{=} \frac{u^{1/\alpha}}{\kappa}\exp\Big[\frac{2\sigma^2}{\alpha^2}\Big]\hat{H}^{1/\alpha}. \tag{4.42}$$

This distribution is different from the unconditional distribution to any of the base stations.

Proof of Theorem 4.2.3 By the displacement theorem (Baccelli & Błaszczyszyn 2009a, Section 1.3.3) $\tilde{\Theta}$ is an independently marked Poisson point process with intensity measure

$$\begin{aligned}
\tilde{Q}'([0,y]\times[0,\rho]\times\tau) &= 2\pi\lambda\mathbb{E}\int_0^\infty \mathbf{1}((r\kappa)^\alpha/H \le y)\mathbf{1}(r\le\rho)\mathbf{1}(T\in\tau)r\,dr \\
&= \pi\lambda\mathbb{E}[\min(\rho^2,(yH)^\delta/\kappa^2))\mathbf{1}(T\in\tau)] \\
&= \frac{\pi\lambda}{\kappa^2}\mathbb{E}[H^\delta\min((\rho\kappa)^2/H^\delta,y^\delta)\mathbf{1}(T\in\tau)] \\
&= \frac{\pi\lambda}{\kappa^2}\mathbb{E}[H^\delta\min((\ell(\rho)H)^{-\delta},y^\delta)\mathbf{1}(T\in\tau)] \\
&= \frac{\delta\pi\lambda}{\kappa^2}\mathbb{E}\left[H^\delta\int_0^y \mathbf{1}(H\le 1/(u\ell(\rho)))\mathbf{1}(T\in\tau)u^{\delta-1}\,du\right] \\
&= \frac{\delta\pi\lambda\mathbb{E}[H^\delta]}{\kappa^2\mathbb{E}[H^\delta]}\int_0^y \mathbb{E}[H^\delta\mathbf{1}(H\le 1/(u\ell(\rho)))\mathbf{1}(T\in\tau)]u^{\delta-1}\,du,
\end{aligned}$$

where in the last line we exchanged the integral and expectation, completing the proof of Theorem 4.2.3. We also prove Lemma 4.2.2 by putting $\rho=\infty$ and $\tau=\mathcal{T}$ and recognizing that

$$\begin{aligned}
Q(du) &= \frac{2\pi\lambda\mathbb{E}[H^\delta]}{\alpha\kappa^2}u^{\delta-1}du \\
&= a\delta u^{\delta-1}du,
\end{aligned}$$

is the intensity measure of the unmarked process Θ.

4.2.3 Poisson network equivalence

The idea of network equivalence essentially means that we can recast a heterogeneous Poisson network model, which has stations of different types, into a homogeneous Poisson model with statistically indistinguishable stations, and for a single user, it will appear stochastically identical.

DEFINITION 4.2.4 (Network equivalence) *Two heterogeneous network models $\tilde{\Phi}_1$ and $\tilde{\Phi}_2$ are said to be* equivalent*, from the point of view of a user at the typical location, if their corresponding projection processes with station types $\widetilde{\Theta}_1$ and $\widetilde{\Theta}_2$ have the same distribution.*

We will see that network equivalence works behind the scenes of many of the results for the Poisson model. Conveniently, we can use Theorem 4.2.3 to construct an equivalent heterogeneous network with some of the previously random marks set to constants. In a way, this gives us an ability to move randomness away from certain marks onto others.

COROLLARY 4.2.5 *For a given heterogeneous network $\widetilde{\Phi}$, there is an equivalent heterogeneous network without random propagation effects*

$$\widetilde{\Phi}^* = \{(\mathsf{x}^*, (H_{\mathsf{x}^*} = 1, T_{\mathsf{x}^*})) : \mathsf{x}^* \in \Phi^*\}, \tag{4.43}$$

with density

$$\lambda^* \stackrel{\triangle}{=} \lambda \mathbb{E}[H^\delta] = a\kappa^2/\pi, \tag{4.44}$$

where given Φ^*, the marks T_{x^*} are independent (across each $\mathsf{x}^* \in \Phi^*$), and with a distribution given by

$$\mathbb{P}(T^* \le t) = \frac{\mathbb{E}\left[H^\delta \mathbf{1}(T \le t)\right],}{\mathbb{E}[H^\delta]}. \tag{4.45}$$

Proof The proof is done by setting the marks to $H = 1$, substituting F_{T^*}, and verifying that the resulting process forms a marked Poisson process with intensity measure $\tilde{Q}([0, y] \times \tau) \stackrel{\triangle}{=} \tilde{Q}'([0, y] \times [0, \rho = \infty] \times \tau)$ as in Eq. (4.34), where we have set $\rho = \infty$ to ignore distances to the stations.

The above result reveals that network equivalence in this setting hinges upon the Poisson assumption and the simple path loss function $\ell(x) = (\kappa \|x\|)^{-\alpha}$.

Example 4.20 *Multitier network with different transmission powers*. For the multitier network model in Example 4.18, we now assume that base station x from tier j, that is, $\mathsf{x} \in \Phi_j$, transmit signals with constant power p_j, where $j = 1, \ldots, J$. We consider the propagation effects that account for these powers and are equal to $H_H \stackrel{\triangle}{=} p_j H'_{\mathsf{x}}$ for stations $\mathsf{x} \in \Phi_k$, where H'_{x} are iid "pure" propagation effects. The power-aware propagation effects H of an arbitrarily chosen station have the distribution of $H' \times P$, where P is a discrete random variable taking the value p_j with probability λ_j/λ, and $\lambda = \sum_{j=1}^{J} \lambda_j$. A quick calculation gives the weighted sum

$$E[H^\delta] = \mathbb{E}[H'^\delta] \sum_{j=1}^{J} p_j^\delta \frac{\lambda_j}{\lambda} . \tag{4.46}$$

We define

$$a_j \stackrel{\triangle}{=} \frac{\pi \mathbb{E}\left[H'^\delta\right]}{\kappa^2} \lambda_j p_j^\delta , \tag{4.47}$$

and observe that $a = \sum_{j=1}^{J} a_j$ corresponds to Eq. (4.31).

Example 4.21 *Randomly selected station versus randomly selected signal*. We continue with the model considered in Example 4.20. Equation (4.35) says that a given received signal comes from a station of type k with probability

$$\begin{aligned} \frac{\mathbb{E}[H^\delta \mathbf{1}(T=k)]}{\mathbb{E}[H^\delta]} &= \frac{\mathbb{E}[H'^\delta] p_k^\delta \lambda_k/\lambda}{\mathbb{E}[H^\delta] \sum_{j=1}^{J} p_j^\delta \lambda_j/\lambda} \\ &= \frac{p_k^\delta \lambda_k}{\sum_{j=1}^{J} p_j^\delta \lambda_j} \\ &= \frac{a_j}{a} . \end{aligned}$$

This probability does not depend on the value of the received power: any given received power, large or small, can come from the station of type k with the same probability a_k/a. This probability is a power-biased fraction of stations of type k and is generally different from λ_k/λ, which is the probability that an arbitrarily chosen station is of type k.

Example 4.22 *Equivalent multitier network*. We now consider a single-tier network model with the intensity $\lambda = \sum_{j=1}^{J} \lambda_j$ equal to the total intensity of the base stations of the multitier network considered in Example 4.20. Assume that all base stations of this single-tier network transmit with *constant* power equal to

$$p^* = \left(\sum_{j=1}^{J} \frac{\lambda_j}{\lambda} p_j^\delta \right)^{1/\delta} . \tag{4.48}$$

It is straightforward to verify that the value of the constant $a = \sum_{j=1}^{J} a_j$ for the multitier network considered in Example 4.20 is equal to

$$a = \frac{\lambda \pi (p^*)^\delta \mathbb{E}[F^\delta]}{\kappa^2} .$$

Consequently, the projection process generated by the single-tier network has the same distribution as the projection process of the multitier network considered in Example 4.20. Furthermore, we can equivalently, so to speak, bring back network tiers to this single-tier network without altering the constant transmission power p^* used by its base stations. Indeed, motivated by the observation made in Example 4.21, let us *independently* mark the stations of the single-tier network by labels $T \in \{1, \ldots, J\}$ having distribution $\mathbb{P}\{T = j\} = a_j/a, j = 1, \ldots, J$, where a_j are given by Eq. (4.47). By direct

calculation, the marked propagation loss process observed in this single-tier network by a user located at the origin has the same intensity measure $\tilde{Q}$ given by Eq. (4.34) with $\rho = \infty$ (that is, ignoring the base station distances) as the initial multitier network. Observe, however, in Eq. (4.37) that the conditional distribution of the distance to the base station R, given its received power (or propagation loss $\mathsf{L} = u$), is different in the two models. This means that the considered equivalent multitier model cannot be used to calculate network characteristics depending explicitly on the distances to the base stations. In conclusion, all the probabilistic characteristics of the multitier network expressible in terms of the process $\tilde{\Phi}$ representing the powers received by a user at the typical location and the types of the corresponding stations (but not their distances!) can be calculated using the considered equivalent multitier network model.

Example 4.23 *Typical and zero-cell in multitier network model.* For the multitier network model considered in Example 4.20, the probability that the typical cell is of type j is equal to

$$\mathbb{P}_o\{T_o = j\} = \lambda_j/\lambda, \qquad j = 1, \ldots, J\,.$$

The probability that any given received power belongs to station j is equal to a_j/a, so the probability that the station of the zero-cell is of type j is

$$\mathbb{P}\{T_* = j\} = a_j/a\,,$$

where a_j is given by Eq. (4.47). Moreover we have an expression for the conditional mean area of the typical cell C_o given its type is equal to j

$$\begin{aligned}\mathbb{E}_o[|C_o| \,|\, T_o = j] &= \frac{\mathbb{E}_o[|C_0|\mathbf{1}(T_o = j)]}{\mathbb{P}_o\{T_o = j\}} \\ &= \frac{1}{\lambda}\frac{\mathbb{P}\{T_o = j\}}{\mathbb{P}^o\{Z_o = j\}} \\ &= \frac{a_j}{a\lambda_j} = \frac{1}{\lambda}\frac{p_j^\delta}{(p^*)^\delta}\,,\end{aligned}$$

where the second equality follows from the exchange formula in Eq. (4.7) with $\chi(\mathsf{x}, y) = \mathbf{1}(y \in C_{\mathsf{x}})\mathbf{1}(T_{\mathsf{x}} = j)$, and p^* is the power in the equivalent homogeneous network given by expression Eq. (4.48). By interpreting $1/\lambda$ as the (unconditional) mean area of the typical cell, given by the formula in Eq. (4.21), we obtain

$$\mathbb{E}_o[|C_o| \,|\, T_o = j] = \mathbb{E}_o[|C_o|]\,\Big(\frac{p_j}{p^*}\Big)^\delta\,.$$

For the mean area of the zero-cell conditioned on its base station's type, we obtain, again by the exchange formula Eq. (4.7), we get the expression

$$\mathbb{E}[|C_o| \,|\, T_o = j] = \frac{\lambda a_j}{a}\mathbb{E}_o[|C_o|^2\mathbf{1}(T_o = j)]\,.$$

Example 4.24 Quite similarly to the cell area considered in Example 4.23, under the multitier network model we obtain the conditional mean typical cell load given the station type

$$\mathbb{E}_o[\theta_o \mid T_o = j] = \mathbb{E}_o[\theta_o] \left(\frac{p_j}{p^*}\right)^{\delta}, \qquad j = 1, \ldots, J,$$

where $\mathbb{E}_o[\theta_o]$ is the unconditional mean cell load of the typical station satisfying Eq. (4.28).

4.2.4 Incorporating propagation terms such as transmission powers and antenna gains

We have not introduced into our model any terms for the power of the transmitted signal or the antenna gains, as discussed in Section 2.2. If these terms are constants, then the fading variables are just rescaled. If they are iid random variables, then mathematically they can be treated in the same manner as the general fading variables $\{H_{\mathsf{x}}\}_{\mathsf{x}\in\Phi}$. For example, given Φ, let the iid random variable $\{P_{tx,\mathsf{x}}\}_{\mathsf{x}\in\Phi}$ be the transmitted powers. For each $\mathsf{x} \in \Phi$, the variables $P_{tx,\mathsf{x}}$ and H_{x} need not be independent of each other, as long as they are independent across $x \in \Phi$, then the product $H_{\mathsf{x}}P_{tx,\mathsf{x}}$ can be treated as a single random variable. The process

$$\hat{\Theta} \stackrel{\triangle}{=} \left\{ \frac{1}{H_{\mathsf{x}}P_{tx,\mathsf{x}}\ell(\|\mathsf{x}\|)} : \mathsf{x} \in \Phi \right\}, \tag{4.49}$$

is also a Poisson point process. Under the path loss model $\ell(x) = (\kappa\|x\|)^{-\alpha}$, the process $\hat{\Theta}$ has an intensity measure given by

$$\hat{Q}((0,t]) = \hat{a}t^{\delta}, \tag{4.50}$$

where $\hat{a} = \pi\lambda\mathbb{E}[(P_{tx}H)^{\delta}]/\kappa^2$. Consequently, propagation terms such as transmission powers and antenna gains can be immediately incorporated into the model.

4.2.5 Intensity measure of a general projection process

For a non-Poisson process of base station locations, it is still possible to derive the intensity measure of the resulting projection process, essentially due to Campbell's theorem. But in general the projection process cannot be characterized by its intensity measure, as is the case for the Poisson model. We now use a general path loss function $\ell(x) = 1/h(\|x\|)$ where h is a left continuous and nondecreasing function with a generalized inverse $h^{-1}(y) = \inf\{x : h(x) > y\}$. We introduce a general point process $\Phi \subset \mathbb{R}^2$ representing the locations of the base stations. For all $\mathsf{x} \in \Phi$, H_{x} form again a collection of iid positive random variables representing random propagation effects. We now consider the point process on $\mathbb{R}^+$, analogous to Θ defined by Eq. (4.29), but induced by our general model of Φ, H, and h. The point process is defined as

$$\begin{aligned}\Theta &\stackrel{\triangle}{=} \left\{ \frac{1}{H_{\mathsf{x}}\ell(\mathsf{x})} : \mathsf{x} \in \Phi \right\} \\ &= \{\mathsf{L}_{\mathsf{x}} : \mathsf{x} \in \Phi\},\end{aligned} \tag{4.51}$$

where we reuse the notation Θ, L_{x}, and Q.

PROPOSITION 4.2.6 *For a general point process Φ with intensity measure $\Lambda(B) = E[\Phi(B)]$, the induced point process Θ has an intensity measure*

$$Q((0,t]) \stackrel{\triangle}{=} \mathbb{E}[\sum_{\mathsf{L}\in\Theta} \mathbf{1}(\mathsf{L} \le t)] \tag{4.52}$$

$$= \mathbb{E}[\Phi(B_0[h^{-1}(tH)])] \tag{4.53}$$

$$= \int_{\mathbb{R}^2} \mathbb{P}\left(0 < \frac{h(\|x\|)}{H} \le t\right) \Lambda(dx)\,. \tag{4.54}$$

If Φ is a general Poisson point process, then the point process Θ is an inhomogeneous Poisson point process with intensity measure Q.

Remark 4.4 The intensity measure Q has the term $\mathbb{E}[\Phi(B_0[h^{-1}(tH)])]$, which can be interpreted as the average number of points of Φ located in a disk centered at the origin with random radius $h^{-1}(tH)$. The expressions for Q, which is a first moment term, will hold if H is replaced with a random propagation field.

Example 4.25 For any stationary point process with density λ, the intensity measure of the projection process Θ is

$$Q((0,t]) = \pi\lambda\mathbb{E}\left[\left(h^{-1}(tH)\right)^2\right]. \tag{4.55}$$

Example 4.26 For $\ell(x) = (\kappa\|x\|)^{-\alpha}$ and the stationary network, the results quickly extend to $\mathbb{R}^d$ by replacing the disk with a d-dimensional ball. The intensity measure becomes

$$Q([0,t)) = a_d t^{d/\alpha}, \tag{4.56}$$

where

$$a_d \stackrel{\triangle}{=} \lambda\mathbb{E}[|b_0(H^{1/\alpha})|] = \lambda|b_0(1)|\mathbb{E}[H^{d/\alpha}], \tag{4.57}$$

which shows again that the propagation effects H is incorporated into the intensity measure of Θ through only one moment $\mathbb{E}[H^{d/\alpha}]$. For $d = 3$, $a_3 = \frac{4\lambda\pi}{3}\mathbb{E}[H^{3/\alpha}]/\kappa^3$.

Example 4.27 Returning to $\mathbb{R}^2$ and writing $r = \|x\|$, the *multislope* model for path loss (see Section 7.1.2) is

$$h(r) = \sum_{i=1}^{k+1} \mathbf{1}(r_{i-1} \le r < r_i) b_i^{-1} r^{\alpha_i},$$

where $0 = r_0 < r_1 < \cdots < r_k < r_{k+1} = \infty$, $\alpha_i > 0$, and $b_i > 0$ are all set such that h is continuous. Each interval $[r_{i-1}, r_i)$ is disjoint with all others, so the inverse of $h(r)$ is simply

$$h^{-1}(s) = \sum_{i=1}^{k+1} \mathbf{1}(s_{i-1} \le s < s_i) c_i s^{1/\alpha_i},$$

where $s_i = b_i^{-1} r_i^{\alpha_i}$ and $c_i = b_i^{1/\alpha_i}$. The induced process Θ has an intensity measure

$$Q((0,t]) = 2\pi\lambda\mathbb{E}(h^{-1}(tH)^2) = 2\pi\lambda \sum_{i=1}^{k} t^{\delta_i} c_i \mathbb{E}\left[\mathbf{1}(s_{i-1} \le tH < s_i) H^{\delta_i}\right]. \tag{4.58}$$

4.3 Networks Appear Poisson Due to Random Propagation Effects

We have seen the advantages of the typical location approach, particularly when the base stations form a Poisson process. But if base station locations do not appear Poisson, then the reason for using the Poisson point process comes down to pure tractability, and not model accuracy. This is, however, not the entire story. Under general model conditions, mathematical results show that the randomness in the signal propagation model can render the network to appear to a single user more Poisson. We present these results, based on the propagation model, further justifying the Poisson assumption. The reader who is more interested in questions on coverage probability can skip the rest of this chapter.

To explain networks appearing Poisson, we first consider a non-Poisson point process or a deterministic point pattern such as a hexagonal lattice. We then apply a random operation or transformation to this collection of points such as moving each point separately in some random direction by some random distance. If this particular point process operation, in this case a random translation, adheres to a couple of certain conditions, such as each point is moved independently of all other points, then each time the operation is applied, the resulting point process will appear, or rather, behave stochastically more Poisson. In the limit as the number of operations applied approaches infinity the point process converges in distribution, that is, weakly to a Poisson point process. Put informally, each operation is adding more randomness to the point process, and the Poisson point process is the inevitable and most random point process, with a nonrandom intensity measure, due to its property of complete independence. The Poisson point process is often the limiting point process of random point process operations, so it plays the role of the Gaussian distribution in point process theory.

Returning to network models, the source of the randomness that makes our non-Poisson network behave more Poisson is the random propagation effects of independent fading, shadowing, randomly varying antenna gains, and so on, or some combination of these. Under the standard propagation model, with a general path loss function and sufficiently large and independent propagation effects, the power values of the signals from the base stations appear *at a single location* as an inhomogeneous Poisson point process on the positive real line. This in turn makes the network configuration appear, from this location, Poisson, because an actual Poisson network with matching intensity measure would produce the same inhomogeneous Poisson point process on the positive real line. In other words, the more the propagation effects weaken the signals, the more the network will appear Poisson to a single user who does not know whether the source

of Poisson behavior is the network configuration, the random propagation effects, or some combination of the two.

For a single user, we show that even in a network with a deterministic configuration of transmitters, the projection process Θ of inverse signal strengths can appear Poisson, provided sufficiently strong and independent propagation effects.

4.3.1 Projection process based on a deterministic configuration of base stations

We assume a single user at the origin o and that the base stations are located according to some locally finite deterministic point pattern ϕ on $\mathbb{R}^2 \setminus \{o\}$, where the origin is removed to prevent the user being located on a transmitter. The local finiteness condition means there is not an infinite number of points in some bounded region of $\mathbb{R}^2$. The point pattern ϕ could be a realization of some random point process Φ. We often use the shorthand $\phi(r) \triangleq \phi(b_o(r))$ to denote the number of points of ϕ located in a disk $b_o(r)$ in $\mathbb{R}^2$ centered at the origin, and $|b_o(r)|$ is its area. We recall the general path loss function $\ell(x) = 1/h(\|x\|)$, where h is a left continuous and nondecreasing function.

The projection process induced by ϕ, H, and h is

$$\Theta \triangleq \left\{ \frac{h(\|\mathsf{x}\|)}{H_{\mathsf{x}}} : \mathsf{x} \in \phi \right\}, \tag{4.59}$$

which has the intensity measure

$$Q((0,t]) = \mathbb{E}\left[\sum_{\mathsf{L} \in \Theta} \mathbf{1}(\mathsf{L} \le t) \right] \tag{4.60}$$

$$= \mathbb{E}[\phi(h^{-1}(tH))], \tag{4.61}$$

where we have once again reused the notation Θ, L, and Q, but this time for a deterministic configuration of base stations. We define $p_{\mathsf{x}}(t) \triangleq \mathbb{P}\{0 < 1/(H_{\mathsf{x}}^{(v)} \ell(\mathsf{x})) \le t\} = \mathbb{P}\{0 < \mathsf{L}_{\mathsf{x}} \le t\}$ for each transmitter $\mathsf{x} \in \phi$, and note that the intensity measure Q, given by Eq. (4.61), is equivalent to

$$Q((0,t]) = \sum_{\mathsf{x} \in \phi} p_{\mathsf{x}}(t). \tag{4.62}$$

4.3.2 Poisson model approximation

For a deterministic point pattern ϕ, we want to approximate the resulting process Θ defined by Eq. (4.59) with a Poisson point process $\bar{\Theta}$ as defined by (4.29) on the positive half-line with intensity measure Q fitting Eq. (4.61) or (4.62). But to speak about approximations, we first need to introduce a suitable probability metric: for two probability measures ν_1 and ν_2 defined on the same probability space with σ-algebra $\mathcal{F}$ of events, the total variation distance is

$$d_{\mathrm{TV}}(\nu_1, \nu_2) \triangleq \sup_{A \in \mathcal{F}} |\nu_1(A) - \nu_2(A)|.$$

This probability metric bounds the largest difference in probabilities between the two distributions. We write $\boldsymbol{L}(E)$ to denote the probability distribution or law of some random element E such as a point process or random variable. Two point processes Θ and $\bar{\Theta}$ are stochastically close if their laws $\boldsymbol{L}(\Theta)$ and $\boldsymbol{L}(\bar{\Theta})$ have a small total variation $d_{\mathrm{TV}}(\boldsymbol{L}(\Theta), \boldsymbol{L}(\bar{\Theta}))$. To obtain meaningful values for the total variation, we must compare the point processes on a finite interval (it is not practical to consider an infinite number of points), which is a small but necessary restriction. For $\tau > 0$, we write $\Theta|_\tau$ and $\bar{\Theta}|_\tau$ to respectively denote the points of the point processes Θ and $\bar{\Theta}$ restricted to the interval $(0, \tau]$, where $1/\tau$ can be interpreted as the smallest possible power value of interest for a user in the network.

We now present an approximation theorem for the truncated point processes $\Theta|_\tau$ and $\bar{\Theta}|_\tau$.

THEOREM 4.3.1

$$\frac{1 \wedge Q(\tau)^{-1}}{32} \sum_{\mathsf{x} \in \phi} p_{\mathsf{x}}(\tau)^2 \leq d_{TV}(\boldsymbol{L}(\Theta|_\tau), \boldsymbol{L}(\bar{\Theta}|_\tau)) \leq \sum_{\mathsf{x} \in \phi} p_{\mathsf{x}}(\tau)^2 \leq Q(\tau) \max_{\mathsf{x} \in \phi} p_{\mathsf{x}}(\tau). \tag{4.63}$$

The theorem says that one wants $p_{\mathsf{x}}(t)$ to be small to use the Poisson approximation. In other words, if the network has many base stations with independent propagation effects and the chance of any transmitter having a strong signal is small, then the point process of signal strengths should be close to Poisson process with the same intensity measure. If x^* is the base station closest to the origin, then $\max_{\mathsf{x} \in \phi} p_{\mathsf{x}}(\tau) = \mathbb{P}\{0 < 1/(H_{\mathsf{x}^*} \ell(\mathsf{x}^*)) \leq \tau\}$. This term is not as small as $\sum_{\mathsf{x} \in \phi} p_{\mathsf{x}}(\tau)^2$, but it is usually easier to calculate.

The upper bound in Theorem 4.3.1 is an increasing function in τ. Interestingly, the smaller we make our window $(0, \tau]$, thereby considering only signals of power values greater than $1/\tau$, then the smaller the bounds, implying the more Poisson the signals in $(0, \tau]$ appear. In other words, the stronger signals behave stochastically more Poisson. This is useful because a functional dependent on k strongest signals can be well approximated with a functional of a Poisson process with intensity $M(t)$. But some care must be taken with such interpretation of Theorem 4.3.1. If τ is made too small, then the bounds lose meaning because there will be no signals in $(0, \tau]$. For large τ the upper bound will also lack meaning because it will be greater than one.

An appealing property of the total variation distance is that it also holds under functions of point processes, for a large family of functions.

Example 4.28 We introduce the point processes formed by the signal power values in the network

$$\Pi \stackrel{\triangle}{=} \left\{ \frac{1}{\mathsf{L}} : \mathsf{L} \in \Theta \right\}, \quad \bar{\Pi} \stackrel{\triangle}{=} \left\{ \frac{1}{\bar{\mathsf{L}}} : \bar{\mathsf{L}} \in \bar{\Theta} \right\}, \tag{4.64}$$

where $\bar{\Pi}$ is also a Poisson process. Both Π and $\bar{\Pi}$ have the intensity measure

$$\bar{Q}([\tau, \infty)) \stackrel{\triangle}{=} \sum_{\mathsf{P} \in \Pi} \mathbf{1}(\mathsf{P} \geq \tau) = Q((0, 1/\tau]).$$

Theorem 4.3.1 also implies, because total variation distance is preserved under one-to-one mappings, that

$$d_{\mathrm{TV}}(\boldsymbol{L}(\Pi|^{\tau}),\boldsymbol{L}(\bar{\Pi}|^{\tau})) \leq \sum_{\mathsf{x}\in\phi} p_{\mathsf{x}}(1/\tau)^2 \leq Q(1/\tau)\max_{\mathsf{x}\in\phi} p_{\mathsf{x}}(1/\tau), \tag{4.65}$$

where $\Pi|^{\tau}$ and $\bar{\Pi}|^{\tau}$, respectively, denote the point processes Π and $\bar{\Pi}$ restricted to $[\tau,\infty)$.

4.3.3 Order statistics of signals

We can use the aforementioned property of the total variation distance to, for example, pick the strongest signal from the two network models, which corresponds to the smallest value, respectively, from Θ and $\bar{\Theta}$. We then have a bound on the probability distance between these two random quantities. We develop this example further by looking at the order statistics of Θ and $\bar{\Theta}$. We recall $\mathsf{L}^{(1)} \leq \mathsf{L}^{(2)} \leq \ldots$, the increasing order statistics of the process Θ, which correspond to the decreasing order statistics of the process of power strengths. Similarly, let $\bar{\mathsf{L}}^{(1)} \leq \bar{\mathsf{L}}^{(2)} \leq \ldots$ be the order statistics of the Poisson process $\bar{\Theta}$. Then there are bounds on the total variation distance between these two sets of order statistics.

THEOREM 4.3.2 *For any $\tau > 0$, the inequality*

$$d_{TV}(\boldsymbol{L}(\mathsf{L}^{(1)},\ldots,\mathsf{L}^{(k)}),\boldsymbol{L}(\bar{\mathsf{L}}^{(1)},\ldots,\bar{\mathsf{L}}^{(k)})) \leq \sum_{\mathsf{x}\in\phi} p_{\mathsf{x}}(\tau)^2 + \sum_{j=0}^{k-1} \frac{Q(\tau)^j \mathrm{e}^{-Q(\tau)}}{j!},$$

holds.

Example 4.29 For a process Θ with intensity measure Q induced by some network model, we can use a Poisson process $\bar{\Theta}$ with the matching intensity measure Q to approximate the distributions of its signal strengths. Because $\bar{\Theta}$ is a Poisson point process, the probability distribution of the smallest value of $\bar{\Theta}$, denoted by $\bar{\mathsf{L}}_{(1)}$, is

$$\mathbb{P}(\bar{\mathsf{L}}_{(1)} \leq t) = 1 - \mathrm{e}^{-Q(t)},$$

which in turn can be used to approximate the distribution of $\mathsf{L}_{(1)}$, that is

$$\mathbb{P}(\mathsf{L}_{(1)} \leq t) \approx 1 - \mathrm{e}^{-Q(t)}.$$

Example 4.30 For continuous random variables defined on $[0,\infty)$, the definition of total variation distance implies

$$d_{\mathrm{TV}}(\boldsymbol{L}(\mathsf{L}_{(1)}),\boldsymbol{L}(\bar{\mathsf{L}}_{(1)})) = 1 - \int_0^{\infty} \min[f_{(1)},\bar{f}_{(1)}]dx,$$

where $f_{(1)}$ and $\bar{f}_{(1)}$ are the respective probability densities of $\mathsf{L}_{(1)}$ and $\bar{\mathsf{L}}_{(1)}$.

4.3.4 Fitting the Poisson model

For a simple model based on the intensity measure $Q(\tau)$, we just need to measure the signal strengths, and not the locations of the base stations, and use the estimate of $Q(\tau)$ to create a Poisson model. The empirical distribution of $\mathsf{L}_{(1)}$, denoted by $\hat{E}(t)$, gives a way to statistically estimate or fit $Q(t)$ by first assuming that the transmitters are positioned according to a Poisson model, even if the transmitters don't appear Poisson. We then measure the largest signal in the network in different locations and then fit the empirical distribution of $\bar{\mathsf{L}}_{(1)}$ to the equation

$$-\log[1-\hat{E}(t)] = \hat{Q(t)},\ t>0,$$

where $\hat{Q(t)}$ is the estimate of $Q(t)$. For a large cellular phone network, the intensity measure Q has been fitted to experimental signal data under a Poisson network model with $\ell(x) = \|x\|^{-\alpha}$ (Błaszczyszyn et al. 2015); also see Błaszczyszyn & Karray (2016) for work with radiation pattern models. Other work (Lu & Di Renzo 2015) has examined the intensity measure of the path loss process (so the process Θ with all $\{H_i\}_{i\geq 0}$ set to some constant) by using a more advanced path loss model and geographic data from cellular networks in two cities. If random H was incorporated into that model, Theorem 4.3.1 suggests that the Poisson model for the network locations would be a very suitable model.

4.3.5 Poisson convergence

Theorem 4.3.1 is an approximation theorem, but it has been used to show that Poisson convergence of the projection process occurs when the random variables representing propagation effects converge to zero. For each $\mathsf{x} \in \phi$, let $\{H_{\mathsf{x}}^{(v)}\}_{v\geq 0}$ be a family of positive random variables indexed by some parameter $v \geq 0$. The process induced by ϕ, $H^{(v)}$, and h is

$$\Theta^{(v)} \stackrel{\triangle}{=} \left\{ \frac{h(\|\mathsf{x}\|)}{H_{\mathsf{x}}^{(v)}} : \mathsf{x} \in \phi \right\}, \tag{4.66}$$

which has a intensity measure $Q^{(v)}(t) = \mathbb{E}[\phi(h^{-1}(tH^{(v)}))]$. Under these model assumptions, Poisson convergence occurs as the propagation effects converge to zero (Keeler, Ross, & Xia 2016a, Theorem 1.1).

THEOREM 4.3.3 *Let $\phi \subset \mathbb{R}^2/\{o\}$ be a locally finite point pattern satisfying*

$$\lim_{r\to\infty} \frac{\phi(r)}{\pi r^2} = \lambda > 0. \tag{4.67}$$

If as $v \to \infty$,

$$(i)\ H^{(v)} \xrightarrow{\mathbb{P}} 0, \quad \textit{and}$$

$$(ii)\ Q^{(v)}(t) = \mathbb{E}[\phi(h^{-1}(H^{(v)}t)] \to L(t),\ t>0,$$

then $\Theta^{(v)}$ converges weakly to a Poisson point process on $\mathbb{R}^+$ with intensity measure L, where $L(t)$ is some continuous function in t.

Remark 4.5 The main component of using Theorem 4.3.3 is finding a suitable combination of H and ℓ (or h) such that in the limit as $v \to \infty$ the intensity measure $Q^{(v)}(t)$ converges to some reasonable function $L(t)$. The condition of Eq. (4.67) is merely saying that the density of ϕ converges to some constant.[5] This condition is immediately satisfied by any (for example, hexagonal) lattice and any realization of an ergodic point process. The parameter v can be any parameter of the distributions of $H^{(v)}$.

Example 4.31 We assume log-normal shadowing and $\ell = (\kappa\|x\|)^{-\alpha}$, and $\phi \subset \mathbb{R}^2/\{o\}$ that meets condition of Eq. (4.67). We use the log-normal shadowing model in Eq. (4.3) by setting $v = \sigma$ and letting a collection of iid log-normal random variables $\hat{H}_{\mathsf{x}}^{(v)}$ be the shadowing between the user and base station x so that

$$\hat{H}_{\mathsf{x}}^{(v)} = \exp(vZ_{\mathsf{x}} - v^2/\alpha), \tag{4.68}$$

where Z_{x} are standard Gaussian random variables. Dropping the subscript, the above form of $\hat{H}^{(v)}$ implies that for any $s > 0$

$$\mathbb{P}\{\,\hat{H}^{(v)} > s\,\} = \mathbb{P}\{\,Z > \log(s)/v + v/\alpha\,\} \to 0 \quad \text{as } v \to \infty\,.$$

In other words, $S \xrightarrow{\mathbb{P}} 0$ as $v \to \infty$. But Eq. (4.16) for the moment $\mathbb{E}(\hat{H}^{\delta})$ implies $\mathbb{E}[(\hat{H}^{(v)})^{\delta}] = 1$ for all $v > 0$, so we have

$$Q^{(v)}(t) = \pi\lambda/\kappa^2 t^{\delta}, \tag{4.69}$$

Then Theorem 4.3.3 gives Poisson convergence as $v \to \infty$ by setting $H_{\mathsf{x}}^{(v)} = \hat{H}_{\mathsf{x}}^{(v)}$.

Remark 4.6 For an interpretation of Example 4.31, the shadowing variables are statistically and individually weakening the received powers from all the base stations. But the form of the intensity measure, given by Eq. (4.69), means that there are always strong-enough signals. These signals, however, come from more and more distant stations, whose exact geographic locations matter less and less. In the limit as v approaches infinity, this loss, so to speak, of the original geometry of the stations ϕ is a condition for obtaining Poisson convergence of the process $\Theta^{(v)}$. We also bring attention to an interesting observation (Błaszczyszyn et al. 2015, Remark 11). For large v, the base station distance $\mathsf{R}_{\mathsf{x}} = \|\mathsf{x}\|$ behaves asymptotically:

$$\mathsf{R}_{\mathsf{x}} \approx \exp\Big(\frac{v^2}{\alpha^2} + \frac{v}{\alpha}Z\Big)\,. \tag{4.70}$$

The distance R_{x} is asymptotically log-normal, which implies that in the limit as $v \to \infty$, the propagation values are detached or decoupled from the underlying geometry of the network.

Remark 4.7 The independence assumption on propagation effects greatly simplifies the model, but Ross & Schuhmacher (2017) have shown mathematically that the assumption can be lifted for the log-normal shadowing and $\ell = (\kappa\|x\|)^{-\alpha}$ case,

[5] The results still hold if density of transmitters is isotropic (Keeler et al. 2016a)

allowing for correlated shadowing. More specifically, they replaced the iid log-normal variables with

$$\hat{H}_x^{(\nu)} = \mathrm{e}^{\nu \mathbf{Z}_x - \nu^\delta},$$

where $\alpha > 2$, $\nu > 0$, and $\mathbf{Z}_x$ is a value of a zero-mean Gaussian random field at point $x \in \mathbb{R}^2$, and then proved Poisson convergence in the limit $\nu \to \infty$, provided some natural conditions on the base station locations and the covariance function of the Gaussian field. If the base station locations cannot be arbitrarily close and the covariance function of the Gaussian field has at least moderate polynomial decay, both very reasonable physical assumptions, then Poisson approximation is still possible for sufficiently large ν.

Example 4.32 For the same point pattern ϕ as in Theorem 4.3.3, we consider a Rayleigh fading model where for each $\mathsf{x} \in \phi$, $E_{\mathsf{x}}^{(\nu)}$ is an iid exponential random variable with mean $1/\nu$. Given $\mathbb{P}(E^{(\nu)} > s) = \mathrm{e}^{-\nu s}$, we have

$$E^{(\nu)} \xrightarrow{\mathbb{P}} 0 \quad \text{as} \quad \nu \to \infty,$$

which is condition (i) in Theorem 4.3.3. But for $\ell = (\kappa \|x\|)^{-\alpha}$, the intensity measure

$$Q^{(\nu)}(t) \to 0 \quad \text{as} \quad \nu \to \infty,$$

due to $\mathbb{E}([E^{(\nu)}]^\delta) = \Gamma(\delta + 1)/\nu^\delta$, so $L = 0$. Although the Rayleigh model does not give a reasonable L in the limit as $\nu \to \infty$, we stress that for large but finite ν the resulting process Θ of inverse signal strengths can still be approximated with a Poisson process with measure $Q^{(\nu)}$.

We can embellish Theorem 4.3.3 by assigning independent marks to each base station $\mathsf{x} \in \phi$, as was done in Section 4.2. One example of possible marks could be fading variables, then we can also apply the result to fading-shadowing models, such as the Suzuki model, formed by a product of a log-normal and some other independent random variable, which seeks to capture both multipath fading and shadowing (Keeler et al. 2016a, Section 3.2).

4.3.6 Possible extensions

We have outlined how the point process of inverse power values can often be approximated well with an inhomogeneous Poisson process, explaining why the Poisson model can be a good statistical fit. But the approximation and convergence results are a relatively new line of inquiry, suggesting more empirical validation is needed. Studies have investigated the claim of whether network layouts actually appear Poisson, but there has been little empirical work on how Poisson the process Θ is under certain models.

We assumed a deterministic or nonrandom initial base station configuration ϕ for the Poisson convergence and approximation results. But what happens when we replace ϕ with a random point process Φ? The results in Theorem 4.3.1 and its consequences still

hold with some straightforward modifications due to the total variation distance holding under expectation (see Keeler et al. 2016a, Theorem 2.7). For the limit as $v \to \infty$, the extra layer of randomness from the random Φ means that the resulting process $\Theta^{(v)}$ can converge to an inhomogeneous Poisson point process with a random intensity measure Q given by Eq. (4.52), hence it can converge to a Cox process. Cox convergence in this setting is a recent observation (Keeler et al. 2016a, Section 2.2), but it is possible under some conditions on Φ to ensure Poisson (and not Cox) convergence (Keeler et al. 2016a, Theorem 2.7).

The independence of propagation effects greatly simplifies the results, but Ross & Schuhmacher showed that Poisson approximation and covergence were still possible for correlated log-normal shadowing and power law path loss model (Ross & Schuhmacher 2017). Future work would involve extending these conditions. A competing model for shadowing with correlation is based on Poisson processes of lines or planes. In this model approach, the randomly placed lines and planes represent buildings in a cellular network, and some SINR results have already been successfully derived (Baccelli & Zhang 2015, Lee, Zhang, & Baccelli 2016, Zhang, Baccelli, & Heath 2015).

We have also insisted on treating the network perceived from only a single location. But what happens when there are multiple users? This line of inquiry could promise interesting results and observations, yet remains almost untouched. It would require a multidimensional treatment of the problem and a specific model for the propagation fields with correlation across different locations. But it is not entirely clear how this would be achieved.

4.4 Bibliographic Notes

A non-Poisson network appearing stochastically Poisson was first observed with log-normal shadowing by Brown (Brown 2000) and then independently by Błaszczyszyn and Karray (2012), both in a simulation setting. For the log-normal setting with power law path loss model, a mathematical proof was presented by Błaszczyszyn, Karray, & Keeler (2013), who later made further observations (Błaszczyszyn et al. 2015). These Poisson convergence results were greatly extended for other propagation effects and path loss models by Keeler, Ross, & Xia (2016a) with total variation bounds being derived. Under this model, it was observed that the strongest signals appear the most Poisson (Keeler et al. 2016b). For the power law path loss model, Poisson convergence was shown to be still possible with correlated log-normal shadowing, where shadowing is modeled with a Gaussian field (Ross & Schuhmacher 2017). Network equivalence property has been formulated for wireless context in Błaszczyszyn & Keeler (2013), but this useful property had been, however, observed previously in several different contexts; see Błaszczyszyn & Keeler (2015) for references.

Part II

SINR Analysis

5 Downlink SINR: Fundamental Results

In this chapter we focus on a fundamental quantity related to the network performance, known as the downlink *signal-to-interference-and-noise ratio* (SINR). Recall from Chapter 1 that we defined this quantity as the ratio between the power received by a mobile user as a useful signal, in the simplest scenario transmitted by one serving base station, to the total power of all unwanted signals received at the user terminal. The latter power is usually expressed as the sum of two terms, one representing the total sum of signal powers of nonserving base stations, called *interference*, and the other, called *noise*, comprising the total power of all signals whose sources are not represented explicitly in the point process model, such as background noise. The SINR is related to the information-theoretic capacity of a communication channel, and is referred to as SIR or SNR when only interference or noise is considered, respectively.

The SINR also characterizes the performance of real communication coding schemes. Consequently, it is an important building block for more complex studies of the performance of wireless networks and the quality of service that they offer.

5.1 General Considerations

We consider a network model with a collection of base stations Φ, assumed for the moment to be a general stationary point process. In this network, we consider a hypothetical user located at some location $y \in \mathbb{R}^2$. From each base station $\mathsf{x} \in \Phi$, the user receives a signal of power $S_{\mathsf{x}}(y)$, usually assumed to be a random function of the base station location x. Typically, $S_{\mathsf{x}}(y) = S_{\mathsf{x}}/\mathsf{L}_{\mathsf{x}}(y)$, where S_{x} is the power transmitted by x and $\mathsf{L}_x(y)$ is the propagation loss between x and the location y, as is represented in, for example, Eq. (4.1). In all cases, we assume that S_{x} are stationary marks of Φ.

The downlink SINR at y with respect to base station x in the ratio

$$\mathrm{SINR}(\mathsf{x}, y) \stackrel{\triangle}{=} \frac{S_{\mathsf{x}}(y)}{\sum\limits_{\mathsf{x}' \in \Phi, \mathsf{x}' \neq \mathsf{x}} S_{\mathsf{x}'}(y) + N}, \tag{5.1}$$

where $N \geq 0$ is an assumed constant or random variable representing the noise. $\mathrm{SINR}(\mathsf{x}, y)$ characterizes the performance of the wireless channel between $\mathsf{x} \in \Phi$ and the location y. The assumed stochastic nature of our network model, through both random propagation effects and base station locations, means that $\{\mathrm{SINR}(\mathsf{x}, y)\}_{\mathsf{x} \in \Phi}$ is a set of random fields indexed by y.

5.1.1 SINR distribution

We are usually interested in describing the probability that the SINR at a given location with respect to some base station exceeds some given threshold $\tau > 0$; $\mathbb{P}\{\text{SINR} > \tau\}$. There are two reasons for this:

- *Constant bit rate communication.* A user seeks to establish a communication channel achieving some given bit rate, given the available technology, which is possible only if the corresponding SINR exceeds some threshold τ. $\mathbb{P}\{\text{SINR} > \tau\}$ is then the *probability of the successful communication* at the required bit rate (also called *success*, *capture*, *coverage*, or *nonoutage* probability, depending on the specific scenario being considered). In a conventional cellular system like long-term evolution (LTE), the user needs to maintain a connection to the serving base station of sufficient quality to be able to reliably decode the lowest-rate transmissions sent to it by the base station. If the user has such a connection, it is said to be *covered* by the base station. It should be emphasized that though we shall mostly use the language of coverage in this book, every result on coverage probability is equally valid as a statement of success probability.
- *Variable bit rate communication.* The communication is established by the value of the SINR, with the bit rate equal to $\mathbf{R}(\text{SINR})$, where $\mathbf{R}(\cdot)$ is some (technology dependent) function relating the SINR to the available bit rate. Knowing the probability $\mathbb{P}\{\text{SINR} > \tau\}$ for all values of $\tau > 0$, namely, the distribution of the SINR, allows us to calculate various characteristics of the bit rate $\mathbf{R}(\text{SINR})$, such as expressing the *mean bit rate* in the channel

$$\mathbb{E}[\mathbf{R}(\text{SINR})] = \int_0^\infty \mathbb{P}\{\mathbf{R}(\text{SINR}) > \tau\}\,\mathrm{d}\tau = \int_0^\infty \mathbb{P}\{\text{SINR} > \mathbf{R}^{-1}(\tau)\}\,\mathrm{d}\tau\,,$$

where we have assumed that $\mathbf{R}(\cdot)$ has an inverse $\mathbf{R}^{-1}(\cdot)$.

5.1.2 Signal-to-total-interference-plus-noise ratio

It is often easier to work with the *signal-to-total-interference-plus-noise ratio* (STINR) defined as follows:

$$\text{STINR}(\mathsf{x}, y) \triangleq \frac{S_{\mathsf{x}}(y)}{\sum_{\mathsf{x}\in\Phi} S_{\mathsf{x}}(y) + N}\,, \tag{5.2}$$

where $\sum_{\mathsf{x}\in\Phi} S_{\mathsf{x}}(y)$ is the total received power from the entire network, which is known as the *total* or *aggregate interference*. The difference is in the denominator, where all the received powers, including that of the useful signal, are considered. The STINR is easily related to the SINR:

$$\text{STINR}(\mathsf{x}, y) = \frac{\text{SINR}(\mathsf{x}, y)}{1 + \text{SINR}(\mathsf{x}, y)}. \tag{5.3}$$

The STINR is a strictly increasing function of SINR on $[0, \infty)$ with the inverse

$$\text{SINR}(\mathsf{x}, y) = \frac{\text{STINR}(\mathsf{x}, y)}{1 - \text{STINR}(\mathsf{x}, y)}. \tag{5.4}$$

The advantage of working with the STINR is that it has the common denominator in Expression (5.2) for all $\mathsf{x} \in \Phi$, allowing for easy algebraic manipulation.

5.1.3 Choice of the base station

We detail some possible scenarios for the choice of the base station serving a given location y.

Typical base station

Consider the point process of base station locations, and focus on the base station located at the origin o under the Palm distribution of this point process. We consider a user at some location y that connects to this base station. Given the corresponding SINR is $\mathrm{SINR}(o, y)$, we define the (Palm) coverage probability $p_y(\tau)$ of the location y, at the SINR threshold τ, by this base station, as

$$p_y(\tau) = \mathbb{P}_o\{\, \mathrm{SINR}(o, y) > \tau \,\}. \tag{5.5}$$

Strongest base station

The user at location y is assumed to receive the signal from the base station offering the strongest received signal, which means

$$\mathrm{SINR}(y) \stackrel{\triangle}{=} \frac{\max_{\mathsf{x}\in\Phi} S_{\mathsf{x}}(y)}{N + \sum_{\mathsf{x}'\in\Phi} S_{\mathsf{x}'}(y) - \max_{\mathsf{x}\in\Phi} S_{\mathsf{x}}(y)}\,. \tag{5.6}$$

In terms of STINR, it is easy to see that the value of $\mathrm{SINR}(y)$ also corresponds to the maximal SINR value.

LEMMA 5.1.1 *We have*

$$\mathrm{SINR}(y) = \max_{\mathsf{x}\in\Phi} \mathrm{SINR}(\mathsf{x}, y)\,. \tag{5.7}$$

Proof By Eq. (5.4) and the fact that the function $t/(1-t)$ is strictly increasing on $(0, 1)$, which is verified by calculating its derivative, we have $\mathrm{SINR}(\mathsf{x}, y) \geq \mathrm{SINR}(\mathsf{x}', y)$ if and only if $\mathrm{STINR}(\mathsf{x}, y) \geq \mathrm{STINR}(\mathsf{x}', y)$. Hence the maximal STINR and SINR are attained at the same value of $\mathsf{x} \in \Phi$. By the common denominator of STINR, its maximal value is attained at $\mathsf{x} \in \Phi$ offering the strongest received power,

$$\max_{\mathsf{x}\in\Phi} \mathrm{STINR}(\mathsf{x}, y) = \frac{\max_{\mathsf{x}\in\Phi} S_{\mathsf{x}}(y)}{N + \sum_{\mathsf{x}'\in\Phi} S_{\mathsf{x}'}(y)}\,,$$

which completes the proof. □

For the coverage probability, we can assume without loss of generality that the receiver is located at the origin o. We define the coverage probability of the typical location by the strongest base station at the SINR threshold τ as

$$\mathcal{P}(\tau) = \mathbb{P}\{\, \mathrm{SINR}(o) > \tau \,\}. \tag{5.8}$$

The closest base station

For a user, the choice of the serving base station is sometimes made so as to minimize its distance to the receiver. This may correspond to the situation, for example, when the strongest received station is chosen with respect to the averaged propagation effects. We write $\mathsf{x}^* \in \Phi$ to denote the station closest to the origin; $|\mathsf{x}^*| \leq \min_{\mathsf{x}' \in \Phi} |\mathsf{x}'|$. The corresponding SINR is then $\mathrm{SINR}(\mathsf{x}^*, o)$. A generalization is possible by averaging over some part of the propagation effects, typically multipath fading, and looking for a base station that maximizes the received signal perturbed by the shadowing.

5.1.4 Simple and multiple coverage regime

For $\mathrm{SINR}(\mathsf{x}, y)$, we now make an important structural observation, which is of purely algebraic nature. Regardless of any probabilistic assumptions, the number of base stations that offer to a given location y an SINR value larger than τ cannot exceed the fixed (nonrandom) value $(1+\tau)/\tau$. We state this more formally.

LEMMA 5.1.2 *For all* $y \in \mathbb{R}^2$

$$\#\{\mathsf{x} \in \Phi : \mathrm{SINR}(\mathsf{x}, y) \geq \tau\} \leq \frac{1+\tau}{\tau}\,.$$

The inequality is strict provided $N > 0$ *or* $0 < \mathrm{SINR}(\mathsf{x}, y) < \tau$ *for some* $\mathsf{x} \in \Phi$.

Remark 5.1 The value $\tau = 1$ corresponds to a phase transition in the complexity of the SINR coverage analysis, owing to the combinatorial nature of the multiple-coverage problem. Indeed, Lemma 5.1.2 says that at any location a user can connect to at most one base station in the network when the SINR $\geq \tau \geq 1$. We call the regime $\tau \geq 1$ the *simple (SINR) coverage regime*. If $1/2 \leq \tau < 1$, then the user can connect to a maximum of two base stations in the network, and so on. This means that all the values $\tau < 1$ correspond to multiple-coverage regime. Assuming $\tau \geq 1$ can greatly simplify the coverage analysis. The bound in Lemma 5.1.2 is sometimes called *pole capacity*, and, because it is purely a consequence of the SINR definition and algebra, it is completely independent of any assumptions on base station locations or propagation models.

Proof For $t > 0$, the function $t/(1+t)$ is increasing in t. Then for all $\mathsf{x} \in \Phi$, $y \in \mathbb{R}^2$, by Eq. (5.3) $\mathrm{SINR}(\mathsf{x}, y) \geq \tau$ if and only if $\mathrm{STINR}(\mathsf{x}, y) \geq \tau' \stackrel{\triangle}{=} \tau/(1+\tau)$. Then we introduce the number

$$K' \stackrel{\triangle}{=} \#\{\mathsf{x} \in \Phi : \mathrm{SINR}(\mathsf{x}, y) > \tau\} = \#\{\mathsf{x} \in \Phi : \mathrm{STINR}(\mathsf{x}, y) > \tau'\}\,,$$

which leads to

$$\tau' K' \leq \sum_{\mathsf{x} \in \Phi} \mathrm{SINR}(\mathsf{x}, y) \mathbf{1}(\mathrm{STINR}(\mathsf{x}, y) > \tau') \leq 1\,,$$

where the last inequality is due to the common denominator. The inequality is strict provided $N > 0$ in Eq. (5.2) or there is $\mathsf{x} \in \Phi$ such that $0 < \mathrm{STINR}(\mathsf{x}, y) < \tau'$. Consequently $K' \leq 1/\tau' = (1+\tau)/\tau$. □

5.1.5 Coverage probability exchange formula in the simple regime

The exchange formula of Theorem 4.1.4 yields an important, general relation between the coverage probability of the typical location by the strongest base station $\mathcal{P}(\tau)$ and the coverage function by the typical base station $S_y(\tau)$, $y \in \mathbb{R}^2$, where both probabilities are defined in Section 5.1.3.

PROPOSITION 5.1.3 *For a general stationary network model, we have for* $\tau \geq 1$

$$\mathcal{P}(\tau) = \int_{\mathbb{R}^2} p_y(\tau)\, \mathrm{d}y\,. \tag{5.9}$$

Proof By setting $\chi(\mathsf{x}, y) = \mathbf{1}(\mathrm{SINR}(\mathsf{x}, y) > \tau)$ in Eq. (4.7), we have

$$\mathbb{E}[\mathbf{N}(\tau)] = \int_{\mathbb{R}^2} p_y(\tau)\, \mathrm{d}y\,,$$

where

$$\mathbf{N}(\tau) \stackrel{\triangle}{=} \#\{\mathsf{x} \in \Phi : \mathrm{SINR}(\mathsf{x}, o) > \tau\}$$

is the number of base station offering to the location o the SINR greater than τ. In the simple coverage regime, when $\tau \geq 1$, Lemma 5.1.2 yields $\mathbf{N}(\tau) = 0$ or 1, so

$$\mathbb{E}[\mathbf{N}(\tau)] = \mathbb{P}\{\mathbf{N}(\tau) \geq 1\} = \mathbb{P}\{\max_{x \in \Phi} \mathrm{SINR}(\mathsf{x}, o) > \tau\} = \mathcal{P}(\tau)\,.$$

□

5.1.6 Increasing model complexity

For a quantitative SINR analysis, we must specify general model assumptions, which naturally affect the complexity of the model and the resulting expressions for the distribution of the SINR. For example, arguably one of the simplest models possible is to have a single tier, where all the base stations have the same parameters, represented by a single homogeneous Poisson process. Adding more tiers, with presumably different networks parameters, increases the model complexity.

5.2 Basic Results for Poisson Network with Singular Path Loss Model

Assuming the singular path loss model in Eq. (4.2) and building upon results of Section 4.2, we present several closed or semiclosed form expressions for distribution of the downlink SINR in a homogeneous Poisson network model.

5.2.1 The singular path loss model

We assume that Φ is a homogeneous Poisson process of intensity λ and we use the simple path loss function in Eq. (4.2), which has a singularity at the origin. More precisely, we assume that the power received at y from a base station $\mathsf{x} \in \Phi$ can be expressed as

$$S_{\mathsf{x}}(y) = \frac{1}{\mathsf{L}_{\mathsf{x}}(y)} = \frac{H_{\mathsf{x}}(y-\mathsf{x})}{(\kappa|\mathsf{x}-y|)^{\alpha}}\,, \tag{5.10}$$

where $\kappa > 0$ is the path loss constant and $\alpha > 0$ is the path loss exponent, and the random propagation field $H_{\mathsf{x}}(\cdot)$, given Φ, is independent across all $\mathsf{x} \in \Phi$ and identically distributed. We consider the SINR field (Eq. [5.1]) by assuming a constant or random noise $N \geq 0$ such that it is always independent of Φ, H_{x} and $\mathsf{x} \in \Phi$.

Remark 5.2 In Expression (5.10), we tacitly assumed that all base stations transmit with unit power. We see in Section 4.2.4 we show a way how this assumption can be easily relaxed to include transmission powers or antenna gains. There are in fact two immediate ways to include, for example, the transmission power.

- *Constant transmission power.* The analysis of the SINR fields when all base stations transmit with the same constant power P_0, meaning that for each base station $\mathsf{x} \in \Phi$, the received power is $P_0 \times S_{\mathsf{x}}$, which in the model is equivalent to having unit-transmitted power and noise N/P_0.
- *Random independent transmission powers.* When base stations use variable transmission powers P_{x} that can be assumed V, independent across all $\mathsf{x} \in \Phi$, then the SINR model can be reduced to this with the random propagation fields $P_{\mathsf{x}} \times H_{\mathsf{x}}(\cdot)$.

The above cases do not cover all power control mechanisms.

5.2.2 SINR with respect to the typical station

We again assume a user is located at some location $y \in \mathbb{R}^2$, receiving a useful signal from the base station located at the origin under Palm distribution of the base stations. We consider the (Palm) coverage probability $S_y(\tau)$ of the location y by the SINR of the typical base station at o defined in Eq. (5.5). Due to the isotropy of the model, S_c depends only on the distance $r = |y|$ between the base station and the location y, so, with a slight abuse of notation, we can write $S_r(\tau)$ instead of $S_y(\tau)$.

PROPOSITION 5.2.1 *Assume exponential marginal distribution of the propagation effects $H = H_{\mathsf{x}}(y)$ with mean $1/\mu$, corresponding to Rayleigh fading. Then*

$$p_r(\tau) = \mathcal{L}_N(\mu\tau(\kappa r)^{\alpha}) \times \exp\left\{-\lambda\pi r^2 \tau^{\delta} C(\delta)\right\}, \tag{5.11}$$

where $\mathcal{L}_N(\xi) = \mathbb{E}[\mathrm{e}^{-\xi N}]$ is the Laplace transform of the noise N, and

$$C(\delta) \stackrel{\triangle}{=} \Gamma(\delta+1)\Gamma(1-\delta) = \frac{\delta\pi}{\sin(\delta\pi)},$$

recalling that $\delta = 2/\alpha$ is the characteristic exponent for $\mathbb{R}^2$.

Remark 5.3 We note that

$$C(\delta) = D(\delta)\mathbb{E}[H^{\delta}]/\pi\,,$$

where $D(\delta) = \pi\Gamma(1-\delta)$ is the spatial contention factor that already appeared in Expression (4.14) of the Laplace transform of the total recevied power I' with the simple

path loss function, and $\mathbb{E}[H^\delta] = \Gamma(\delta+1) = \delta\Gamma(\delta)$, corresponding to the exponential assumption on H.

Proof of Proposition 5.2.1 By the independence of N and $H_o(0)$ as well as the exponential distribution of the latter, we have

$$\begin{aligned}
&\mathbb{P}_o\{\,\mathrm{SINR}(o,y) > \tau\,\} \\
&= \mathbb{P}_o\left\{ H_o(y) > (\kappa r)^\alpha \tau \left[(N + \sum_{\mathsf{x}\in\Phi,\mathsf{x}\neq o} H_{\mathsf{x}}(y-x)/(\kappa|y-x|)^\alpha \right] \right\} \\
&= \mathbb{E}_N\Big[\exp\{-\mu\tau(\kappa r)^\alpha N\}\Big] \times \mathbb{E}_o\left[\exp\left\{ -\mu\tau(\kappa r)^\alpha \sum_{\mathsf{x}\in\Phi,\mathsf{x}\neq o} H_{\mathsf{x}}(y-x)/(\kappa|y-x|)^\alpha \right\} \right] \\
&= \mathcal{L}_N(\mu\tau(\kappa r)^\alpha) \times \mathcal{L}_{I'(y)}(-\mu\tau(\kappa r)^\alpha)\,,
\end{aligned}$$

where $\mathcal{L}_{I'(y)}$ is the Laplace transform of the total received power as in Section 4.1.4, and where in the last equality we have used the fact that under $\mathbb{P}_o$ points and marks of $\Phi \setminus o$ have the stationary distribution, due to Slivnyak's theorem. By the stationarity, $\mathcal{L}_{I'(y)} = \mathcal{L}_{I'(0)}$, and then the result follows from Eq. (4.14). □

For a general distribution of $H_{\mathsf{x}}(\cdot)$, the evaluation of $S_r(\tau)$ from the Laplace transform $\mathcal{L}_{I'(o)}$ and $\mathcal{L}_N$ is not so straightforward. An integral formula, based on the Plancherel-Parseval theorem, has been proposed when H has a square integrable density (Baccelli, Błaszczyszyn, & Mühlethaler 2009). This approach, however, does not apply to the no-fading case $H = 1$. In this case, another approach is needed (Blaszczyszyn & Mühlethaler 2015), which uses a numerical method based on a Bromwich contour integral (Abate & Whitt 1995).

5.2.3 SINR with respect to the strongest station in the simple coverage regime

We consider the scenario when the receiver is located at the origin o and the coverage probability $\mathcal{P}(\tau)$ dictated by the strongest SINR value, as defined by Eq. (5.8). In the simple coverage regime ($\tau \geq 1$), we use Proposition 5.1.3 and Proposition 5.2.1 to explicitly evaluate $\mathcal{P}(\tau)$.

Lemma 4.2.2 presented the constant

$$a \stackrel{\triangle}{=} \frac{\lambda\pi\mathbb{E}(H^\delta)}{\kappa^2}\,, \tag{5.12}$$

which appears often in the rest of this chapter and in Chapter 6.

The next result requires the gamma function, which is defined in Section C.2.7.

PROPOSITION 5.2.2 *Under assumptions of Proposition 5.2.1, for $\tau \geq 1$ we have*

$$\mathcal{P}(\tau) = \frac{2\tau^{-\delta}}{\Gamma(1+\delta)} \int_0^\infty u\mathrm{e}^{-u^2\Gamma(1-\delta)} \mathcal{L}_N(a^{-1/\delta}u^\alpha)\,\mathrm{d}u\,, \tag{5.13}$$

where Γ *is the gamma function. For Rayleigh fading, that is, exponential propagation effects H, the propagation constant is*

$$a = \lambda\pi\Gamma(\delta+1)/\kappa^2\,.$$

Furthermore, for the zero-noise ($N = 0$) *case, the above expression reduces to*

$$\mathcal{P}(\tau) = \frac{\tau^{-\delta}}{C(\delta)}\,, \tag{5.14}$$

where $C(\delta) = \Gamma(1+\delta)\Gamma(1-\delta)$.

Remark 5.4 (*Relaxing exponential propagation effects)* The probability $\mathcal{P}(\tau)$ is a functional of the point process Θ. From Lemma 4.2.2 and Remark 4.3, the distribution of this process depends on the marginal distribution of the propagation effects H only through the moment $\mathbb{E}[H^\delta]$. Consequently, the Expression (5.13) remains valid for an arbitrary distribution of H for which $\mathbb{E}[H^\delta] < \infty$, with the constant a, defined in Eq. (4.31), accounting for this moment of H. When $N = 0$, the model depends only on the path loss exponent α and none of the other physical parameters λ, $\mathbb{E}[H^\delta]$, or κ. Unfortunately, there is no such invariance for the typical station coverage probability $p_r(\tau)$, which cannot be expressed as a functional of Θ.

5.2.4 Coverage probability by the closest base station

Based on the material covered in Section 5.1.3, we denote by

$$\tilde{\mathcal{P}}(\tau) = \mathbb{P}\{\,\mathrm{SINR}(\mathbf{x}^*, o) > \tau\,\}$$

the SINR coverage probability by the closest base station to the origin, which is our typical location. In this case, $\tilde{\mathcal{P}}(\tau)$ is not a functional solely of the point process Θ, so its expression may depend on the particular assumption on the distribution of H.

PROPOSITION 5.2.3 *Under assumptions of Proposition 5.2.1, in particular exponential propagation effects H, the coverage probability is*

$$\tilde{\mathcal{P}}(\tau) = \delta \int_0^\infty t^{\delta-1}\mathcal{L}_N(t\tau a^{-1/\delta})\mathrm{e}^{-t^\delta} \exp\left(-t^\delta\tau\frac{\delta}{(1-\delta)}\,{}_2F_1(1,1-\delta;2-\delta;-\tau)\right)\mathrm{d}t\,, \tag{5.15}$$

where

$$a = \frac{\lambda\pi}{\kappa^2}$$

and ${}_2F_1(a,b;c;z)$ *is the (Gauss) hypergeometric function given by equation 5.11 in the Digital Library of Mathematical Functions* (National Institute of Standards and Technology 2012).

Proof Assuming exponential H and conditioning on the distance to $\mathbf{x}^*$, whose probability density function is $f_{|\mathbf{x}^*|}(r) = 2\lambda\pi\, r\mathrm{e}^{-\lambda\pi r^2}$ (recall Example 4.15), gives

$$\tilde{P}(\tau) = \int_0^\infty \mathbb{P}\Big\{ H_{\mathsf{x}^*} \geq (\kappa r)^\alpha \tau (N + I_{(r,\infty)}) \Big\} f_{|\mathsf{x}^*|}(r)\, \mathrm{d}r$$

$$= \int_0^\infty \mathcal{L}_N(\tau(\kappa r)^\alpha) \times \mathcal{L}_{I_{(r,\infty)}}(\tau(\kappa r)^\alpha) f_{|\mathsf{x}^*|}(r)\, dr\,, \tag{5.16}$$

where $I_{(r,\infty)}$ is the random variable representing conditional interference given the closest, serving station is located at a distance r. By the Markov property of the Poisson process, the Laplace transform of this conditional interference can be explicitly evaluated:

$$\mathcal{L}_{I_{(r,\infty)}}(\xi) = \exp\left(-2\pi\lambda \int_r^\infty u\Big(1 - \mathcal{L}_H(\xi(\kappa u)^{-\alpha})\Big)\, \mathrm{d}u \right) \tag{5.17}$$

$$= \exp\left(\frac{-a\xi r^\delta}{r} \frac{\delta}{(1-\delta)}\ {}_2F_1(1, 1-\delta; 2-\delta; -\xi/r) \right). \tag{5.18}$$

Plugging into integral Eq. (5.16) completes the proof. □

Remark 5.5 Expression (5.15) can be easily evaluated numerically. Setting $N = 0$ yields a compact analytic solution

$$\tilde{\mathcal{P}}(\tau) = \frac{1}{{}_2F_1(1, -\delta; 1-\delta; -\tau)}\,, \tag{5.19}$$

through the identity

$${}_2F_1(1, -\delta; 1-\delta; -\tau) = 1 + \frac{\tau\delta}{(1-\delta)} {}_2F_1(1, 1-\delta; 2-\delta; -\tau)\,,$$

which can be derived through ${}_2F_1$ identities such as those shown in Gradshteyn & Ryzhik (2007, equation 9.137.4). We see again that the coverage probability does not depend on the base station density λ.

5.2.5 Alternative derivation of coverage probability by the closest base station

Observe that the scenario in Proposition 5.2.3 is equivalent to the case where the user of interest receives from the *nearest* base station, and there is iid Rayleigh fading on all links to the user from all base stations in the tier. In this section, we arrive at the same result as Proposition 5.2.3, but this time using the Campbell-Mecke theorem (see Eq. [3.32]).[1]

Notation: x^* is the nearest base station to the user (who is assumed located at the origin), Φ is the Poisson point process describing the locations of the base stations of the single tier, $Y_{\mathsf{x}} \stackrel{\triangle}{=} H_{\mathsf{x}} \ell(\|\mathsf{x}\|)$ is the received power at the user location from the base

[1] This section is optional and may be skipped on a first reading. We shall later make greater use of the Campbell-Mecke theorem in deriving Eq. (A.2) for the case where the user of interest receives from the *strongest* base station and there is arbitrary iid fading on all links to the user from all base stations in the tier.

station at $\mathsf{x} \in \mathbb{R}^2$, $\mathrm{SINR}_{\mathsf{x}}$ is the SINR at the user location when receiving from the base station at $\mathsf{x} \in \mathbb{R}^2$, $\tilde{\Phi} = \{\|\mathsf{x}\| : \mathsf{x} \in \Phi\}$. We begin with the following sequence of derivations:

$$\begin{aligned}
\tilde{\mathcal{P}}_c(\tau) &= \mathbb{E} \sum_{\mathsf{x}\in\Phi} 1\{\mathsf{x}^* = \mathsf{x}, \mathrm{SINR}_{\mathsf{x}} > \tau\} = \mathbb{E}\left[\mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} 1\{\mathsf{x}^* = \mathsf{x}\}\, 1\{\mathrm{SINR}_{\mathsf{x}} > \tau\} \;\middle|\; \Phi\right]\right] \\
&= \mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} 1\{\mathsf{x}^* = \mathsf{x}\}\, \mathbb{P}\left\{H_{\mathsf{x}} > \tau(\kappa\|\mathsf{x}\|)^{\alpha}\left(N + \sum_{\mathsf{x}'\in\Phi\setminus\{\mathsf{x}\}} \frac{H_{\mathsf{x}'}}{(\kappa\|\mathsf{x}'\|)^{\alpha}}\right) \;\middle|\; \Phi\right\}\right] \\
&= \mathbb{E}\left[\sum_{\mathsf{x}\in\Phi} 1\{\mathsf{x}^* = \mathsf{x}\} \exp\left[-\tau N(\kappa\|\mathsf{x}\|)^{\alpha}\right] \prod_{\mathsf{x}'\in\Phi\setminus\{\mathsf{x}\}} \mathcal{L}_H\left(\tau\frac{\|\mathsf{x}\|^{\alpha}}{\|\mathsf{x}'\|^{\alpha}}\right)\right] \\
&= \mathbb{E}\left[\sum_{r\in\tilde{\Phi}} \exp\left[-\tau N(\kappa r)^{\alpha}\right] \prod_{r'\in\tilde{\Phi}\setminus\{r\}} 1(r' > r)\mathcal{L}_H\left(\tau\left(\frac{r}{r'}\right)^{\alpha}\right)\right] \\
&= \mathbb{E}\left[\sum_{r\in\tilde{\Phi}} h(r, \tilde{\Phi}\setminus\{r\})\right] = \int_0^{\infty} \mathbb{E}[h(u,\tilde{\Phi})]\tilde{\lambda}(u)\,\mathrm{d}u, \qquad (5.20)
\end{aligned}$$

where Eq. (5.20) comes from the Campbell-Mecke theorem Poisson point process (see Eq. [3.32]), and for any print measure ψ on $\mathbb{R}^+$, we define

$$h(r,\psi) = \exp\left[-\tau N(\kappa r)^{\alpha}\right] \prod_{r'\in\psi} 1_{(r'>r)}\mathcal{L}_H\left(\tau\left(\frac{r}{r'}\right)^{\alpha}\right),$$

$\mathcal{L}_H(s) = \mathbb{E}\exp(-sH)$ where H has the common exponential distribution of the iid random variables $H_{\mathsf{x}'}$, $\mathsf{x}' \in \Phi\setminus\{\mathsf{x}\}$, and $\tilde{\lambda}(x) = 2\pi\lambda x$, $x \geq 0$. From equation 3.30 we have

$$\begin{aligned}
\mathbb{E}[h(u,\tilde{\Phi})] &= \exp\left[-\tau N(\kappa u)^{\alpha}\right]\mathbb{E}\left[\prod_{v\in\tilde{\Phi}} 1(v > u)\mathcal{L}_H\left(\tau\left(\frac{u}{v}\right)^{\alpha}\right)\right] \\
&= \exp\left[-\tau N(\kappa u)^{\alpha}\right]\exp\left\{-\int_0^u \tilde{\lambda}(v)\,\mathrm{d}v - \int_u^{\infty}\left[1 - \mathcal{L}_H\left(\tau\left(\frac{u}{v}\right)^{\alpha}\right)\right]\tilde{\lambda}(v)\,\mathrm{d}v\right\} \\
&= \exp\left[-\tau N(\kappa u)^{\alpha}\right]\exp(-\pi\lambda u^2)\exp\left\{-\int_u^{\infty}\tilde{\lambda}(v)\left[1 - \mathcal{L}_H\left(\tau\left(\frac{u}{v}\right)^{\alpha}\right)\right]\right\}\mathrm{d}v.
\end{aligned}$$

The result of Proposition 5.2.3 follows immediately by (5.16) substituting Eq. (5.17) in Eq. (5.20).

5.2.6 Coverage probability with shadowing separated from fading

Proposition 5.2.3 can be extended. We suppose that the received power S_{x} from each station $x \in \Phi$ given by Expression (5.10) accounts for two types of propagation effects

$$S_{\mathsf{x}}(y) = \frac{\hat{H}_{\mathsf{x}}(y-\mathsf{x})H_{\mathsf{x}}(y-\mathsf{x})}{(\kappa|\mathsf{x}-y|)^{\alpha}}\,, \tag{5.21}$$

with fields $\hat{H}_{\mathsf{x}}(\cdot)$ and $H_{\mathsf{x}}(\cdot)$ being independent of each other, and both independent marks of Φ. We now suppose that the serving base station is the one that maximizes the received power accounted for $\hat{H}_{\mathsf{x}}$ (call it shadowing) but not H_{x} (call it fading). We denote by x^{*s} the location of this station, that is,

$$S_{\mathsf{x}^{*s}}(o)/H_{\mathsf{x}^{*s}}(o-\mathsf{x}^{*s}) \geq \max_{x\in\Phi} S_{\mathsf{x}}(o)H_{x}(o-\mathsf{x})\,.$$

With respect to the base station at x^{*s}, we consider the coverage probability

$$\breve{\mathcal{P}}(\tau) = \mathbb{P}\{\,\mathrm{SINR}(\mathsf{x}^{*s}, o) > \tau\,\}\,.$$

PROPOSITION 5.2.4 *Assume a general marginal distribution of the shadowing $\hat{H}$ with $\mathbb{E}[\hat{H}^{\delta}] < \infty$ and assume an exponential, with mean $1/\mu$, marginal distribution of fading H. Then $\breve{\mathcal{P}}(\tau) = \tilde{\mathcal{P}}(\tau)$ given by Expression (5.15) with*

$$a = \frac{\lambda\pi\mathbb{E}[\hat{H}^{\delta}]}{\kappa^2}\,.$$

Proof Note that $\breve{\mathcal{P}}(\tau)$ can be seen as a functional of the process Θ with shadowing, which is then further independently marked by independent exponential fading variables. Lemma 4.2.2 and Remark 4.3 allow us to reduce the problem to the coverage probability by the closest station (without shadowing) with appropriate modification of the constant a. □

In the zero-noise case, the dependence of $\breve{\mathcal{P}}(\tau)$ on a and, consequently, the shadowing $\hat{H}$ disappears.

5.3 Multiple Coverage in Poisson Network with Singular Path Loss Model

We extend the results of Section 5.2.3 on the coverage probability of the typical location by the strongest base station to the multiple coverage regime. In Section 5.1.4, we see when the SINR threshold $\tau < 1$ that there might be several base stations offering sufficiently high SINR values. The goal is to characterize the distribution of the strongest SINR in the whole domain $\tau > 0$. Throughout this section we keep assuming the singular path loss model as in Section 5.2.1.

5.3.1 Coverage number and k-coverage probability

DEFINITION 5.3.1 (Coverage number) *The* coverage number *of the typical location is the number of base stations that (the user at) the typical location can connect to at SINR level τ, namely*

$$N(\tau) = \sum_{\mathsf{x}\in\Phi} \mathbf{1}[\mathrm{SINR}(\mathsf{x}, o) > \tau]\,. \tag{5.22}$$

DEFINITION 5.3.2 (k-Coverage probability) *The k-*coverage probability *is*

$$\mathcal{P}^{(k)}(\tau) = \mathbb{P}\{\,N(\tau) \geq k\,\}\,, \tag{5.23}$$

which is the probability of the typical location being connected to or covered by at least k base stations.

Naturally, $\mathcal{P}^{(1)}(\tau) = \mathcal{P}(\tau)$ already considered in the regime $\tau \geq 1$. To treat the $\tau > 0$ case, we need to introduce the combinatorial concept of a symmetric sum.

DEFINITION 5.3.3 (Symmetric sum) *For $n \geq 1$, the nth* symmetric sum *is*

$$\mathcal{S}_n(\tau) \stackrel{\triangle}{=} \mathbb{E}\left[\sum_{\substack{\mathsf{x}_1,\ldots,\mathsf{x}_n \in \Phi \\ distinct}} \mathbf{1}\{\,\mathrm{SINR}(\mathsf{x}_j, o) > \tau,\, j = 1,\ldots,n\,\}\right]. \tag{5.24}$$

We set $\mathcal{S}_0(\tau) \equiv 1$, and note that $\mathcal{S}_n(\tau)$ is the expected number of ways that the typical user can connect to n base stations, each with an SINR greater than τ. We note that Lemma 5.1.2 influences $\mathcal{S}_n$ by setting its values to zero if n is above a certain value.

LEMMA 5.3.4 *For $n \geq 1$, $\mathcal{S}_n(\tau) = 0$ whenever $\tau \geq 1/(n-1)$.*

The motivation for introducing $\mathcal{S}_n$ is the direct connection with $\mathbf{N}(\tau)$ and $\mathcal{P}^{(k)}$.

LEMMA 5.3.5 *For $k \geq 1$*

$$\mathcal{P}^{(k)}(\tau) = \sum_{n=k}^{\lceil 1/\tau \rceil} (-1)^{n-k} \binom{n-1}{k-1} \mathcal{S}_n(\tau)\,, \tag{5.25}$$

$$\mathbb{P}\{\boldsymbol{N}(\tau) = k\} = \sum_{n=k}^{\lceil 1/\tau \rceil} (-1)^{n-k} \binom{n}{k} \mathcal{S}_n(\tau)\,, \tag{5.26}$$

$$\mathbb{E}[z^{\boldsymbol{N}(\tau)}] = \sum_{n=0}^{\lceil 1/\tau \rceil} (z-1)^n \mathcal{S}_n(\tau)\,, \quad z \in [0,1]\,, \tag{5.27}$$

$$\mathbb{E}[\boldsymbol{N}(\tau)] = \mathcal{S}_1(\tau)\,, \tag{5.28}$$

where for real t, $\lceil t \rceil$ denotes ceiling of t (the smallest integer not less than t).

These combinatorial identities are modifications of the famous inclusion-exclusion principle (for example, see Feller (Feller 1968, Section IV.3 and IV.5)), but they have been simplified here with Lemma 5.3.4 so they have only finite sums. The Schuette-Nesbitt formula also links the distribution of $\mathbf{N}$ and the symmetric sums $\mathcal{S}_n$.

To express the symmetric sums $\mathcal{S}_n$ and consequently k-coverage probabilities, we must introduce two families of special functions. The first family of functions, the simpler of the two, allows one to capture the impact of the noise N. In the zero-noise ($N = 0$) case, it reduces to constants involving gamma functions. The second family of functions captures the combinatorial complexity of the k-coverage problem.

For all $n \geq 1, 2, \ldots$ and $t \geq 0$, we define the first integral

$$\mathcal{I}_{n,\delta}(t) \stackrel{\triangle}{=} \frac{2\delta^{n-1}}{(C(\delta))^n (n-1)!} \int_0^\infty u^{2n-1} \mathrm{e}^{-u^2 - tu^\alpha \Gamma(1-\delta)^{-1/\delta}} \mathrm{d}u, \tag{5.29}$$

where

$$C(\delta) = \frac{\delta\pi}{\sin(\delta\pi)} = \Gamma(1+\delta)\Gamma(1-\delta). \tag{5.30}$$

Note that for $t = 0$, corresponding to the zero-noise case, (5.29) boils down to the gamma function so

$$\mathcal{I}_{n,\delta}(0) = \frac{\delta^{n-1}}{(C(\delta))^n}. \tag{5.31}$$

For all $n = 1, 2, \ldots$ and $t \in [0, 1]$ define

$$\bar{\mathcal{J}}_{n,\delta}(t) \stackrel{\triangle}{=} (1/n + t) \int\limits_{[0,1]^{n-1}} \frac{\prod_{i=1}^{n-1} v_i^{i(\delta+1)-1}(1-v_i)^\delta}{\prod_{i=1}^{n}(t+\eta_i)} \, \mathrm{d}v_1 \ldots \mathrm{d}v_{n-1} \tag{5.32}$$

where η_i's are the following functions of v_i's

$$\begin{cases} \eta_1 &= v_1 v_2 \ldots v_{n-1} \\ \eta_2 &= (1-v_1) v_2 \ldots v_{n-1} \\ \eta_3 &= (1-v_2) v_3 \ldots v_{n-1} \\ &\ldots \\ \eta_n &= 1 - v_{n-1}. \end{cases} \tag{5.33}$$

Remark 5.6 For $n = 1$, $\mathcal{J}_{1,\delta}(t) = 1$. For $n = 2$, a closed-form solution exists

$$\mathcal{J}_{2,\delta}(x) = \frac{1}{x} B(\delta+1, \delta+1) \; {}_2F_1(1, \delta+1; 2(\delta+1); -1/x),$$

where B is the beta function (see Section C.2.8 or equation 5.12.1 in the Digital Library of Mathematical Functions [National Institute of Standards and Technology 2012]) and ${}_2F_1$ is the hypergeometric function. For higher n, numerical integration techniques are needed to evaluate these functions (Keeler 2014).

For $0 < \tau < 1/(n-1)$, we introduce the shorthand

$$\bar{\tau}_n \stackrel{\triangle}{=} \bar{\tau}_n(\tau) = \frac{\tau}{1-(n-1)\tau}, \tag{5.34}$$

We can present now expressions for the symmetric sums $\mathcal{S}_n$.

LEMMA 5.3.6 *For the homogeneous Poisson network Φ with path loss $\ell(x) = (\kappa\|x\|)^{-\alpha}$ and iid random fading marks $\{H_{\mathsf{x}}\}_{\mathsf{x}\in\Phi}$, the symmetric sum $\mathcal{S}_n$ is given by*

$$\mathcal{S}_n(\tau) = \bar{\tau}_n^{-n\delta} \mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})] \bar{\mathcal{J}}_{n,\delta}(\bar{\tau}_n) \tag{5.35}$$

for $0 < \tau < 1/(n-1)$ and $\mathcal{S}_n(\tau) = 0$ otherwise, where the expectation in the right-hand side of Eq. (5.35) is with respect to the distribution of the noise N and where $a = \pi\lambda\mathbb{E}(H^\delta)/\kappa^2$, $\bar{\tau}_n$ is given by Eq. (5.34) and $\bar{\mathcal{J}}_{n,\delta}$ is given by Eq. (5.32).

Remark 5.7 We note that $\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})]$ can be expressed in terms of the Laplace transform $\mathcal{L}_N$ of the noise N. By Eq. (5.29), we have

$$\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})] = \frac{2\delta^{n-1}}{(C(\delta))^n(n-1)!}\int_0^\infty u^{2n-1}e^{-u^2} \times \mathcal{L}_N(a^{-1/\delta}u^\alpha\Gamma(1-\delta)^{-1/\delta})\,\mathrm{d}u\,. \tag{5.36}$$

The proof of Lemma 5.3.6 is quite long and appears Appendix A. Lemma 5.3.6 coupled with Lemma 5.3.5 and Lemma 5.3.4 gives an expression for the k-coverage probability.

COROLLARY 5.3.7 *Under the assumptions of Lemma 5.3.6,*

$$\mathcal{P}^{(k)}(\tau) = \sum_{n=k}^{\lceil 1/\tau\rceil}(-1)^{n-k}\binom{n-1}{k-1}\bar{\tau}_n^{-n\delta}\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})]\bar{\mathcal{J}}_{n,\delta}(\bar{\tau}_n)\,. \tag{5.37}$$

In particular for $k = 1$ and $\tau \geq 1$

$$\mathcal{P}^{(1)}(\tau) = \tau^{-\delta}\mathbb{E}[\mathcal{I}_{1,\alpha}(Na^{-1/\delta})]. \tag{5.38}$$

Note that the special case Eq. (5.38) of $k = 1$ with $\tau \geq 1$ corresponds to the single-tier network result in Proposition 5.13. We bring attention to the convenient ability to factor out the noise in Expressions (5.35) and (5.37). It is reminiscent of factoring out the noise Laplace transform in the distribution of the SINR of the typical station (see Eq. [5.11]) under Rayleigh fading, an assumption that is not required, however, in our present setting.

Example 5.1 *Multitier network with open access.* We consider the independent Poisson multitier network model introduced in Example 4.18, where base stations of tier $\{1,\ldots,J\}$ are modeled by independent Poisson point processes of intensity λ_i, respectively. As in Example 4.20 we assume that base stations of tier $j \in \{1,\ldots,J\}$ transmit with power S_j. Suppose that the user at the typical location is seeking a base station *in all tiers* offering it an SINR value greater than τ. The probability that there are at least k such stations is equal to the k-coverage probability $\mathcal{P}^{(k)}(\tau)$ given by Eq. (5.37) with $a = \sum_{j=1}^J a_j$ and a_j given by Eq. (4.47). To calculate this probability, we can consider the equivalent single-tier network representation introduced in Example 4.22 and apply directly the result of Corollary 5.3.7 with the aforementioned constant a.

Example 5.2 *Multitier network with closed access.* We continue with the independent Poisson multitier network model with different transmitted powers of Examples 4.18 and 4.20. We suppose now that the user at the typical location is seeking a base station offering it an SINR value greater than τ, but uniquely in some *subset of tiers* $\mathbb{J} \subset \{1,\ldots,J\}$. We assume that the base stations in the other tiers, that is tiers $\{1,2,\ldots,J\}\backslash\mathbb{J}$, act as interferers. Letting $a = \sum_{j=1}^J a_j$ with a_j given by Eq. (4.47), as in the previous example, characterizes the equivalent single-tier network representation. We denote also by $a_{\mathbb{J}} := \sum_{j\in\mathbb{J}} a_j$. This latter constant characterizes the equivalent single-tier

representation of tiers $\mathbb{J}$. The probability that there are at least k stations *in tiers* $\mathbb{J}$ with SINR greater than τ, called k-coverage probability with access opened only to $\mathbb{J}$ and closed to other tiers, is equal to

$$\mathcal{P}_{\mathbb{J}}^{(k)}(\tau) = \sum_{n=k}^{\lceil 1/\tau \rceil} (-1)^{n-k} \left(\frac{a_{\mathbb{J}}}{a}\right)^n \binom{n-1}{k-1} \bar{\tau}_n^{-n\delta} \mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})] \bar{\mathcal{J}}_{n,\delta}(\bar{\tau}_n)\,. \tag{5.39}$$

In particular for $k = 1$ and $\tau \geq 1$ (single coverage regime)

$$\mathcal{P}^{(1)}(\tau) = \tau^{-\delta} \frac{a_{\mathbb{J}}}{a} \mathbb{E}[\mathcal{I}_{1,\delta}(Na^{-1/\delta})]. \tag{5.40}$$

We note that every term n, ($k \leq n \leq \lceil 1/\tau \rceil$) of Eq. (5.39) is equal to the same term of Eq. (5.37) multiplied by the factor $(\frac{a_{\mathbb{J}}}{a})^n$. This factor corresponds precisely to the probability that n given received signals come from tiers in $\mathbb{J}$ in the equivalent homogeneous network model with independently introduced tiers as explained in Example 4.22. With this observation, we can develop Expression (5.39). However, we shall prove it more directly by again using Corollary 5.3.7, separating the total interference caused by stations in $\mathbb{J}$ from the total interference caused by tiers $\{1, \ldots, J\} \setminus \mathbb{J}$, which we denote by I'. We can then safely treat this interference I' as an additional noise term. Using this approach and the independence of the network tiers we can account for closed-access model constraints by replacing $\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})]$ in Eq. (5.37) by $\mathbb{E}[\mathcal{I}_{n,\delta}((N+I')a_{\mathbb{J}}^{-1/\delta})]$. The independence of N and I and Expressions (5.29) and (4.14) lead to

$$\begin{aligned}
&\mathbb{E}[\mathcal{I}_{n,\delta}(N + I')a_{\mathbb{J}}^{-1/\delta})] \\
&= \frac{2\delta^{n-1}}{(C(\delta))^n (n-1)!} \int_0^\infty u^{2n-1} \mathrm{e}^{-u^2} \mathcal{L}_N(a_{\mathbb{J}}^{-1/\delta} u^\alpha \Gamma(1-\delta)^{-1/\delta}) \mathcal{L}_{I'}(a_{\mathbb{J}}^{-1/\delta} u^\alpha \Gamma(1-\delta)^{-1/\delta}) \mathrm{d}u \\
&= \frac{2\delta^{n-1}}{(C(\delta))^n (n-1)!} \int_0^\infty u^{2n-1} \mathrm{e}^{-u^2 a/a_{\mathbb{J}}} \mathcal{L}_N(a_{\mathbb{J}}^{-1/\delta} u^\alpha \Gamma(1-\delta)^{-1/\delta})\, \mathrm{d}u \\
&= \frac{2\delta^{n-1} \left(\frac{a_{\mathbb{J}}}{a}\right)^n}{(C(\delta))^n (n-1)!} \int_0^\infty u^{2n-1} \mathrm{e}^{-u^2} \mathcal{L}_N(a^{-1/\delta} u^\alpha \Gamma(1-\delta)^{-1/\delta})\, \mathrm{d}u \\
&= \left(\frac{a_{\mathbb{J}}}{a}\right)^n \mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})],
\end{aligned}$$

with the last equation following by appropriate substitution in the integral.

Example 5.3 *Cellular network model with device-to-device communications.* The multitier network model can be adapted to account for device-to-device communications, that is, the possibility for users to communicate directly with each other. Let Φ_{B} and Φ_{D} be two independent homogeneous Poisson processes with densities λ_{B} and λ_{D}, respectively, representing the locations of base stations and users, called in this context devices. For every base station $\mathsf{x} \in \Phi$, we let H_i be a positive random variable for the general propagation effects perturbing the base-station-to-device signal. Similarly, for

every device $\mathsf{x} \in \Phi_{\mathrm{D}}$, let E_i be an exponential random variable with unit mean for Rayleigh fading experienced by the device-to-device signal. Let $\tau_{\mathrm{B}} > 0$ be the SIR threshold of base-station-to-device transmission. We define the base station coverage probability as the tail distribution of the maximum SIR value

$$\mathcal{P}_{\mathrm{B}}(\tau_{\mathrm{B}}) := \mathbb{P}\left(\max_{\mathsf{x}\in\Phi_{\mathrm{B}}}\left\{\frac{H_i\ell(\mathsf{x})}{I'_{\mathrm{D}} + [I'_{\mathrm{B}} - H_i\ell(\mathsf{x})]}\right\} > \tau\right), \tag{5.41}$$

where I'_{D} and I'_{B} are the total interference terms from devices and base stations, respectively. Then for $\tau_{\mathrm{B}} \geq 1$, the probability of a device connecting to a base station, given there is interference from other base stations and devices, is

$$\mathcal{P}_{\mathrm{B}}(\tau_{\mathrm{B}}) = \frac{{\tau_{\mathrm{B}}}^{-\delta}}{\Gamma(1+\delta)\Gamma(1-\delta)}\frac{a_{\mathrm{B}}}{a_{\mathrm{B}} + a_{\mathrm{D}}}, \tag{5.42}$$

where the constants a_{B} and a_{D} are given by Expression (4.31) with the appropriate parameters used. The relatively simple expression stems from I'_{B} and I'_{D} being independent stable variables.

Let us also consider the SIR with respect to a typical device, assuming that a potential receiver (another device) is within a distance r to it. For convenience, we assume that this receiver is located at the origin and the typical device, denoted by x'_0, located at the distance r from the origin, independent of the Poisson process of other devices. We assume there is Rayleigh fading between the origin and the device x'_0, represented again by an exponential variable E with unit mean. Based on this model, the coverage probability of the typical device connecting to another device with the distance r is equal to

$$\mathcal{P}_{\mathrm{D}}(\tau_{\mathrm{D}}) := \mathbb{P}\left[\frac{E\ell(r)}{I'_{\mathrm{B}} + I'_{\mathrm{D}}} > \tau_{\mathrm{D}}\right] = \mathrm{e}^{-(a_{\mathrm{B}}+a_{\mathrm{D}})r^2\Gamma(1-\delta){\tau_{\mathrm{D}}}^{\delta}}, \tag{5.43}$$

where τ_{D} is the device-to-device SIR threshold, which stems from Proposition 5.2.1.

We now suppose at any moment of time only some proportion p of devices is allowed to communicate to other devices so as not to let I'_{D} become too large, which would degrade the quality of the downlink (base-station-to-device) communications. Meanwhile, the remaining proportion $1 - p$ of devices are nonactive, and so they are potentially receiving signals from base stations or other devices. We assume that the probability of any device being active or not is independent of all the other devices. In our Poisson model, this assumption leads to the modification of Expressions (5.42) and (5.43) by simply replacing a_{D} with pa_{D}.

We can then pose an optimization problem of maximizing the *device-to-device spatial throughput*, defined as

$$p\lambda_{\mathrm{D}}\mathcal{P}_{\mathrm{D}}(\tau_{\mathrm{D}}, p) = p\lambda_{\mathrm{D}}\, e^{-(a_{\mathrm{B}}+pa_{\mathrm{D}})r^2\Gamma(1-\delta){\tau_{\mathrm{D}}}^{\delta}},$$

while ensuring that the drop of the downlink coverage probability is not larger than some parameter $\epsilon \leq 1$ when compared to the situation when no device-to-device communication is allowed. In other words, we have the optimization problem

$$\begin{aligned} &\underset{0\le p\le 1}{\text{maximize}} \quad p\lambda_{\mathrm{D}}\mathcal{P}_{\mathrm{D}}(\tau_{\mathrm{D}},p) \\ &\text{subject to} \quad \inf_{\tau_{\mathrm{B}}\ge 1}\frac{\mathcal{P}_{\mathrm{B}}(\tau_{\mathrm{B}},p)}{\mathcal{P}_{\mathrm{B}}(\tau_{\mathrm{B}},0)} \ge \epsilon\,, \end{aligned}$$

where $\mathcal{P}_{\mathrm{B}}(\tau_{\mathrm{B}},p)$ is the downlink coverage probability when only fraction p of devices is allowed to transmit.

Under our device-to-device network model, this optimization problem has a surprisingly simple solution shedding light on how much coverage probability is lost when device-to-device communication is allowed in a network. The value of p that maximizes the device spatial throughput, thus solving Eq. (5.44), is given by

$$p^* = \min[1, p_1^*, p_2^*]\,, \tag{5.44}$$

where

$$p_1^* = \frac{{\tau_{\mathrm{D}}}^{-\delta}}{a_{\mathrm{D}} r^2 \Gamma(1-\delta)}, \tag{5.45}$$

$$p_2^* = \frac{a_{\mathrm{B}}}{a_{\mathrm{D}}}\left(\frac{1}{\epsilon}-1\right). \tag{5.46}$$

See Błaszczyszyn, Keeler, & Mühlethaler (2017) for the original model and optimization problem.

5.3.2 Multiple coverage in heterogeneous network

We now assume a Poisson network model $\tilde{\Phi}$ allowing different stations requesting different SINR thresholds for the connection. More specifically, we let

$$\tilde{\Phi} = \{\mathsf{x}, (H_{\mathsf{x}}, T_{\mathsf{x}}) : \mathsf{x} \in \Phi\}$$

be an independently marked Poisson point process of intensity λ, where for each $\mathsf{x} \in \Phi$, the mark T_{x} corresponds to the SINR threshold required by the station x. We assume that, given Φ, the SINR threshold T_{x} dos not depend on the location of base station x but can, in general, depend on the propagation effects H_{x}, which can account for the specific transmission power of station x. In short, given the point process Φ, the marks $(H_{\mathsf{x}}, T_{\mathsf{x}})$ are iid across all $\mathsf{x} \in \Phi$. This scenario may be cast into the general heterogeneous network model introduced in Section 4.2.2, in which case the types of base stations then correspond to their SINR thresholds and the set of all types $\mathcal{T}$ is any subset of $\mathbb{R}^+$.

For heterogeneous networks, we generalize the quantities we introduced in Section 5.3.1.

DEFINITION 5.3.8 (Coverage number) *The* coverage number *of the typical location is the random variable*

$$N \stackrel{\triangle}{=} \sum_{(\mathsf{x},T_{\mathsf{x}})\in\tilde{\Phi}} \mathbf{1}(\mathrm{SINR}(\mathsf{x},o) > T_{\mathsf{x}})\,. \tag{5.47}$$

Again, this is simply the number of base stations to which (the user at) the typical location can connect, but now each SINR threshold is specific to the base station. The definition of k-coverage probability remains the same, $\mathcal{P}^{(k)} \stackrel{\triangle}{=} \mathbb{P}\{\,\mathbf{N} \geq k\,\}$, namely the probability that at least k base stations offer the SINR value larger than their specific threshold. The symmetric sums also extend to the heterogeneous network setting.

DEFINITION 5.3.9 (Symmetric sum) *For $n \geq 1$, nth* symmetric sum *is*

$$\mathcal{S}_n \stackrel{\triangle}{=} \mathbb{E}\left[\sum_{\substack{(\mathsf{x}_1,T_{\mathsf{x}_1}),\ldots,(\mathsf{x}_n,T_{\mathsf{x}_n})\in\widetilde{\Phi}\\ \text{distinct}}} \mathbf{1}(\,\mathrm{SINR}(\mathsf{x}_j,o) > T_{\mathsf{x}_n},\, j=1,\ldots,n\,)\right]. \tag{5.48}$$

Again, we set $\mathcal{S}_0 \stackrel{\triangle}{=} 1$ and note that $\mathcal{S}_n$ is the expected number of ways of choosing n different base stations $\mathsf{x}_1,\ldots,\mathsf{x}_n$ such that the signal of each base station x_j is received by the user at the origin with an SINR larger than a specific value T_{x_j}.

Remark 5.8 The results of Lemma 5.3.5 hold with $\lceil 1/\tau \rceil$ replaced in general by ∞ for $\mathbf{N}$, $\mathcal{P}^{(k)}$, $k \geq 1$, and $\mathcal{S}_n$ defined in (5.46), (5.47), and (5.48), respectively. However, if there exists $\tau_{\min} > 0$ such that $\mathbb{P}(T \geq \tau_{\min}) = 1$, then by Lemma 5.1.2, $\mathcal{S}_n = 0$ for $n \geq 1/(\tau_{\min}) + 1$, which implies that one can replace ∞ by $\lceil 1/(\tau_{\min}) \rceil$ in the sums in the expressions given in Lemma 5.3.5.

To express the symmetric sums $\mathcal{S}_n$ and consequently k-coverage probabilities in the heterogeneous network model first we need to extend the family of functions $\bar{\mathcal{J}}$ defined by the integral in Eq. (5.32), which captures the combinatorial complexity of the k-coverage problem. For all $n \geq 1, 2, \ldots$ and all $t_i \geq 0$, we define

$$\mathcal{J}_{n,\delta}(t_1,\ldots,t_n) \stackrel{\triangle}{=} \frac{(1+\sum_{j=1}^{n} t_j)}{n} \int\limits_{[0,1]^{n-1}} \frac{\prod_{i=1}^{n-1} v_i^{i(\delta+1)-1}(1-v_i)^{\delta}}{\prod_{i=1}^{n}(t_i+\eta_i)} \mathrm{d}v_1 \ldots \mathrm{d}v_{n-1}, \tag{5.49}$$

where η_i are related to v_i by Eq. (5.33). For $n = 1$, Eq. (5.49) should be interpreted as $\mathcal{J}_{n,\delta}(t) = \bar{\mathcal{J}}_{n,\delta}(t) = 1$.

We need some more notation. For a vector $(t_1,\ldots,t_n)$, with $t_i > 0$, we define

$$t_i' = t_i'(t_i) = \frac{t_i}{1+t_i}, \tag{5.50}$$

and

$$\hat{t}_i = \hat{t}_i(t_1',\ldots,t_n') \stackrel{\triangle}{=} \frac{t_i'}{1-\sum_{j=1}^{n} t_j'}. \tag{5.51}$$

We define the functions

$$M^{(n)}(t_1,\ldots,t_n) \stackrel{\triangle}{=} \begin{cases} n!\left(\prod_{i=1}^{n} \hat{t}_i^{-\delta}\right)\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})]\mathcal{J}_{n,\delta}(\hat{t}_1,\ldots,\hat{t}_n), & \text{when } \sum_{i=1}^{n} t_n' < 1,\\ 0 & \text{otherwise,} \end{cases} \tag{5.52}$$

where the expectation $\mathbb{E}[\mathcal{I}_{n,\delta}(...)]$ is with respect to the noise term N as in Eq. (5.36).

PROPOSITION 5.3.10 *For the generalized heterogeneous network $\widetilde{\Phi}$, the nth symmetric sum is*

$$\mathcal{S}_n = \frac{1}{n!} \int\limits_{(\mathbb{R}_+)^n} \left(\prod_{i=1}^{n} \mathrm{CDF}_{T^*}(t_i) \right) M^{(n)}(\mathrm{d}t_1, \ldots, \mathrm{d}t_n), \tag{5.53}$$

which is equivalent to

$$\mathcal{S}_n = \frac{1}{n!} \int\limits_{(\mathbb{R}_+)^n} M^{(n)}(t_1, \ldots, t_n) \prod_{i=1}^{n} \mathrm{CDF}_{T^*}(\mathrm{d}t_i), \tag{5.54}$$

where $M^{(n)}$ is given by Eq. (5.52) and

$$\mathrm{CDF}_{T^*}(t) \stackrel{\triangle}{=} \mathbb{P}(T^* \leq t) = \frac{\mathbb{E}\left[H^{\delta} \mathbf{1}(T \leq t)\right]}{\mathbb{E}[H^{\delta}]}, \tag{5.55}$$

is the CDF of the SINR threshold in the equivalent homogeneous network model (see Corollary 4.2.5) or, equivalently, the conditional CDF of the SINR threshold requested by a base station, given its received power (see Eq. [4.38]).

We skip the proof of the above result, which is essentially an extension of the proof of Lemma 5.3.6, and refer the reader to Błaszczyszyn & Keeler (2015). We illustrate Proposition 5.3.10 results recalling the multitier network model with different transmission powers.

Example 5.4 *Open access multitier network with variable SINR threshold.* Consider the multitier network model with open access to all tiers studied in Example 5.1. Additionally, we assume now that each station from tier k requires some specific SINR threshold τ_k, $k = 1, \ldots, J$. Following the calculations of Expression (4.46), the discrete distribution function of the SINR threshold T^* in the equivalent homogeneous network can be evaluated

$$\mathbb{P}(T^* = \tau_k) = \frac{\mathbb{E}[H^{\delta} \mathbf{1}(T = \tau_k)]}{\mathbb{E}[H^{\delta}]} = \frac{a_k}{a}$$

with a_k as in Eq. (4.47) and $a = \sum_{j=1}^{J} a_j$. Consequently, by Eq. (5.54)

$$\mathcal{S}_n = \frac{1}{n!} \left[\sum_{i_1, \ldots, i_n \in \{1, \ldots, J\}} M^{(n)}\left(\tau_{i_1}, \ldots, \tau_{i_n}\right) \prod_{j=1}^{n} \frac{a_{i_j}}{a} \right] \tag{5.56}$$

and the probability that there are at least k stations covering the given location by the SINR larger than the threshold required for their tier is

$$\begin{aligned} \mathcal{P}^{(k)} = & \sum_{n=k}^{\left\lceil \frac{1}{\min(\tau_1, \ldots, \tau_J)} \right\rceil} (-1)^{n-k} \binom{n-1}{k-1} \mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})] \, \frac{1}{a^n} \\ & \times \sum_{\substack{i_1, \ldots, i_n \in \{1, \ldots, J\} \\ \frac{\tau_{i_1}}{1+\tau_{i_1}} + \cdots + \frac{\tau_{i_n}}{1+\tau_{i_n}} < 1}} \left(\prod_{j=1}^{n} a_{i_j} \hat{\tau}_{i_j}^{-\delta} \right) \mathcal{J}_{n,\delta}(\hat{\tau}_{i_1}, \ldots, \hat{\tau}_{i_n}), \end{aligned} \tag{5.57}$$

where, for $j = 1, \ldots, J$,

$$\hat{\tau}_{i_j} = \hat{\tau}_{i_j}(\tau_{i_1}, \ldots, \tau_{i_n}) \stackrel{\triangle}{=} \frac{\tau_{i_j}/(1+\tau_{i_j})}{1 - \sum_{l=1}^{n} \tau_{i_l}/(1+\tau_{i_l})} .$$

The number $n = \lceil 1/\min(\tau_1, \ldots, \tau_J) \rceil$ is precisely the largest value of n for which there exist $i_1, \ldots, i_n \in \{1, \ldots, J\}$ such that $\tau_{i_1}/(1+\tau_{i_1}) + \cdots + (\tau_{i_n}/(1+\tau_{i_n}) < 1$, and hence the second sum in Eq. (5.57) is not zero. In fact, in this case $\tau_{i_j} = \min(\tau_1, \ldots, \tau_J)$, for all $j = 1, \ldots, n$.

In particular, for $k = 1$ in the single coverage regime ($\tau_j \geq 1$), for all $j = 1, \ldots, J$ we have

$$\mathcal{P}^{(1)} = \mathbb{E}[\mathcal{I}_{1,\delta}(Na^{-1/\delta})] \frac{1}{a} \sum_{j=1}^{J} a_j \tau_j^{-1/\alpha} . \tag{5.58}$$

Example 5.5 *Closed-access multitier network with variable SINR threshold.* Following Example 5.4, suppose now that the user at the typical location seeks a base station yielding SINR greater than some threshold τ_j specific for the tier j, but only among the subset of tiers $j \in \mathbb{J} \subset \{1, \ldots, J\}$. This is an extension of the closed-access network model considered in Example 5.2 allowing for variable SINR thresholds. Using the same arguments as in this latter example one can show that the k-coverage probability in such a closed-access network is given by Expression (5.57) *with the second sum running over n-tuples of ties* $i_1, \ldots, i_n \in \mathbb{J}$ *instead in all* $\{1, \ldots, J\}$. Indeed, by replacing a^{-n} by $a_{\mathbb{J}}^{-n}$ and $\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})]$ by $\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})](\frac{a_{\mathbb{J}}}{a})^n$, we obtain $\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})](\frac{1}{a})^n$ as in the original Expression (5.57).

For $k = 1$ in the single coverage regime, that is, when $\tau_j \geq 1$ for all $j \in \mathbb{J}$, we have

$$\mathcal{P}^{(1)} = \mathbb{E}[\mathcal{I}_{1,\delta}(Na^{-1/\delta})] \frac{1}{a} \sum_{j \in \mathbb{J}} a_j \tau_j^{-1/\alpha} . \tag{5.59}$$

5.3.3 Matrix formulation of the multiple coverage event

Let us begin with the event defining coverage at threshold τ of the typical location by n base stations at distances of $r_1, \ldots, r_n$:

$$\bigcap_{i=1}^{n} \{\mathrm{SINR}(r_i) > \tau\} . \tag{5.60}$$

It is straightforward to rewrite $\bigcap_{i=1}^{n} (\mathrm{SINR}(r_i) > \tau)$ in Eq. (5.60) in the following equivalent matrix form:

$$\bigcap_{i=1}^{n} \{\mathrm{SINR}(r_i) > \tau\} = \left\{ \mathbf{C}\boldsymbol{E} > W\boldsymbol{\tau}' \right\} , \tag{5.61}$$

where

$$\mathbf{C} = \mathbf{I} - \boldsymbol{\tau}'\mathbf{1}^\top, \tag{5.62}$$

$$\boldsymbol{\tau}' = \tau'\mathbf{1}, \tag{5.63}$$

$$\tau' = \tau/(1+\tau), \tag{5.64}$$

$$W = N + I', \tag{5.65}$$

$$\boldsymbol{E} = [H_1\ell(r_1), \ldots, H_n\ell(r_n)]^\top, \tag{5.66}$$

and I' is the total interference power, i.e., the total received power at the typical location from all base stations other than the n base stations that provide coverage. For any two vectors $\boldsymbol{v} = [v_1, \ldots, v_n]^\top$ and $\boldsymbol{u} = [u_1, \ldots, u_n]^\top$, define $\boldsymbol{v} > \boldsymbol{u} \Leftrightarrow v_l > u_l$, $l = 1, \ldots, n$, with similar definitions for $\boldsymbol{v} \geq \boldsymbol{u}$, $\boldsymbol{v} < \boldsymbol{u}$, and $\boldsymbol{v} \leq \boldsymbol{u}$.

Note that the square matrix $\mathbf{C}$ has all off-diagonal entries nonpositive. Such matrices are called Z-matrices in the literature. For any nonzero $\boldsymbol{b} \geq \mathbf{0}$, the set $\{\boldsymbol{x} : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\}$ where $\mathbf{A}$ is a Z-matrix has the following structure.

LEMMA 5.3.11 *Given $\boldsymbol{b} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ and an $n \times n$ Z-matrix $\mathbf{A}$ (that is, every off-diagonal entry in $\mathbf{A}$ is ≤ 0), the set $\{\boldsymbol{x} \in \mathbb{R}^n_+ : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\}$ is nonempty if and only if $\mathbf{A}$ is an M-matrix*[2] *(that is, $\mathbf{A}^{-1}$ exists, and every entry in $\mathbf{A}^{-1}$ is ≥ 0).*

Proof Suppose $\mathbf{A}$ is not an M-matrix. From (I_{28}) in Berman and Plemmons (Berman & Plemmons 1994, Chapter 6, Theorem 2.3, p. 136), it follows that for all $\boldsymbol{x} \in \mathbb{R}^n$, if $\mathbf{A}\boldsymbol{x} > \mathbf{0}$ then it cannot be true that $\boldsymbol{x} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$. Further, $\mathbf{0} \notin \{\boldsymbol{x} \in \mathbb{R}^n_+ : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\}$ because $\boldsymbol{b} \neq \mathbf{0}$. Thus

$$\{\boldsymbol{x} \in \mathbb{R}^n_+ : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\} = \{\boldsymbol{x} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\} : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\} \subseteq \{\boldsymbol{x} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\} : \mathbf{A}\boldsymbol{x} > \mathbf{0}\} = \emptyset.$$

On the other hand, if $\mathbf{A}$ is an M-matrix, then every entry of $\mathbf{A}^{-1}$ is nonnegative and $\mathbf{A}^{-1}$ exists, so $\mathbf{A}^{-1}$ cannot have an all-zero row. In other words, $\mathbf{A}^{-1}\boldsymbol{\beta} > \mathbf{0}$ for all $\boldsymbol{\beta} > \mathbf{0}$. Therefore

$$\begin{aligned}\{\boldsymbol{x} \in \mathbb{R}^n_+ : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\} &= \mathbf{A}^{-1}\boldsymbol{b} + \{\boldsymbol{x} \in \mathbb{R}^n_+ : \mathbf{A}\boldsymbol{x} > \mathbf{0}\} \\ &= \mathbf{A}^{-1}\boldsymbol{b} + \{\boldsymbol{x} \in \mathbb{R}^n_+ : (\exists \boldsymbol{\beta} > \mathbf{0})\, \mathbf{A}\boldsymbol{x} = \boldsymbol{\beta}\} \\ &= \mathbf{A}^{-1}\tilde{\boldsymbol{b}} + \{\mathbf{A}^{-1}\boldsymbol{\beta} : \boldsymbol{\beta} > \mathbf{0}\} \subseteq \mathbb{R}^n_{++}, \end{aligned} \tag{5.67}$$

from which it follows that $\{\boldsymbol{x} \in \mathbb{R}^n_+ : \mathbf{A}\boldsymbol{x} > \boldsymbol{b}\}$ is nonempty and contains only vectors with all-positive entries. □

There are many different results giving conditions under which a Z-matrix can be an M-matrix. For our purposes, it is sufficient to use the following result.

LEMMA 5.3.12 *(Berman & Plemmons 1994, (E_{17}), Ch. 6, Thm. (2.3), p. 135) An $n \times n$ Z-matrix $\mathbf{A}$ is an M-matrix if and only if for every $k = 1, \ldots, n$, $\det \mathbf{A}_{[k]} > 0$, where $\mathbf{A}_{[k]}$ is the $k \times k$ submatrix of $\mathbf{A}$ comprising only the entries in the first k rows and first k columns.*

[2] Some authors refer to this definition as that of a nonsingular M-matrix.

Note that $\mathbf{A}_{[n]} = \mathbf{A}$. For any $n \times n$ Z-matrix $\mathbf{A}$ of the form

$$\mathbf{A} = \mathbf{I} - \boldsymbol{u}\boldsymbol{v}^\top, \quad \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\},$$

it turns out that

$$\det \mathbf{A}_{[1]} > \det \mathbf{A}_{[2]} > \cdots > \det \mathbf{A}_{[n]} \equiv \det \mathbf{A}, \tag{5.68}$$

because for each $k = 1, \ldots, n$,

$$\det\left[\mathbf{I}_k - (\boldsymbol{u})_k (\boldsymbol{v})_k^\top\right] = 1 - (\boldsymbol{v})_k^\top (\boldsymbol{u})_k,$$

where $(\boldsymbol{u})_k = [u_1, \ldots, u_k]^\top$, $(\boldsymbol{v})_k = [v_1, \ldots, v_k]^\top$, $k = 1, \ldots, n$. Hence such a matrix $\mathbf{A}$ is an M-matrix if and only if $\det \mathbf{A} > 0$.

It follows that the matrix $\mathbf{C}$ in Eq. (5.62) is an M-matrix if and only if $\det \mathbf{C} = 1 - n\tau' > 0$, which is an alternative proof of Lemma 5.1.2. We also note that $\mathbf{C}^{-1}$ can be computed in closed form using the Sherman-Morrison formula as

$$\mathbf{C}^{-1} = \mathbf{I} + \frac{\boldsymbol{\tau}' \mathbf{1}^\top}{1 - n\tau'}.$$

The probability of the event in Eq. (5.61) can now be computed as follows, with the exponential assumption for $\boldsymbol{E}$.

THEOREM 5.3.13 *Suppose $E_k \sim \text{Exp}(1/\mu_k)$ with $\mu_k > 0$, $k = 1, \ldots, n$; W is a random variable taking positive values with probability one, and $W, E_1, \ldots, E_n$ are all independent. Define $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_n]^\top$. Then the probability $\mathbb{P}\{\mathbf{C}\boldsymbol{E} > W\boldsymbol{\tau}'\}$ is given by*

$$\begin{aligned} \mathbb{P}\{\mathbf{C}\boldsymbol{E} > W\boldsymbol{\tau}'\} &= \mathbb{E}\left[\mathbb{P}\{\mathbf{C}\boldsymbol{E} > W\boldsymbol{\tau}' \mid W\}\right] \\ &= \begin{cases} 0, & \text{if } n\tau' \geq 0, \\ \left(\prod_{k=1}^n \mu_k\right) h_{\mathbf{C}^{-1}}(\boldsymbol{\mu}) \mathcal{L}_W(\boldsymbol{\mu}^\top \mathbf{C}^{-1} \boldsymbol{\tau}'), & \text{if } n\tau' < 0, \end{cases} \end{aligned} \tag{5.69}$$

where

$$h_{\mathbf{C}^{-1}}(\boldsymbol{\mu}) = \frac{1}{1 - n\tau'} \prod_{k=1}^n \frac{1}{\boldsymbol{\mu}^\top \mathbf{C}^{-1} \boldsymbol{e}_k}, \tag{5.70}$$

$$\mathbf{C}^{-1} = \mathbf{I} + \frac{\boldsymbol{\tau}' \mathbf{1}^\top}{1 - n\tau'}, \tag{5.71}$$

and $\boldsymbol{e}_k$ is the n-dimensional vector with ith entry $\delta_{k,i}$, $i = 1, \ldots, n$.

Proof The result when $n\tau' \geq 0$, i.e. when $\mathbf{C}$ is not an M-matrix, follows immediately from Lemma 5.3.11. If $\mathbf{C}$ is an M-matrix, we write $\mathbf{C}^{-1} = [\boldsymbol{a}_1^{(-1)}, \ldots, \boldsymbol{a}_n^{(-1)}]^\top$, where $\boldsymbol{a}_k^{(-1)} = \mathbf{C}^{-1}\boldsymbol{e}_k \in \mathbb{R}^n_+ \setminus \{\mathbf{0}\}$ is the kth column of $\mathbf{C}^{-1}$, $k = 1, \ldots, n$. Note that for any $\boldsymbol{\beta} = [\beta_1, \ldots, \beta_n]^\top$, $\mathbf{C}^{-1}\boldsymbol{\beta} = \sum_{k=1}^n \beta_k \boldsymbol{a}_k^{(-1)}$. The joint probability density function of $\boldsymbol{E} = [E_1, \ldots, E_n]^\top$ is $f_{\boldsymbol{E}}(\boldsymbol{x}) = (\prod_{k=1}^n \mu_k) \exp(-\boldsymbol{\mu}^\top \boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^n_+$. Then from Eq. (5.67), we can write

$$
\begin{aligned}
&\mathbb{P}\{\mathbf{C}\boldsymbol{E} > W\boldsymbol{\tau}' \mid W\} \\
&\quad = \int\limits_{\{\boldsymbol{x}\in\mathbb{R}_+^n:\mathbf{C}\boldsymbol{x}>W\boldsymbol{\tau}'\}} \left(\prod_{k=1}^{n}\mu_k\right)\exp(-\boldsymbol{\mu}^\top\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&\quad = \exp\left(-\boldsymbol{\mu}^\top\mathbf{C}^{-1}\boldsymbol{\tau}'W\right)\int\limits_{\{\mathbf{C}^{-1}\boldsymbol{\beta}:\boldsymbol{\beta}\in\mathbb{R}_{++}^n\}} \left(\prod_{k=1}^{n}\mu_k\right)\exp(-\boldsymbol{\mu}^\top\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \\
&\quad = \exp\left(-\boldsymbol{\mu}^\top\mathbf{C}^{-1}\boldsymbol{\tau}'W\right)\det\mathbf{C}^{-1}\int\limits_{\mathbb{R}_{++}^n}\left(\prod_{k=1}^{n}\mu_k\right)\exp\left\{-\sum_{k=1}^{n}\left[\boldsymbol{\mu}^\top\boldsymbol{a}_k^{(-1)}\right]\beta_k\right\}\,\mathrm{d}\boldsymbol{\beta} \\
&\quad = \left(\prod_{k=1}^{n}\mu_k\right)\exp\left(-\boldsymbol{\mu}^\top\mathbf{C}^{-1}\boldsymbol{\tau}'W\right)\det\mathbf{C}^{-1}\prod_{k=1}^{n}\int_0^\infty\exp\left\{-\left[\boldsymbol{\mu}^\top\boldsymbol{a}_k^{(-1)}\right]\beta_k\right\}\,\mathrm{d}\beta_k \\
&\quad = \left(\prod_{k=1}^{n}\mu_k\right)\exp\left(-\boldsymbol{\mu}^\top\mathbf{C}^{-1}\boldsymbol{\tau}'W\right)h_{\mathbf{C}^{-1}}(\boldsymbol{\mu}),
\end{aligned}
\tag{5.72}
$$

which proves the result. □

Remark 5.9 The matrix formulation of the multiple coverage event can be used to shorten the proof of Lemma 5.3.6 in Appendix A as follows: note that $\bigcap_{i=1}^{n}(\mathrm{STINR}(r_i) > \tau')$ in Eq. (A.4) is equivalent to Eq. (5.61). Theorem 5.3.13 then yields Eq. (5.69), and it is relatively straightforward to verify that Eq. (5.69) is the same as Eq. (A.12).

5.4 Bibliographic Notes

In the last few years there has been a surge in papers with SINR models based on the Poisson process, but we only mention some key papers here. An early SINR model based on a Poisson process was formulated by Zorzi and Pupolin in 1995 (Zorzi & Pupolin 1995), but these results remained largely unknown until years later. A stochastic geometric framework for the analysis of cellular networks was first proposed in (Baccelli, Klein, Lebourges, & Zuyev, 1997), "Stochastic geometry and architecture of communication networks" Telecommunications Systems, 7, pp. 209–227, 1997. It was general, not specific to wireless networks. Baccelli & Błaszczyszyn introduced the SINR coverage model based on the Poisson process in a stochastic geometry setting (Baccelli & Błaszczyszyn 2001). The Poisson-based model was then applied to ad hoc networks (Baccelli, Błaszczyszyn, & Mühlethaler 2003), and then developed further (Baccelli, Błaszczyszyn, & Mühlethaler 2006). Later Andrews, Baccelli, & Ganti (2011) adapted this approach to a cellular network model with an association policy in which the user connects to the closest base station (Andrews et al. 2011). Poisson multitier models of cellular networks soon followed, both with the previous

association policy (Jo et al. 2011, Mukherjee 2012, Nigam, Minero, & Haenggi 2014), and with the association policy of the user connecting to the instantaneously strongest signal (Dhillon, Ganti, Baccelli, & Andrews 2012) inspired Examples 5.1 and 5.2 (Mukherjee 2011, Madhusudhanan et al. 2011). Returning to the single-tier case, Vu and colleagues and Keeler and colleagues derived expressions for the coverage probability in Poisson models in which the user connects to the strongest signal, based on just the shadowing effects, and then Rayleigh fading is added, so the shadowing is separated from the fading (Keeler, Błaszczyszyn, & Karray 2013, Vu, Decreusefond, & Martins 2014). For further references and details, we recommend the following works: Baccelli & Błaszczyszyn 2009a, 2009b, Błaszczyszyn & Keeler 2015, ElSawy, Hossain, & Haenggi 2013, Haenggi 2012, Haenggi & Ganti 2008, Haenggi et al. 2009, Mukherjee 2014.

6 Downlink SINR: Advanced Results

6.1 More Advanced Results for Poisson Network with Singular Path Loss Model

Chapter 5, Sections 5.2 and 5.3, covers tractable coverage results for the single-tier and multitier network models based on the homogeneous Poisson process with the singular path loss model. Keeping the same model, we now present more complex concepts and results including the distribution of the SINR with implemented interference cancellation and signal combination schemes. These results follow from the study of the SINR process at the typical location and the factorial moment measures of the process. We implicitly used these measures in Expression (5.53) to express symmetric sums in the heterogeneous coverage scenario. They are useful tools in point process theory allowing, in particular, for expansion expressions for very general functionals of point processes. We present further results based on a simple relation between this SINR process and a point process known as the Poisson–Dirichlet process. We skip more technical proofs, and refer the reader to Błaszczyszyn and Keeler (Błaszczyszyn & Keeler 2015) for details and further citations.

6.1.1 SINR and STINR point processes at the typical location and their factorial moment measures

As in previous sections we consider the propagation model with the singular path loss function $\ell(\cdot)$ given by Expression (4.2). In Chapter 4, Section 4.2.1, we introduced the point process $\Theta = \{1/[H_{\mathsf{x}}\ell(\|\mathsf{x}\|)]\}_{\mathsf{x}\in\Phi}$ for such a propagation model and saw with Lemma 4.2.2 that Θ is an inhomogeneous Poisson point process with intensity measure $Q((0,t]) = at^{\delta}$, where $a = \pi\lambda\mathbb{E}[H^{\delta}]/\kappa^2$. For this network model, and ones that build off it, we always assume the finite moment condition in Eq. (4.30) is satisfied, namely $\mathbb{E}[H^{\delta}] < \infty$.

Based on Θ, we introduce the *SINR process* on the positive half-line $\mathbb{R}^+$

$$\Psi = \{\mathsf{Z} : \mathsf{Z} \in \Psi\} \stackrel{\triangle}{=} \left\{ \frac{1/\mathsf{L}}{N + (I' - 1/\mathsf{L})} : \mathsf{L} \in \Theta \right\} = \left\{ \frac{H_{\mathsf{x}}\ell(\mathsf{x})}{N + I' - H_{\mathsf{x}}\ell(\mathsf{x})} : \mathsf{x} \in \Phi \right\}, \tag{6.1}$$

where

$$I' \stackrel{\triangle}{=} \sum_{\mathsf{L}\in\Theta} (1/\mathsf{L}) = \sum_{\mathsf{x}\in\Phi} [H_{\mathsf{x}}\ell(\mathsf{x})], \tag{6.2}$$

is the power received from the entire network, called the total or aggregate interference. In other words, for base station $\mathsf{x} \in \Phi$ with the received signal strength $1/\mathsf{L} = \ell(\mathsf{x})H_{\mathsf{x}}$, the term $I' - \mathsf{L}^{-1}$ is the interference of x.

To study Ψ, as observed in Section 5.1.2, it is useful to consider also the *STINR process* on $(0, 1]$ as

$$\Psi' = \{\mathsf{Z}' : \mathsf{Z}' \in \Psi'\} \stackrel{\triangle}{=} \left\{ \frac{1/\mathsf{L}}{N + I'} : \mathsf{L} \in \Theta \right\}. \tag{6.3}$$

We can always recover one process from the other

$$\Psi = \left\{ \frac{\mathsf{Z}'}{1 - \mathsf{Z}'} : \mathsf{Z}' \in \Psi' \right\}, \qquad \Psi' = \left\{ \frac{\mathsf{Z}}{1 + \mathsf{Z}} : \mathsf{Z} \in \Psi \right\}. \tag{6.4}$$

For the zero-noise case $N = 0$, these relations still hold for the SIR and STIR process, where the latter becomes a point process with many interesting properties, and we still use the symbols Ψ and Ψ', explicitly specifying the value of N if the results depend on it.

To offer an analogy, we recall that for an integer-valued random variable X, its nth factorial moment is defined as $\mathbb{E}[X(X-1)\dots(X-k+1)^+]$. The factorial moments appear as coefficients of the Taylor expansion of the probability generating function $\mathbb{E}[s^X]$ of random variable X at $s = 1$. The corresponding objects for point processes are factorial moment measures (see Section 3.2.6). These measures also appear as coefficients of expansions of functionals of point processes (see Błaszczyszyn & Keeler 2015 and the references therein).

For $n \geq 1$ the nth factorial moment measure of the SINR process Ψ can be defined by specifying its values on the sets $(t_1, \infty] \times \cdots \times (t_n, \infty]$, with $t_1, \dots, t_n \geq 0$

$$M^{(n)}(t_1, \dots, t_n) \stackrel{\triangle}{=} M^{(n)}((t_1, \infty] \times \cdots \times (t_n, \infty]) \tag{6.5}$$

$$\stackrel{\triangle}{=} \mathbb{E}\left(\sum_{\substack{(\mathsf{Z}_1,\dots,\mathsf{Z}_n)\in(\Psi)^{\times n} \\ \text{distinct}}} \prod_{j=1}^{n} \mathbf{1}(\mathsf{Z}_j > t_j) \right). \tag{6.6}$$

Similarly for the STINR process Ψ' and $0 \leq t'_1, \dots, t'_n \leq 1$, we define the factorial moment measures

$$M'^{(n)}(t'_1, \dots, t'_n) \stackrel{\triangle}{=} M'^{(n)}\left((t'_1, 1] \times \cdots \times (t'_n, 1]\right) \tag{6.7}$$

$$\stackrel{\triangle}{=} \mathbb{E}\left(\sum_{\substack{(\mathsf{Z}'_1,\dots,\mathsf{Z}'_n)\in(\Psi')^{\times n} \\ \text{distinct}}} \prod_{j=1}^{n} \mathbf{1}(\mathsf{Z}'_j > t'_j) \right), \tag{6.8}$$

By Expression (6.4)

$$M^{(n)}(t_1,\ldots,t_n,) = M'^{(n)}\left(\frac{t_1}{1+t_1},\ldots,\frac{t_n}{1+t_n}\right), \tag{6.9}$$

hence we can easily recover factorial moment measures of the SINR process from those of STINR and the other way around. The following result gives the factorial moment measures of the SINR process.

THEOREM 6.1.1 *The factorial moment measures $M'^{(n)}$ of the STINR process Ψ satisfy for all $n \geq 1$ and $t'_1,\ldots,t'_n \in (0,1]$*

$$M'^{(n)}(t'_1,\ldots,t'_n) = \begin{cases} n!\left(\prod_{i=1}^{n}\hat{t}_i^{-\delta}\right)\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})]\mathcal{J}_{n,\delta}(\hat{t}_1,\ldots,\hat{t}_n), & \text{when } \sum_{i=1}^{n} t'_n < 1, \\ 0 & \text{otherwise,} \end{cases} \tag{6.10}$$

where

$$\hat{t}_i = \hat{t}_i(t'_1,\ldots,t'_n) \stackrel{\triangle}{=} \frac{t'_i}{1-\sum_{j=1}^{n} t'_j} \tag{6.11}$$

and the expectation $\mathbb{E}[\mathcal{I}_{n,\delta}(...)]$ is with respect to the noise N as in Eq. (5.36).

For the proof we refer the reader to (Błaszczyszyn & Keeler 2015). By Expression (6.9), the factorial moment measures $M^{(n)}$ of the SINR process Ψ satisfy Expression (5.52) for all $n \geq 1$ and $t_1,\ldots,t_n > 0$. We recall, via Expression (5.53), that these measures allow us to express symmetric sums in the heterogeneous coverage scenario.

Surprisingly, we can express the densities of $M'^{(n)}$ and $M^{(n)}$ quite explicitly. To this end, we normalize the functions $\mathcal{I}_{n,\delta}(t)$ given by Expression (5.29), which capture the noise N impact in Expressions (6.10) and (5.52), so as they take value 1 when for $t = 0$

$$\bar{\mathcal{I}}_{n,\delta}(t) \stackrel{\triangle}{=} \frac{\mathcal{I}_{n,\delta}(t)}{\mathcal{I}_{n,\delta}(0)} = \frac{2\int_0^\infty u^{2n-1}\mathrm{e}^{-u^2-tu^\alpha\Gamma(1-\delta)^{-1/\delta}}\mathrm{d}u}{(n-1)!}. \tag{6.12}$$

Also, for $n \geq 0$, we write

$$c_{n,\delta} \stackrel{\triangle}{=} \frac{\delta^{n-1}\Gamma(n)}{\Gamma(n\delta)(\Gamma(1-\delta))^n}.$$

THEOREM 6.1.2 *For $n \geq 1$, the density of the nth factorial moment measure $M'^{(n)}$ of the STINR process Ψ' satisfies*

$$\begin{aligned}\mu'^{(n)}(t'_1,\ldots,t'_n) &\stackrel{\triangle}{=} (-1)^n\frac{\partial^n M'^{(n)}(t'_1,\ldots,t'_n)}{\partial t'_1\ldots\partial t'_n} \\ &= c_{n,\delta}\,\mathbb{E}[\bar{\mathcal{I}}_{n,\delta}(Na^{-1/\delta})]\left(\prod_{i=1}^{n}(t'_i)^{-(\delta+1)}\right)\left(1-\sum_{j=1}^{n}(t'_j)\right)^{n\delta-1}\end{aligned} \tag{6.13}$$

where the expectation $\mathbb{E}[\bar{\mathcal{I}}_{n,\delta}(...)]$ *is taken with respect to the noise* N *and can be expressed using its Laplace similarly as in Expression (5.36).*

By Expression (6.9) we obtain immediately the result for the SINR process.

COROLLARY 6.1.3 *The density* $\mu^{(n)}$ *of the nth factorial moment measure* $M^{(n)}$ *of the SINR process* Ψ *is related to this of STINR by*

$$\mu^{(n)}(t_1,\ldots,t_n) = \prod_{j=1}^{n} \frac{1}{(1+t_j)^2}\ \mu'^{(n)}\left(\frac{t_1}{1+t_1},\ldots,\frac{t_n}{1+t_n}\right).$$

Remark on the proof of Theorem 6.1.2 The result can be proved by factoring out in Eq. (6.10) the impact of the noise

$$\mathbb{E}[\mathcal{I}_{n,\delta}(Na^{-1/\delta})] = \mathbb{E}[\bar{\mathcal{I}}_{n,\delta}(Na^{-1/\delta})] \times \mathcal{I}_{n,\delta}(0)]$$

and relating $M'^{(n)}$ with $\mathcal{I}_{n,\delta}(0)$ of the STINR process Ψ' in the case of no noise $N = 0$ to some Poisson–Dirichlet process whose product or (factorial) moment densities can be more straightforwardly evaluated using an equivalent definition, detailed in Section 6.2.[1] □

6.1.2 Order statistics of the STINR process

We now examine the order statistics of the STINR process, which later allow us to derive expressions for coverage probability when signals are removed from the interference or combined. We denote the order statistics of the STINR process Ψ' by

$$\mathsf{Z}'_{(1)} > \mathsf{Z}'_{(2)} > \mathsf{Z}'_{(3)} \ldots,$$

such that $\mathsf{Z}'_{(1)}$ is the largest STINR value in Ψ'.

If we denote in a similar way the order statistics $\mathsf{Z}_{(1)} > \mathsf{Z}_{(2)} > \mathsf{Z}_{(3)} \ldots$ of the SINR process Ψ, then by the monotonicity of the relation in Eq. (5.3)

$$\mathsf{Z}_i = \frac{\mathsf{Z}'_{(i)}}{(1+\mathsf{Z}'_{(i)})}$$

for all $i \geq 1$. With this property on mind, we focus now only on the STINR process.

We recall from Section 5.1.2, Section 5.1.3, and Section 5.2.3 that

$$\mathcal{P}^{(k)}(\tau) = \mathbb{P}\left(\mathsf{Z}_{(k)} > \tau\right) = \mathbb{P}\left(\mathsf{Z}'_{(k)} > \tau'\right),$$

where $\tau' = \tau/(1+\tau)$. Consequently, we already know the marginal distributions of the order statistics of both processes.

To study the joint distribution of order statistics of the STINR process, we define some auxiliary functions: For $i \geq 1$, we write

$$\mu'^{(k+i)}_k(z'_1,\ldots,z'_k) \triangleq \int_{z'_k}^{1} \cdots \int_{z'_k}^{1} \mu'^{(k+i)}(z'_1,\ldots,z'_k,\zeta'_1,\ldots,\zeta'_i)\, d\zeta'_i \ldots d\zeta'_1, \tag{6.14}$$

[1] We are unaware of anyone directly showing the equivalence of Theorem 6.1.1 and Theorem 6.1.2, by either differentiating the measure (Eq. [6.10]) or integrating the product density (Eq. [6.12]).

where $\mu'^{(n)}$ is the density of the factorial moment measure $M'^{(n)}$ of the STINR process Ψ' given in Theorem 6.1.2. This is equivalent to differentiating k times $M'^{(n)}$

$$\mu_k'^{(k+i)}(z'_1,\ldots,z'_k) = (-1)^k \frac{\partial^k M'^{(k+i)}(t'_1,\ldots,t'_k,z'_k,\ldots,z'_k)}{\partial t'_1 \ldots \partial t'_k}(t'_1 = z'_1,\ldots,t'_k = z'_k). \tag{6.15}$$

We also define $\mu_k'^{(k+i)}$ for $i = 0$ as the density of $M'^{(k)}$; $\mu_k'^{(k)} \triangleq \mu'^{(k)}$. We call $\mu_k'^{(n)}$ $(k \leq n)$ the partial densities of $M'^{(n)}$.

PROPOSITION 6.1.4 *Under the assumptions of Theorem 6.1.1, the joint probability density of the vector of k strongest values of the STINR process* $(\mathsf{Z}'_{(1)},\ldots,\mathsf{Z}'_{(k)})$ *is equal to*

$$f'_{(k)}(z'_1,\ldots,z'_k) = \sum_{i=0}^{i_{\max}} \frac{(-1)^i}{i!} \mu_k'^{(k+i)}(z'_1,\ldots,z'_k), \tag{6.16}$$

for $z'_1 > z'_2 > \cdots > z'_k$ *and* $f'_{(k)}(z'_1,\ldots,z'_k) = 0$ *otherwise, where the upper summation limit is bounded by*

$$i_{\max} < \frac{1}{z'_k} - k. \tag{6.17}$$

The bound on $i_{\max}$ given in Eq. (6.17) is only an upper bound of the maximum index of non-zero terms of the expansion of $f'_{(k)}(t'_1,\ldots,t'_k)$ in terms of the partial densities of the factorial moment measures. For some particular domains of $(t'_1,\ldots,t'_k)$, a smaller value of $i_{\max}$ can be given.

By combining Theorem 6.1.2 and Proposition 6.1.4, we obtain more explicit expression for the joint densities $f'_{(k)}$.

COROLLARY 6.1.5 *Under the assumptions of Theorem 6.1.1, we have*

$$\begin{aligned} f'_{(k)}(z'_1,\ldots,z'_k) = & \left(\prod_{j=1}^{k}(z'_j)^{-(\delta+1)}\right) \\ & \times \sum_{i=0}^{\lceil 1z'_k+k-1\rceil} \frac{(-1)^i}{i!} c_{k+i,\alpha}\, \mathbb{E}[\bar{\mathcal{I}}_{k+i,\alpha}(Na^{-1/\delta})^{k+i}] \\ & \times \int_{z'_k}^{1}\cdots\int_{z'_k}^{1}\left(\prod_{j=k+1}^{k+i}(\zeta'_j)^{-(\delta+1)}\right) \\ & \times \left(1-\sum_{j=1}^{k}(z'_j)-\sum_{j=k+1}^{k+i}(\zeta'_j)\right)^{\delta(k+i)-1} d\zeta'_1\ldots d\zeta'_i, \end{aligned}$$

for $z'_1 > z'_2 > \cdots > z'_k$ *and* $f'_{(k)}(z'_1,\ldots,z'_2) = 0$ *otherwise.*

Example 6.1 As an example of the application of Proposition 6.1.4 or Corollary 6.1.5, we note that for any $\tau > 0$ and $\tau' = \tau/(1+\tau)$

$$\int_{\tau'}^{1} \cdots \int_{\tau'}^{1} f'_{(k)}(z'_1, \ldots, z'_k)\, dz'_1, \ldots, dz'_k = \mathcal{P}^{(k)}(\tau) \tag{6.18}$$

is equal to the k-coverage probability given in Corollary 5.3.7.

6.1.3 SINR with general interference cancellation and signal combination

We now demonstrate how to use the order statistics of the STINR process to express SINR values accounting for general interference cancellation (or reduction) and signal combination techniques.

We observe first by Equation (6.3) that $\mathsf{Z}'_{(i)}$, and hence $\mathsf{Z}_{(i)}$, correspond to the STINR and SINR, respectively, with the useful signal path loss being the ith order increasing order statistic $\mathsf{L}_{(i)}$ of the projection process $\mathsf{L}_{(1)} < \mathsf{L}_{(2)} < \cdots \in \Phi$

$$\mathsf{Z}'_{(i)} = \frac{1/\mathsf{L}_{(i)}}{N + I'} \qquad \mathsf{Z}_{(i)} = \frac{1/\mathsf{L}_{(i)}}{N + I' - 1/\mathsf{L}_{(i)}}\,.$$

We now consider a set of $k \geq 1$ strongest signals received at the typical location $1/\mathsf{L}_{(1)} > \cdots > 1/\mathsf{L}_{(k)}$. For a subset $\mathbf{K} \subset [k] \stackrel{\triangle}{=} \{1, \ldots, k\}$ of this set, we define the SINR under *interference cancellation(IC) and signal combination* (ICSC-SINR) as

$$\mathrm{SINR}_{\mathbf{K},k} = \frac{\sum_{i\in\mathbf{K}} (1/\mathsf{L}_{(i)})}{N + I' - \sum_{j=1}^{k} (1/\mathsf{L}_{(j)})}\,, \tag{6.19}$$

where signals in $\mathbf{K}$ are combined and interference from all the remaining interferers in $[k] \setminus \mathbf{K}$ is canceled. We see that the ICSC-SINR coverage events can be expressed using the order statistics, so that

$$\begin{aligned}
\{\, \mathrm{SINR}_{\mathbf{K},k} > \tau \,\} &= \left\{ \sum_{i\in\mathbf{K}} (\mathsf{L}_{(i)})^{-1} > \tau N + \tau I' - \tau \sum_{j=1}^{k} (\mathsf{L}_{(j)})^{-1} \right\} \\
&= \left\{ (1+\tau) \sum_{i\in\mathbf{K}} \mathsf{Z}'_{(i)} + \tau \sum_{j\in[k]\setminus\mathbf{K}} \mathsf{Z}'_{(j)} > \tau \right\} \\
&= \left\{ \sum_{i\in\mathbf{K}} \mathsf{Z}'_{(i)} + \tau' \sum_{j\in[k]\setminus\mathbf{K}} \mathsf{Z}'_{(j)} > \tau' \right\},
\end{aligned} \tag{6.20}$$

where, $\tau' = \tau/(1+\tau)$. Expression (6.20) may ignore (practical) conditions under which signal cancellation can be effectively performed. To take them into account, one may multiply the canceled interference term $\tau' \sum_{j\in[k]\setminus\mathbf{K}} \mathsf{Z}'_{(j)}$ by the indicator $\mathbf{1}(\mathrm{IC})$ of a

suitable condition IC for the feasibility of the interference cancellation. This leads us to the following definition of the general ICSC-SINR coverage probability

$$\mathcal{P}^{(\mathbf{K},k)}(\tau) \stackrel{\triangle}{=} \mathbb{P}\{\,\mathrm{SINR}_{\mathbf{K},k} > \tau\} = \mathbb{P}\left\{ \sum_{i\in\mathbf{K}} \mathsf{Z}'_{(i)} + \mathbf{1}(\mathrm{IC}) \times \tau' \sum_{j\in[k]\setminus\mathbf{K}} \mathsf{Z}'_{(j)} > \tau' \right\}. \quad (6.21)$$

Expression (6.21) is the probability that the ratio of the combined signals to the noise and noncanceled interference is larger than the threshold τ.

Natural conditions for IC can also be expressed in terms of the values $\mathsf{Z}'_{(i)}$ $i = 1, \ldots, k$, and so $\mathcal{P}_{\mathrm{IC}}^{(\mathbf{K},k)}$ can be calculated using Proposition 6.1.4. We could in a similar fashion introduce conditions for the feasibility of the signal combination. We give two examples of natural IC conditions.

Example 6.2 *Successive interference cancellation (SIC) condition.* SIC consists of first decoding the strongest interfering signal (among k strongest), subtracting it from the interference, then decoding the second-strongest one, and so on. We assume that we can decode a signal if its SINR is larger than some threshold ϵ, which we call the *IC threshold*. For the interfering signals to be decoded and subtracted, we denote their indices by $j_1 < \cdots < j_{k'}$, where $k' = k - \|\mathbf{K}\|$, starting from the strongest one. The SIC condition can be written then as the superposition of the following conditions

$$\begin{aligned} \mathsf{Z}'_{(j_1)} &> \epsilon' \\ \mathsf{Z}'_{(j_2)} + \epsilon' \mathsf{Z}'_{(j_1)} &> \epsilon' \\ &\cdots \\ \mathsf{Z}'_{(j_{k'})} + \epsilon' \sum_{l=1}^{k'-1} \mathsf{Z}'_{(j_l)} &> \epsilon', \end{aligned}$$

where $\epsilon' = \epsilon/(1+\epsilon)$.

Example 6.3 *Independent interference cancellation (IIC) condition.* IIC is a weaker condition assuming that all interfering signals to be canceled can be decoded independently: $\mathsf{Z}'_{(j)} > \epsilon'$ for all $j \in [k] \setminus \mathbf{K}$.

We give a specific example of the ICSC scenario and express the corresponding ICSC-SINR coverage probabilities.

Example 6.4 *Coverage by the kth strongest signal with cancellation of all stronger signals.* Consider coverage by the kth strongest signal with all $k-1$ stronger signals canceled from the interference. (The remaining interference is called the *residual* one.) That is we take $\mathbf{K} = \{k\}$. We do not consider any *IC* condition; hence $\mathbf{1}(IC) \equiv 1$. Denote by $\mathcal{P}^{(\{k\},k)}(\tau)$ the corresponding ICSC-SINR coverage probability.

The next result follows from Proposition 6.1.4.

PROPOSITION 6.1.6 *Under the assumptions of Theorem 6.1.1*

$$\mathcal{P}^{(\{k\},k)}(\tau) = \sum_{i=0}^{i_{\max}} \frac{(-1)^i}{i!} \int_0^1 \cdots \int_0^1 \mathbf{1}\left(\tau' \sum_{i=1}^{k-1} z'_i + z'_k > \tau'\right) \times \mathbf{1}(z'_1 > \cdots > z'_k)\, \mu_k^{(k+i)}(z'_1, \ldots, z'_k)\, dz'_k \ldots dz'_1 , \tag{6.22}$$

where the upper summation limit is bounded by

$$i_{\max} < 1/\tau' - 1 = 1/\tau . \tag{6.23}$$

Despite the appearance of the integral in Eq. (6.22), strikingly, this result greatly simplifies in a similar way to the k-coverage results when $\tau \geq 1$.

COROLLARY 6.1.7 *For $\tau \geq 1$ we have $i_{\max} = 0$ and*

$$\mathcal{P}^{(\{k\},k)}(\tau) = \frac{\mathbb{E}[\bar{\mathcal{I}}_{k,\delta}(Na^{-1/\delta})]}{\tau^{k\delta}\Gamma(1+k\delta)(\Gamma(1-\delta))^k}$$

where $\bar{\mathcal{I}}_{k,\delta}$ is given by Eq. (6.12). For $N = 0$, we have

$$\mathcal{P}^{(\{k\},k)}(\tau) = \frac{1}{\tau^{k\delta}\Gamma(1+k\delta)(\Gamma(1-\delta))^k}. \tag{6.24}$$

We now, in a certain sense, inverse the scenario and examine how the coverage by the strongest signal can be improved by removing $k-1$ successive strongest signals from the interference or by combining them with the strongest signal. We consider two examples.

Example 6.5 *Coverage by the strongest signal with cancellation of subsequent $k-1$ strongest signals.* We assume that the receiver, being served by the strongest base station, is able to remove the interference created by the subsequent $k-1$ strongest stations, so we set $\mathbf{K} = \{1\}$. For the interference cancellation condition IC, we assume the IIC scenario of Example 6.3, so signals $2, \ldots, k$ are independently (decoded and) canceled only when their SINR level exceeds ϵ. For any ϵ, τ, with $0 < \epsilon < \tau$, and $\epsilon' = \epsilon/(1+\epsilon)$, $\tau' = \tau/(1+\tau)$, we define the corresponding ICSC-SINR coverage probability with interference cancellation

$$\mathcal{P}_{IC}^{(k)}(\tau,\epsilon) \stackrel{\triangle}{=} \mathbb{P}\left\{ \begin{array}{ll} \mathrm{SINR}_{\{1\},k} > \tau & \text{when } \mathsf{Z}_{(k)} > \epsilon \\ \mathsf{Z}_{(1)} > \tau & \text{otherwise} \end{array} \right\}$$

$$= \mathbb{P}\left\{ \mathsf{Z}'_{(1)} + \tau'\left(\mathsf{Z}'_{(2)} + \cdots + \mathsf{Z}'_{(k)}\right)\mathbf{1}(\mathsf{Z}'_{(k)} > \epsilon') > \tau' \right\}.$$

Example 6.6 *Coverage by the combination of the k strongest signals.*

We assume now that the receiver adds (combines) all $k-1$th strongest signals into a useful signal of the strongest station and suffers only from the residual interference, so we set $\mathbf{K} = \{1, \ldots, k\}$. However, we assume that the combination of a signal i, $2 \leq i \leq k$, is possible only when its SINR level exceeds ϵ. (This is hence a condition

similar to IIC but now applied for signal combination.) For any ϵ, τ, with $0 < \epsilon < \tau$, and $\epsilon' = \epsilon/(1+\epsilon)$, $\tau' = \tau/(1+\tau)$, we define the corresponding ICSC-SINR coverage probability with signal combination

$$\begin{aligned}\mathcal{P}_{SC}^{(k)}(\tau,\epsilon) &\stackrel{\triangle}{=} \mathbb{P}\left\{ \begin{array}{ll} \mathrm{SINR}_{[k],k} > \tau & \text{when } \mathsf{Z}_{(k)} > \epsilon \\ \mathsf{Z}_{(1)} > \tau & \text{otherwise} \end{array} \right\} \\ &= \mathbb{P}\left\{ \mathsf{Z}'_{(1)} + \left(\mathsf{Z}'_{(2)} + \cdots + \mathsf{Z}'_{(k)}\right)\mathbf{1}(\mathsf{Z}'_{(k)} > \epsilon') > \tau' \right\}.\end{aligned}$$

Note that

$$\mathcal{P}_{SC}^{(k)}(\tau,\epsilon) \geq \mathcal{P}_{IC}^{(k)}(\tau,\epsilon) \geq \mathcal{P}^{(k)}(\tau) \tag{6.25}$$

for all $0 < \epsilon \leq \tau$. To study the coverage improvement by the signal combination or interference cancellation, we consider the difference between the coverage probabilities

$$\mathcal{P}_{IC}^{(k)}(\tau,\epsilon) = \mathcal{P}(\tau) + \Delta_{IC}^{(k)}(\tau,\epsilon) \tag{6.26}$$

$$\mathcal{P}_{SC}^{(k)}(\tau,\epsilon) = \mathcal{P}(\tau) + \Delta_{SC}^{(k)}(\tau,\epsilon), \tag{6.27}$$

where

$$\Delta_{IC}^{(k)}(\tau,\epsilon) \stackrel{\triangle}{=} \mathbb{P}\left\{ \mathsf{Z}'_{(1)} + \tau'\left(\mathsf{Z}'_{(2)} + \cdots + \mathsf{Z}'_{(k)}\right) > \tau' \text{and } \epsilon' < \mathsf{Z}'_{(i)} < \tau', i = 1 \ldots k \right\} \tag{6.28}$$

$$\Delta_{SC}^{(k)}(\tau,\epsilon) \stackrel{\triangle}{=} \mathbb{P}\left\{ \mathsf{Z}'_{(1)} + \cdots + \mathsf{Z}'_{(k)} > \tau' \text{ and } \epsilon' < \mathsf{Z}'_{(i)} < \tau', i = 1 \ldots k \right\}. \tag{6.29}$$

We have the following result for theses differences.

PROPOSITION 6.1.8 *Under the assumptions of Theorem 6.1.1,*

$$\begin{aligned}\Delta_{IC}^{(k)}(\tau,\epsilon) = \sum_{i=0}^{i_{\max}} \frac{(-1)^i}{i!} \int_{\epsilon'}^{\tau'} \cdots \int_{\epsilon'}^{\tau'} \mathbf{1}\left(z'_1 + \tau'(z'_2 + \cdots + z'_k) > \tau'\right) \\ \times \mathbf{1}(z'_1 > \cdots > z'_k)\, \mu_k'^{(k+i)}(z'_1, \ldots, z'_k)\, dz'_k \ldots dz'_1\end{aligned} \tag{6.30}$$

and

$$\begin{aligned}\Delta_{SC}^{(k)}(\tau,\epsilon) = \sum_{i=0}^{i_{\max}} \frac{(-1)^i}{i!} \int_{\epsilon'}^{\tau'} \cdots \int_{\epsilon'}^{\tau'} \mathbf{1}\left(z'_1 + \cdots + z'_k) > \tau'\right) \\ \times \mathbf{1}(z'_1 > \cdots > z'_k)\, \mu_k'^{(k+i)}(z'_1, \ldots, z'_k)\, dz'_k \ldots dz'_1\end{aligned} \tag{6.31}$$

where the upper summation limit is bounded by

$$i_{\max} < \frac{1}{\epsilon'} - k. \tag{6.32}$$

We note that the parameter ϵ, corresponding to the condition on the feasibility of the signal cancellation or combination, allows mathematically the expansions of $\Delta_{IC}^{(k)}$ and $\Delta_{SC}^{(k)}$ to be finite ones, with $i_{\max} < \infty$.

6.1.4 Some numerical results

We illustrate the mathematical methods covered so far by presenting numerical results for several cellular network models. For all results, we have set $\kappa = 1$ and assumed zero noise ($N = 0$), the results do not depend on base station density λ, meaning they are scale invariant. Usually, a positive noise term makes negligible difference except for very low network density values. This sitution changes for multitier networks, where we are more specific about λ values. To illustrate the impact of the strength of the path loss we consider two path loss exponents $\alpha = 3$ and $\alpha = 5$. The SINR thresholds τ and ϵ are expressed in decibels (dB), that is, $\tau(\text{dB}) = 10\log_{10}(\tau)\text{dB}$.

For some integrals, numerical integration was required, and for dimensions beyond four, the best approach, arguably, is Monte Carlo or quasi-Monte Carlo methods. Despite the name, there is nothing random about quasi-Monte Carlo methods. They rely upon a mathematical result that sampling integral kernels at certain points, by using sequences called quasi-random numbers, speeds up the integration procedure. Quasi-random numbers are known as low-discrepancy numbers, which in loose terms means that they do not cluster, allowing them to more evenly and more efficiently sample the integral kernel values. For the best performance, the nature of the integral kernel dictates what type of quasi-random numbers to be used for the integral estimation. We will not go into such details, but recommend the readable introduction (Kuo & Sloan 2005) as well as the detailed survey article (Dick, Kuo, & Sloan 2013).

We performed our quasi-Monte Carlo integration with Sobol points generated by MATLAB (Keeler 2014). We found the evaluation of the $\mathcal{J}_{n,\delta}$ integral is achieved quickly on a standard machine (the number of sample points mostly ranges from 10^3 to 10^4 points). The structure of integrals such as those in Eqs. (6.30) and (6.31) allows the integration with respect to the z_i' and v_i (from $\mathcal{J}_{n,\delta}$) to be done in the same computation step, which takes slightly longer to evaluate than $\mathcal{J}_{n,\delta}$.

Example 6.7 *Single-tier network.* Figures 6.1 and 6.2 present the k-coverage probability $\mathcal{P}^{(k)}(\tau)$ for $k = 1, 2, 3$ with $\alpha = 3$ and 5, respectively. $\mathcal{P}^{(k)}(\tau)$ is given by Eq. (5.37) and is the tail distribution function of the SINR related to the kth strongest signal.

Example 6.8 *Two-tier network.* We have a two-tier network model with $\lambda_1 = \lambda_2/2$, and $\alpha = 3$. The (constant) transmission power terms $P_1 = 100P_2$ are introduced by replacing $\mathbb{E}(H^\delta)$ with $\mathbb{E}[(PH^\delta)]$ in all the expressions. For the values $\tau_2 = 1$dB and $\tau_2 = -2$ dB, we presents plots of 1-coverage and 2-coverage probability $\mathcal{P}^{(k)}$, $k = 1, 2$ as functions of τ_1; see Figures 6.3–6.5. For comparison purposes, on Figures 6.3–6.5, we consider also a single-tier network model with intensity $\lambda^* = \lambda_1 P_1^\delta + \lambda_2 P_2^\delta$, which gives equivalent versions of processes Θ and Ψ. For this model,

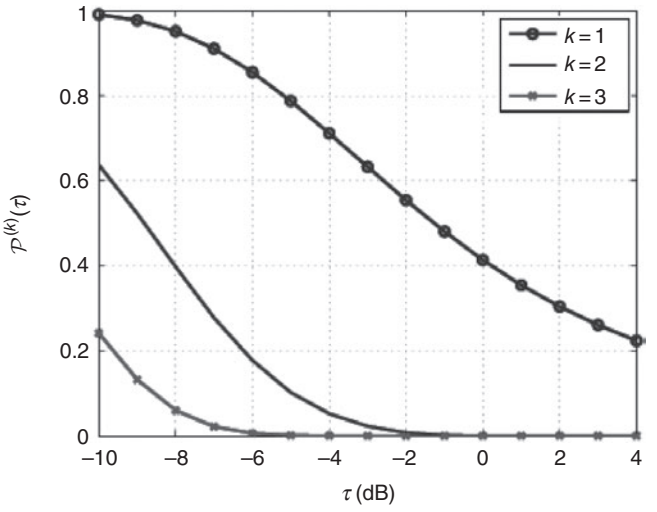

Figure 6.1 k-Coverage probability $\mathcal{P}^{(k)}(\tau)$ for a single-tier network, where $\alpha = 3$.

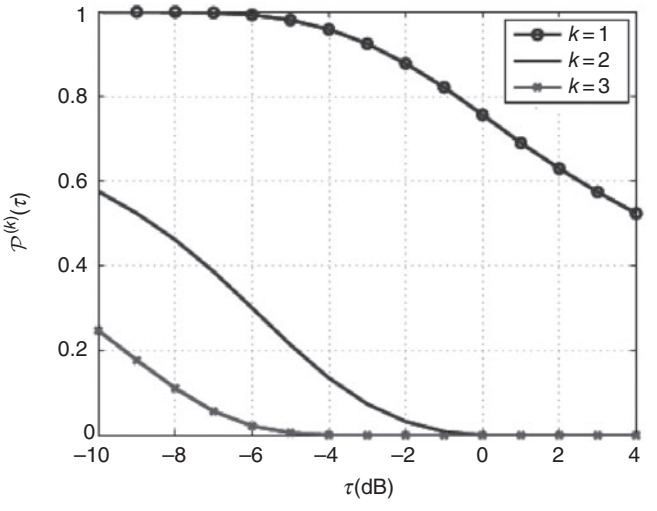

Figure 6.2 k-Coverage probability $\mathcal{P}^{(k)}(\tau)$ for a single-tier network, where $\alpha = 5$.

we calculate the similar k-coverage probabilities with constant SINR thresholds $\tau = \mathbb{E}[T^*] = (\tau_1\lambda_1 P_1^{\delta} + \tau_2\lambda_2 P_2^{\delta})/\lambda^*$, equal to the (spatial) mean of the random SINR threshold mark in the truly equivalent model. For our parameter values, this model offers quite reasonable approximation of the original two-tier one, in particular for $\tau > 1$. Although to the eye the corresponding plots almost coincide, it is still an approximation.

Example 6.9 *Coverage with interference cancellation and cooperation.* Following Examples 6.5 and 6.6 in Figures 6.6 and 6.7, we see the increase of the coverage

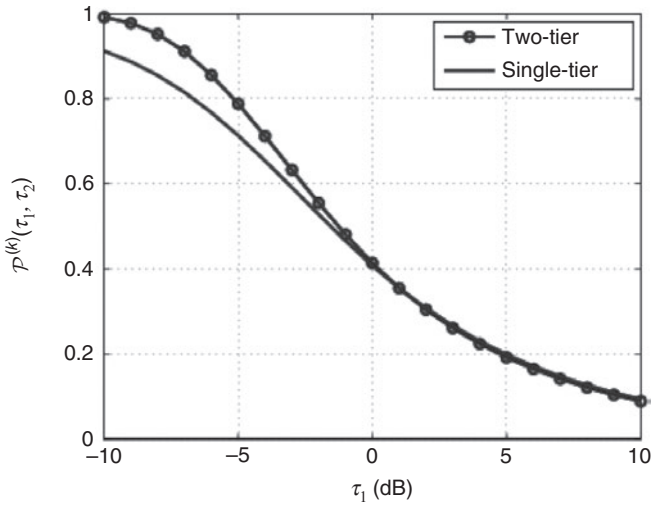

Figure 6.3 Coverage probability $\mathcal{P}^{(1)}$ as a function of τ_1 of a two-tier network with $\lambda_1 = \lambda_2/2$, $S_1 = 100P_2$, and $\tau_2 = 1$ dB, compared with $\mathcal{P}^{(1)}$ of a single-tier network model with equivalent process Θ and constant SINR thresholds, where $\alpha = 3$.

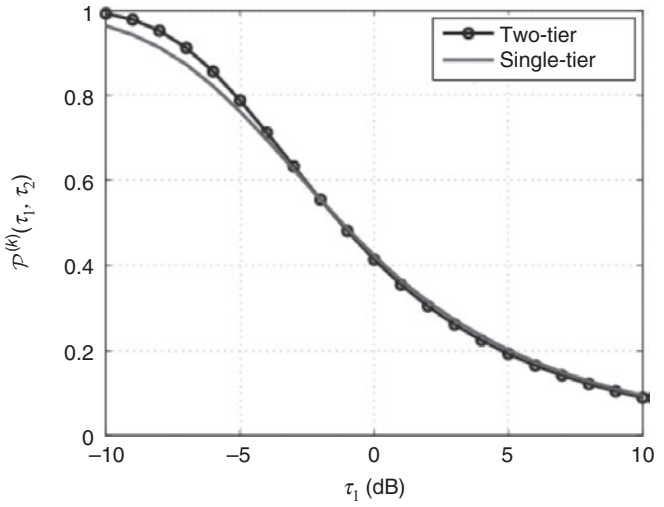

Figure 6.4 Coverage probability $\mathcal{P}^{(1)}$ as a function of τ_1 of a two-tier network with $\lambda_1 = \lambda_2/2$, $S_1 = 100P_2$, and $\tau_2 = -2$ dB, compared with $\mathcal{P}^{(1)}$ of a single-tier network model with equivalent process Θ and constant SINR thresholds, where $\alpha = 3$.

probability of the strongest signal induced by the interference cancellation $\Delta_{IC}^{(2)}(\tau, \epsilon)$ or signal combination $\Delta_{SC}^{(2)}(\tau, \epsilon)$, given by Eqs. (6.30) and (6.31), respectively, with the second-strongest station, that is $k = 2$. We consider two values of the path loss exponent $\alpha = 3$ and $\alpha = 5$ and the decoding threshold (for cancellation and combination) $\epsilon = -9.5424$ dB and -12.7875 dB. The plots of the probabilities show

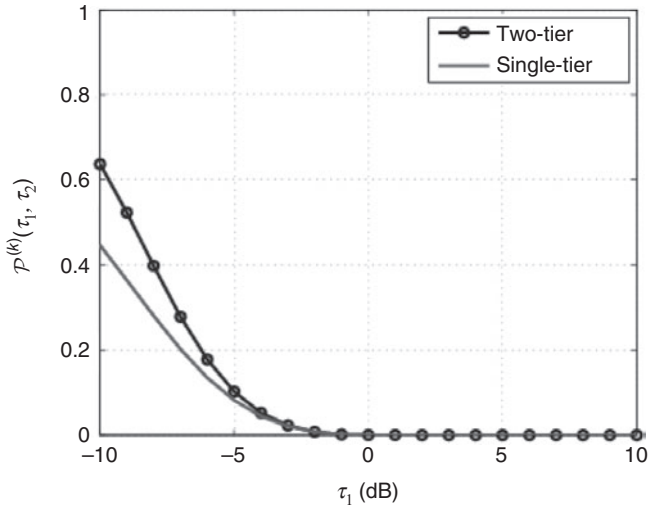

Figure 6.5 Coverage probability $\mathcal{P}^{(2)}$ as a function of τ_1 of a two-tier network with $\lambda_1 = \lambda_2/2$, $S_1 = 100P_2$, and $\tau_2 = -2$ dB, compared with $\mathcal{P}^{(2)}$ of a single-tier network model with equivalent process Θ and constant SINR thresholds, where $\alpha = 3$.

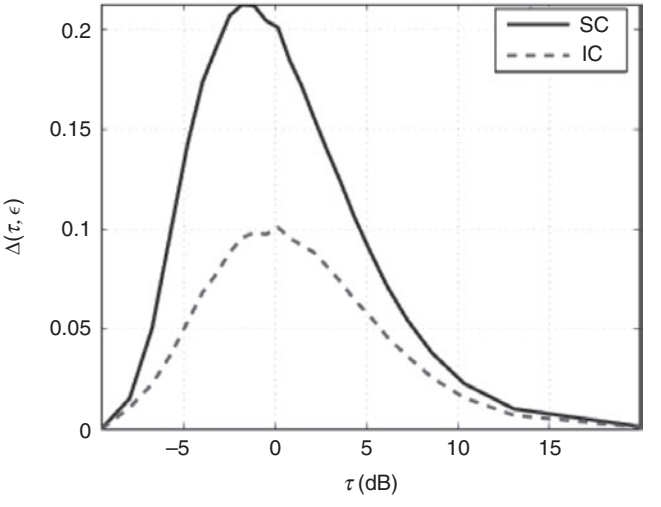

Figure 6.6 For $\alpha = 3$, the increase of the coverage probability of the strongest signal induced by the interference cancellation $\Delta^{(2)}_{IC}(\tau, \epsilon)$ or signal combination $\Delta^{(2)}_{SC}(\tau, \epsilon)$ with the second-strongest station assuming $\epsilon' = 0.1$ (or $\epsilon = -9.5424$ dB).

that the increase of coverage probability is largest around the threshold $\tau = 1$ (0 dB). As the path loss exponent α increases, the difference between the gain under IC and SC decreases (compare Figure 6.6 with 6.7). This is explained by realizing that the second-largest interfering signal, which is removed under IC, weakens as α increases.

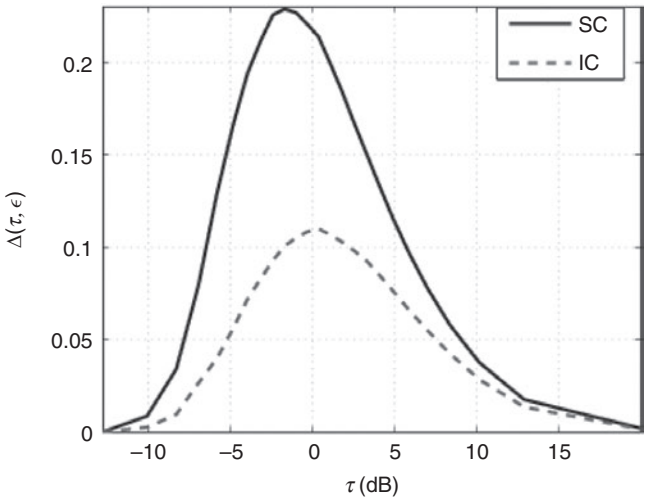

Figure 6.7 The same functions as on Figure 6.6 with $\alpha = 3$ and $\epsilon' = 0.05$ (or $\epsilon = -12.7875$ dB).

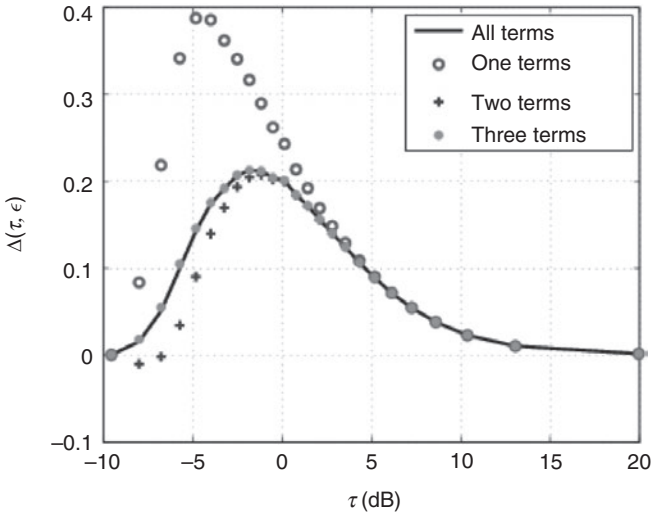

Figure 6.8 For $\alpha = 3$, the complete (10-term) expression and one-, two-, and three-term approximations for the $\Delta_{SC}^{(2)}$ probability with $\epsilon' = 0.1$ (or $\epsilon = -9.5424$ dB).

We recall that the decoding threshold ϵ allows mathematically the expansion of $\Delta_{IC}^{(k)}$ and $\Delta_{SC}^{(k)}$ in Eqs. (6.30) and (6.31) to be finite ones, so these expressions are complete when $i_{\max} < \infty$ number of terms are included. As ϵ decreases, more terms are needed for the exact expression, however, not all the integral terms are necessary to gain a good approximation (Figure 6.8). The numerical results support the intuition that as the threshold ϵ decreases the integral kernel approaches the singularity at $t_i = 0$ (for some integer $i \in [n]$), so the peak of the resulting integral increases.

6.1.5 Two-tier heterogeneous network with tier bias

We consider now a more developed example: a two-tier heterogeneous or "macro-pico" cellular network, where a user is served by the single strongest base station from the picotier, or by the strongest macro base station, or by the two strongest macro base stations with *maximal ratio combining* (MRC) transmitting in a *coordinated multipoint* (CoMP) scheme to be explained in what follows. The selection of the serving base station or pair of base stations is made on the basis of received reference symbol powers, modulated by the range expansion bias (REB) of the picotier.[2]

The SINR threshold is usually set higher for the picotier than for the macrotier and a REB is employed to "boost" the *reference symbol received power* at the user terminal from picocell base stations for comparison with the reference symbol received power from macro base stations to determine the tier and base station that will serve this user. In other words, the REB-based cell association criterion attempts to increase the fraction of users served by the small-cell tier. However, owing to the large difference in transmit powers between the base stations of the two tiers, users that are served by the small-cell tier according to this criterion may not have a high-enough data traffic rate because of interference from nearby macro base stations. The solution usually employed to solve this issue in 3GPP-LTE is the use of periodic *almost-blank subframes*, during which the macro base stations transmit no data to reduce the interference to users served by small cells. We emphasize that almost-blank subframes are intended to improve the data rate for traffic payload and not necessarily coverage. On the other hand, a user that is not very close to a macro base station, while being just outside the coverage region of a small-cell base station, is likely to struggle to meet even the reduced SINR threshold for coverage by the macrotier because of interference from the small-cell tier. Further, this problem also affects the data rate for this user.

CoMP transmission from more than one selected macro base station to the user has been proposed in single-tier macrocellular networks to enhance traffic data rate at users, and is also applicable to a two-tier heterogeneous cellular network. Further, the simple and robust form of CoMP defined by simultaneous transmission of identical symbols from more than one macro base station followed by MRC at the user receiver is easily applicable to improve not only traffic data rate but also coverage.

We focus on coverage probability at an arbitrarily located user. For simplicity, we also ignore the thermal noise at the receiver of this user. More precisely the user will be covered by *two* base stations, provided that (a) they are the two strongest base stations from the same tier; (b) their received powers are within, say, ξ dB of one another; and (c) the sum of the STIR values from these two base stations (equivalently, the total SIR after MRC at the user receiver) exceeds some threshold. We restrict this capability to the macrotier, as it is reasonable to assume that the lower-cost pico base stations are incapable of the coordinated transmission required to support this feature.

[2] This is a popular architectural choice for next-generation wireless cellular networks such as Third-Generation Partnership Project long-term evolution (3GPP-LTE) Release 12 (see Appendix B) and later. The aim is to support densely located high-data-rate users with a dense tier of picocells with low power in relation to macrocells, while a baseline level of service, which is typically set at voice calling with handover support for mobile users, is provided via the macrotier.

Example 6.10 *Coverage in macro-pico network with CoMP and REB.* We consider a two-tier Poisson network and assume that the user cannot tell, from a received signal, whether the transmitting base station is a macro or pico base station. Further, let us assume that the user can store only the three highest received power measurements from all base stations of the overall two-tier network. Denote the (decreasing) ordered STIR values from tier i by $\mathsf{Z}'_{i,(1)} \geq \mathsf{Z}'_{i,(2)} \geq \ldots, i = 1, 2$, and the overall (decreasing) ordered STIR values for the HetNet by $\mathsf{Z}'_{(1)} \geq \mathsf{Z}'_{(2)} \geq \ldots$. The user communicates the three highest STIR values $(\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \mathsf{Z}'_{(3)})$ and the corresponding decoded reference symbols to any base station that it wants to camp on, whence this information moves up into the system, where the identities and tiers of the base stations using these reference symbols are known, and used to select the tier and base station (or pair of base stations) that will serve this user. We have five scenarios, labeled $s = 1, \ldots, 5$:

1. All three of the strongest overall STIR values $(\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \mathsf{Z}'_{(3)})$ are from the picotier (tier 2). We assume *no coordination across pico base stations*, so the user can be served by (equivalently, covered by) only one of these pico base stations, namely the one that is received with STIR value $\mathsf{Z}'_{2,(1)} \equiv \mathsf{Z}'_{(1)}$.
2. Two of the three strongest overall STIR values are from the picotier, and one is from the macrotier (tier 1). In this case, the base station can be served by either the strongest-STIR pico base station or the strongest-STIR macro base station. The choice of serving tier (and therefore the serving base station) is made after accounting for the REB $\theta > 1$ of the picotier: the serving base station is the strongest-STIR macro base station if $\mathsf{Z}'_{1,(1)} > \theta \mathsf{Z}'_{2,(1)}$, and the strongest-STIR pico base station otherwise.
3. Of the three strongest overall STIR values, one is from the picotier and two are from the macrotier, but the gap between the two strongest macro STIRs exceeds the margin $\xi > 1$ for joint coverage, so $\mathsf{Z}'_{1,(1)} > \xi \mathsf{Z}'_{1,(2)}$. In this case, the choice of serving base station is again between the strongest-STIR macro and the strongest-STIR pico base stations, using the REB θ as in Scenario 2.
4. Of the three strongest overall STIR values, one is from the picotier and two are from the macrotier, with both macro base stations having STIR values within the margin ξ, so $\mathsf{Z}'_{1,(1)} \leq \xi \mathsf{Z}'_{1,(2)}$. In this case, the choice of serving base station is either the strongest-STIR pico base station or the pair of strongest-STIR and second-strongest-STIR macro base stations (combined with MRC at the user receiver), with the latter choice if $\mathsf{Z}'_{1,(1)} + \mathsf{Z}'_{1,(2)} > \theta \mathsf{Z}'_{2,(1)}$ and the former otherwise.
5. All three of the strongest overall STIR values are from the macrotier. In this case, the user will be served by the pair of strongest-STIR and second-strongest-STIR macro base stations if $\mathsf{Z}'_{1,(1)} \leq \xi \mathsf{Z}'_{1,(2)}$, and by only the strongest-STIR macro base station otherwise.

We note that the above criteria are similar to, but different from, the selection rule for the set of cooperating base stations (Nigam, Minero, & Haenggi 2015), where, say, the three base stations with largest-sum STIR would be selected to serve the user without using the REB to bias the selection in favor of the picotier.

Remark 6.1 Discussion of the roles of θ and ξ. The role of θ, the REB of the picotier, is clear: it is usually chosen large (6 dB or greater) to *offload* as much traffic as possible from the relatively sparse macrotier onto the relatively dense picotier. The role of the macrotier CoMP parameter ξ is more subtle. It is the maximum difference (in decibels) between the STIR of the strongest and second-strongest macro base stations. Consequently, if ξ is chosen to be large, then the second-strongest macro base station may be considerably weaker than the strongest macro base station, in which case its contribution to the MRC at the user receiver is minimal. Further, the use of MRC ties up resources at the second-strongest macro base station even though this base station cannot contribute much to MRC at the user. Then it seems reasonable to choose ξ relatively small, say 3 dB or less. On the other hand, choosing a relatively small ξ means that the second-strongest macro base station is comparable in strength with the strongest macro base station. This in turn suggests using the second-strongest macro base station to serve the user in addition to the strongest macro base station, so as not to turn the former into a strong interferer when the user receives from the latter. This is accomplished via CoMP. To summarize: θ is chosen relatively large, ξ relatively small: $\xi < \theta$.

Recall that coverage is defined by the STIR (not received power) value exceeding some threshold (which in general is different for each serving tier). The scenarios described in Example 6.10 define a mapping $(J, \mathsf{Z}'_J) = g(\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \mathsf{Z}'_{(3)}, J_1, J_2, J_3)$ from the three largest STIR values $\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \mathsf{Z}'_{(3)}$ and the corresponding tiers J_1, J_2, J_3 of the base stations yielding them, to J, the serving tier, and Z'_J, the STIR from the serving base station in tier J (or serving pair of base stations, if $J = 1$ and the strongest two macro base stations are selected to serve the user). This mapping is explicitly described in Table 6.1.

We let t_1 and t_2, respectively, denote the coverage thresholds on STIR from the macrotier (whether the user is served by a single macro base station or a pair of macro base stations using MRC) and the picotier. Then the coverage probability is computed over the five scenarios in Table 6.1 as

$$\mathbb{P}\{Z_J > t_J\} = \sum_{s=1}^{5} \mathcal{P}_s, \tag{6.33}$$

where for $s = 1, \ldots, 5$ (first column of Table 6.1),

$$\mathcal{P}_s = \sum_{(a,b,c) \text{ in Scenario } s} \mathbb{P}\{(J_1, J_2, J_3) = (a, b, c)\}\, \mathcal{P}^s_{abc}, \tag{6.34}$$

and for each $(a, b, c) \in \{1, 2\}^3$ belonging to Scenario s (second column of Table 6.1),

$$\mathcal{P}^s_{abc} = \mathbb{P}\{Z_J > t_J \mid (J_1, J_2, J_3) = (a, b, c)\} \tag{6.35}$$

is computed by integrating the joint probability density function of $(Z_{(1)}, Z_{(2)}, Z_{(3)})$ conditioned on $(J_1 = a, J_2 = b, J_3 = c)$ over the region defined by the third column of Table 6.1. This joint probability density function is obtained using results from Section 6.1.2, as shown in what follows.

Table 6.1. Explicit description of the mapping $(\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \mathsf{Z}'_{(3)}, J_1, J_2, J_3) \mapsto (J, \mathsf{Z}'_J)$

s	(a,b,c)	$(J, \mathsf{Z}'_J) = g(\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \mathsf{Z}'_{(3)}, J_1 = a, J_2 = b, J_3 = c)$
1	$(2,2,2)$	$(2, \mathsf{Z}'_{(1)})$
2	$(2,2,1)$ $(2,1,2)$	$(2, \mathsf{Z}'_{(1)})$
	$(1,2,2)$	$\begin{cases} (2, \mathsf{Z}'_{(2)}), & \mathsf{Z}'_{(2)} > \frac{\mathsf{Z}'_{(1)}}{\theta}, \\ (1, \mathsf{Z}'_{(1)}), & \text{otherwise.} \end{cases}$
3	$(1,1,2)$	$\mathbf{1}_{(\xi,\infty)}(\frac{\mathsf{Z}'_{(1)}}{\mathsf{Z}'_{(2)}}) \times \begin{cases} (2, \mathsf{Z}'_{(3)}), & \mathsf{Z}'_{(3)} > \frac{\mathsf{Z}'_{(1)}}{\theta}, \\ (1, \mathsf{Z}'_{(1)}), & \text{otherwise.} \end{cases}$
	$(1,2,1)$	$\mathbf{1}_{(\xi,\infty)}(\frac{\mathsf{Z}'_{(1)}}{\mathsf{Z}'_{(3)}}) \times \begin{cases} (2, \mathsf{Z}'_{(2)}), & \mathsf{Z}'_{(2)} > \frac{\mathsf{Z}'_{(1)}}{\theta}, \\ (1, \mathsf{Z}'_{(1)}), & \text{otherwise.} \end{cases}$
	$(2,1,1)$	$\mathbf{1}_{(\xi,\infty)}(\frac{\mathsf{Z}'_{(2)}}{\mathsf{Z}'_{(3)}}) \times (2, \mathsf{Z}'_{(1)})$
4	$(1,1,2)$	$\mathbf{1}_{(0,\xi)}(\frac{\mathsf{Z}'_{(1)}}{\mathsf{Z}'_{(2)}}) \times \begin{cases} (2, \mathsf{Z}'_{(3)}), & \mathsf{Z}'_{(3)} > \frac{\mathsf{Z}'_{(1)}+\mathsf{Z}'_{(2)}}{\theta}, \\ (1, \mathsf{Z}'_{(1)} + \mathsf{Z}'_{(2)}), & \text{otherwise.} \end{cases}$
	$(1,2,1)$	$\mathbf{1}_{(0,\xi)}(\frac{\mathsf{Z}'_{(1)}}{\mathsf{Z}'_{(3)}}) \times \begin{cases} (2, \mathsf{Z}'_{(2)}), & \mathsf{Z}'_{(2)} > \frac{\mathsf{Z}'_{(1)}+\mathsf{Z}'_{(3)}}{\theta}, \\ (1, \mathsf{Z}'_{(1)} + \mathsf{Z}'_{(3)}), & \text{otherwise.} \end{cases}$
	$(2,1,1)$	$\mathbf{1}_{(0,\xi)}(\frac{\mathsf{Z}'_{(2)}}{\mathsf{Z}'_{(3)}}) \times \begin{cases} (2, \mathsf{Z}'_{(1)}), & \mathsf{Z}'_{(1)} > \frac{\mathsf{Z}'_{(2)}+\mathsf{Z}'_{(3)}}{\theta}, \\ (1, \mathsf{Z}'_{(2)} + \mathsf{Z}'_{(3)}), & \text{otherwise.} \end{cases}$
5	$(1,1,1)$	$\begin{cases} (1, \mathsf{Z}'_{(1)}), & \mathsf{Z}'_{(1)} > \xi \mathsf{Z}'_{(2)}, \\ (1, \mathsf{Z}'_{(1)} + \mathsf{Z}'_{(2)}), & \text{otherwise.} \end{cases}$

In the notation of Corollary 4.2.5 and Eq. (5.55), the overall equivalent network $\Phi^* \equiv \Phi_1^* \cup \Phi_2^*$ is also a homogeneous Poisson point process with density $\lambda^* = \lambda_1^* + \lambda_2^*$, and the base stations of tier i can be modeled by independently thinning the points of Φ^* by the retention probability

$$p_i^* = \frac{\lambda_i^*}{\lambda^*}, \quad i = 1, 2. \tag{6.36}$$

We calculate the coverage probability in Eq. (6.33) on the equivalent network Φ^* and derived PD$(\delta, 0)$ process of STIRs, so for the vector $(Z^*_{(1)}, Z^*_{(2)}, Z^*_{(3)})$ of three strongest overall STIRs, and the corresponding tiers (J_1^*, J_2^*, J_3^*). Due to the independent thinning property shown in Eq. (6.36) on Φ^*, the probability that these STIRs came from base stations in tiers a, b, c respectively, with $a, b, c \in \{1, 2\}$, is $\mathbb{P}\{(J_1^*, J_2^*, J_3^*) = (a, b, c)\} = p_a^* p_b^* p_c^*$, *independently* of the values of $(Z^*_{(1)}, Z^*_{(2)}, Z^*_{(3)})$.

Thus the conditioning in Eq. (6.35) goes away and from Eq. (6.34) and the third column of Table 6.1, we have:

$$\mathcal{P}_1 = (p_2^*)^3\, \mathbb{P}\{\, Z^*_{(1)} > t_2 \,\}, \tag{6.37}$$

$$\mathcal{P}_2 = p_1^*(p_2^*)^2(\mathcal{P}^{[2]}_{221} + \mathcal{P}^{[2]}_{212} + \mathcal{P}^{[2]}_{122}), \tag{6.38}$$

$$\begin{aligned}
\mathcal{P}^{[2]}_{212} &= \mathcal{P}^{[2]}_{221} \\
&= \mathbb{P}\{\, Z^*_{(1)} > t_2\}, \\
\mathcal{P}^{[2]}_{122} &= \mathbb{P}\{\, Z^*_{(1)} > \max(t_1, \theta Z^*_{(2)}) \,\} \\
&\quad + \mathbb{P}\{\, Z^*_{(2)} > \max(t_2, Z^*_{(1)}/\theta) \,\},
\end{aligned}$$

$$\mathcal{P}_3 = (p_1^*)^2 p_2^*(\mathcal{P}^{[3]}_{211} + \mathcal{P}^{[3]}_{121} + \mathcal{P}^{[3]}_{112}), \tag{6.39}$$

$$\begin{aligned}
\mathcal{P}^{[3]}_{211} &= \mathbb{P}\{\, Z^*_{(2)} > \xi Z^*_{(3)},\, Z^*_{(1)} > t_2 \,\}, \\
\mathcal{P}^{[3]}_{121} &= \mathbb{P}\{\, Z^*_{(1)} > \max(\xi Z^*_{(3)}, \theta Z^*_{(2)}, t_1) \,\} \\
&\quad + \mathbb{P}\{\, Z^*_{(1)} > \xi Z^*_{(3)},\, Z^*_{(2)} > \max(t_2, Z^*_{(1)}/\theta) \,\}, \\
\mathcal{P}^{[3]}_{112} &= \mathbb{P}\{\, Z^*_{(1)} > \max(\xi Z^*_{(2)}, \theta Z^*_{(3)}, t_1)\} \\
&\quad + \mathbb{P}\{\, Z^*_{(1)} > \xi Z^*_{(2)},\, Z^*_{(3)} > \max(t_2, Z^*_{(1)}/\theta) \,\},
\end{aligned}$$

$$\mathcal{P}_4 = (p_1^*)^2 p_2^*(\mathcal{P}^4_{211} + \mathcal{P}^{[4]}_{121} + \mathcal{P}^{[4]}_{112}), \tag{6.40}$$

$$\begin{aligned}
\mathcal{P}^{[4]}_{211} &= \mathbb{P}\{\, Z^*_{(1)} > \max(t_2, \frac{Z^*_{(2)} + Z^*_{(3)}}{\theta}),\, \xi Z^*_{(3)} > Z^*_{(2)} \,\} \\
&\quad + \mathbb{P}\{\, Z^*_{(2)} + Z^*_{(3)} > \max(t_1, \theta Z^*_{(1)}),\, \xi Z^*_{(3)} > Z^*_{(2)} \,\}, \\
\mathcal{P}^{[4]}_{121} &= \mathbb{P}\{\, Z^*_{(2)} > \max(t_2, \frac{Z^*_{(1)} + Z^*_{(3)}}{\theta}),\, \xi Z^*_{(3)} > Z^*_{(1)} \,\} \\
&\quad + \mathbb{P}\{\, Z^*_{(1)} + Z^*_{(3)} > \max(t_1, \theta Z^*_{(2)}),\, \xi Z^*_{(3)} > Z^*_{(1)}\}, \\
\mathcal{P}^{[4]}_{112} &= \mathbb{P}\{\, Z^*_{(3)} > \max(t_2, \frac{Z^*_{(1)} + Z^*_{(2)}}{\theta}),\, \xi Z^*_{(2)} > Z^*_{(1)} \,\} \\
&\quad + \mathbb{P}\{\, Z^*_{(1)} + Z^*_{(2)} > \max(t_1, \theta Z^*_{(3)}),\, \xi Z^*_{(2)} > Z^*_{(1)}\},
\end{aligned}$$

$$\begin{aligned}
\mathcal{P}_5 = (p_1^*)^3 \Big(&\mathbb{P}\{\, Z^*_{(1)} > \max(t_1, \xi Z^*_{(2)}) \,\} \\
&+ \mathbb{P}\{\, Z^*_{(1)} + Z^*_{(2)} > t_1,\, \xi Z^*_{(2)} > Z^*_{(1)} \,\} \Big).
\end{aligned} \tag{6.41}$$

Each of these probabilities can be evaluated numerically using Corollary 6.1.5. (An alternative approach is presented in Section 6.2.3; compare with Eq. [6.58].)

6.2 STINR and Poisson-Dirichlet Process

We have seen the tractability that the Poisson assumption and singular path loss function can yield. Interesting and convenient mathematical properties abound as results can often be written down, sometimes in quite simple forms, for important characteristics. It is then perhaps not surprising that some of these results have been observed independently in other fields. In fact, the point process that we call the STINR process (or STIR with zero noise) in Section 6.1.1 happens to be exactly a type of Poisson-Dirichlet point process considered in physics, more precisely in probabilistic modeling of spin glasses (Panchenko 2013, Section 2.2). It is also a special case of a two-parameter family of processes known under the same name, whose another special case is the well-known, and historically considered first, Kingman's Poisson-Dirichlet process. The whole two-parameter family of Poisson-Dirichlet processes has been studied extensively since the 1990s but its connection to SINR problems is a more recent observation (Keeler & Błaszczyszyn 2014). We present a few properties of the STINR process, which can be directly borrowed from the existing results on Poisson-Dirichlet processes, without detailing their proofs.

6.2.1 Poisson-Dirichlet processes

There are two different but equivalent ways of defining (or constructing) the two-parameter Poisson-Dirichlet process. We present arguably the simpler one, involving random beta variables, which does not have apparent, immediate connection to the SINR setting considered so far in this chapter.

For two given parameters $0 \leq \alpha' < 1$ and $\theta' > -\alpha'$, we define sequence of random variables $\{V_i\}_{i\geq 1}$ by

$$V_1 \stackrel{\triangle}{=} U_1, \quad V_i \stackrel{\triangle}{=} (1-U_1)\dots(1-U_{i-1})U_i, \quad i \geq 2, \tag{6.42}$$

where $U_1, U_2, \dots$ are independent beta variables with U_i having $B(1-\alpha', \theta' + i\alpha')$ distribution, which is defined in Appendix C, Section C.2.8.

Consider the decreasing-order statistics $\{V_{(i)}\}_{i\geq 1}$ of $\{V_i\}_{i\geq 1}$ such that $V_{(1)} \geq V_{(2)} \geq \cdots$. The two-parameter Poisson-Dirichlet distribution with parameters α' and θ', which we abbreviate as $\mathrm{PD}(\alpha', \theta')$, is defined as the distribution of $\{V_{(i)}\}_{i\geq 1}$.

If we consider $\{V_{(i)}\}_{i\geq 1}$ or, equivalently, $\{V_i\}_{i\geq 1}$ as atoms of a point process, then we see $\mathrm{PD}(\alpha', \theta')$ as a distribution of a point process. This definition of the $\mathrm{PD}(\alpha', \theta')$ process does not yet allow one to discover any relation to the STIR process $\Psi' = \{Z'_i\}$ of Eq. (6.3) with $N = 0$, except that, as for Ψ', the values of the Poisson-Dirichlet process with probability one add up to one

$$\sum_{i=1}^{\infty} V_i = 1\,,$$

which follows easily from the representation in Eq. (6.42).

The advantage of this definition is also that it allows one, quite straightforwardly, to derive the product densities of the PD(α', θ') process by using Handa's Theorem 2.1 (Handa 2009). From this result follows our expression for the product densities of the STINR process given in Theorem 6.1.2 (see also the remark on its proof on page 100), after having identified PD(α', θ') with $\alpha' = 1/\delta$ (recall, α is the path loss exponent) and $\theta' = 0$ as the STIR process Ψ'.

6.2.2 STIR is a PD$(1/\delta, 0)$ point process

The missing relation stems from an equivalent representation of PD$(\alpha', 0)$ in terms of particular Poisson point process. For $0 < \alpha' < 1$ let $\Pi = \{P : P \in \Pi\}$ be a Poisson point process on $(0, \infty)$ of intensity measure $\Lambda(\mathrm{d}r) = r^{-\alpha'-1}\mathrm{d}r$. We define the random variable $I' \stackrel{\triangle}{=} \sum_{P\in\Pi} P$ and the point process

$$\mathcal{V} = \{V : V \in \mathcal{V}\} \stackrel{\triangle}{=} \{P/I' : P \in \Pi\}\,. \tag{6.43}$$

PROPOSITION 6.2.1 *The random variable I' is finite almost surely and the point process $\mathcal{V}$ is a PD$(\alpha', 0)$ process.*

The proof requires the more advanced probabilistic concept of Lévy subordinators, which we skip here (see Keeler & Błaszczyszyn 2014 for an outline and Pitman & Yor 1997 for the complete treatment). The other special case of Poisson-Dirichlet process, namely PD$(0, \theta')$, with $\theta' > 0$, has a similar representation with the Poisson point process on $(0, \infty)$ of intensity measure $\Lambda(\mathrm{d}r) = \theta' r^{-1}\mathrm{e}^{-t}\mathrm{d}r$. More subtly, a doubly stochastic Poisson construction is needed to represent the PD(α', θ') process in full generality.

COROLLARY 6.2.2 *The STIR process Ψ' given by Expression (6.3) with $N = 0$ is a PD$(1/\delta, 0)$ point process. Consequently, the decreasing-order statistics $\mathsf{Z}'_{(1)}, \mathsf{Z}'_{(2)}, \ldots$ of the STIR process have PD$(1/\delta, 0)$ distribution of the order statistics $V_{(1)}, V_{(2)}, \ldots$.*

Proof From Chapter 4, Section 4.2.1, the projection process $\Theta = \{\mathsf{L} : \mathsf{L} \in \Theta\} = \{1/[H_{\mathsf{x}}\ell(\|\mathsf{x}\|)]\}_{\mathsf{x}\in\Phi}$ with the singular path loss function $\ell(\cdot)$ given by Eq. (4.2) is an inhomogeneous Poisson point process with an intensity measure $Q((0, t]) = at^{\delta}$, where $a = \pi\lambda\mathbb{E}(H^{\delta})/\kappa^2$. By the Poisson mapping theorem, we see that the point process $\{\mathsf{L}^{-1} : \mathsf{L} \in \Theta\}$ is an inhomogeneous Poisson process with intensity measure $(2a/\alpha)t^{-1-\delta}\mathrm{d}t$ on $(0, \infty)$. By Eq. (6.3), the STIR process is constructed from $\{\mathsf{L}^{-1} : \mathsf{L} \in \Theta\}$ in the same way as $\mathcal{V}$ in Eq. (6.43) from the Poisson process of intensity $t^{-1-\alpha}\mathrm{d}t$. For the distribution of the STIR (and SIR) process, we can take $a = 1/\delta$ so that $2a/\alpha = 1$, because the constant a is irrelevant. The results then follow from Proposition 6.2.1. □

Remark 6.2 The definition in Eq. (6.42) is directly related to the concept of size-biased sampling. This is a random sampling of points V_i of the point process $\mathcal{V} = \{V_i\}$, without replacement, with a probability proportional to the value V_i. More specifically, we denote by $\tilde{V}_1$ the first point selected among $\mathcal{V}$, so that it is equal to V_i with probability V_i, so $\mathbb{P}\{\tilde{V}_1 = V_i\} = V_i$. Next we independently select a point $\tilde{V}_2$ among the remaining

points $\mathcal{V}\backslash\{\tilde{V}_1\}$, with probability $\mathbb{P}\{\tilde{V}_2 = V_i\} = V_i/(1-\tilde{V}_1)$, with $V_i \neq \tilde{V}_1$. We continue and obtain a random permutation $\{\tilde{V}_i\}$ of $\{V_i\}$ called *size-biased permutation*. A special property of the sequence $\{V_i\}$, defined by Expression (6.43), is that its size-biased permutation $\{\tilde{V}_i\}$ has the same probability distribution. This distributional invariance with respect to the size-biased sampling is even a property that characterizes Poisson-Dirichlet distributions $\mathrm{PD}(\alpha', \theta')$ within the class of infinite, sequences $\{V_i\}$ that satisfy Eq. (6.42) with U_i being independent, but not necessarily beta variables. The distribution of the output of the size-biased permutation does not depend on the original numbering of points. In fact, in our sampling we use the values of points, not their indexes. Consequently, the size-biased permutation of the order statistics $\{V_{(i)}\}$ has the same distribution as $\{V_i\}$. In other words, one obtains $\{V_{(i)}\}$ from $\{V_i\}$ by considering order-statistic permutation. Taking size-biased permutation of $\{V_{(i)}\}$ allows us to recover the original distribution of $\{V_i\}$.

Example 6.11 *Size-biased choice of the base station.* The size-biased permutation can be interpreted for the STIR process as follows. When (the user at) the typical location chooses its serving base station, assume that instead of looking for the strongest received signal $1/\mathsf{L}_{(1)}$, it makes a randomized decision, so selecting a base station x with a bias proportional to value of $1/\mathsf{L}_{\mathsf{x}}$. This implies that stronger base stations, in terms of received power values, have higher chances of being selected, but it is not necessarily the strongest station that is chosen. Then, by Corollary 6.2.2 and Remark 6.2, the STIR with respect to the selected station has the distribution of V_1, which is the beta distribution $B(1-\delta, \delta)$.

Example 6.12 *Hot spot of users with size-biased choice of the base station.* We assume that n users, number them by $1, \ldots, n$, are positioned in the same, typical location and experience the same propagation effects with respect to all stations. We further assume that user number 1 applies size-biased choice of the base station described in Example 6.11. Then, user number 2 applies independently the same randomized policy to choose one of the remaining bases stations, and similarly for the remaining users $3, \ldots, n$. Then the joint distribution of the STIR values experienced by these n users is equal to that of $(V_1, \ldots, V_n)$ in Eq. (6.42), given in terms of independent beta variables $U_1, \ldots, U_n$.

6.2.3 Further useful results

We list some further interesting results of $\mathrm{PD}(\delta, 0)$ for the network setting, some of which still hold if the noise-term $N \geq 0$ is reintroduced.

PROPOSITION 6.2.3 *For the STINR process Ψ' ($N \geq 0$), the ratios of the two successive STIR values (equivalently the ratios of the two successive received powers)*

$$\bar{R}_i \stackrel{\triangle}{=} \frac{\mathsf{Z}'_{(i+1)}}{\mathsf{Z}'_{(i)}} = \frac{\mathsf{L}_{(i)}}{\mathsf{L}_{(i+1)}} \tag{6.44}$$

have, respectively, $B(2i/\alpha, 1)$ distributions such that $\mathbb{P}(\bar{R}_i \leq r) = r^{i\delta}$ for $0 \leq r \leq 1$. Moreover, $\bar{R}_i$ are mutually independent.

For the next result, we introduce the random variables

$$A_i \stackrel{\triangle}{=} \frac{\mathsf{Z}'_{(1)} + \cdots + \mathsf{Z}'_{(i)}}{\mathsf{Z}'_{(i+1)}} = \frac{1/\mathsf{L}_{(1)} + \cdots + 1/\mathsf{L}_{(i)}}{1/\mathsf{L}_{(i+1)}} \tag{6.45}$$

$$\Sigma_i \stackrel{\triangle}{=} \frac{\mathsf{Z}'_{(i+1)} + \mathsf{Z}'_{i+2} + \cdots}{\mathsf{Z}'_i} = \frac{1/\mathsf{L}_{(i+1)} + 1/\mathsf{L}_{(i+2)} + \cdots}{1/\mathsf{L}_{(i)}}, \tag{6.46}$$

which are defined for integers $i \geq 1$. For $\zeta \geq 0$, we introduce the functions

$$\bar{f}_\alpha(\zeta) \stackrel{\triangle}{=} \delta \int_1^\infty \mathrm{e}^{-\zeta x} x^{-\delta-1} \mathrm{d}x, \tag{6.47}$$

$$\bar{g}_\alpha(\zeta) \stackrel{\triangle}{=} \Gamma(1-\delta)\zeta^\delta + \bar{f}_\alpha(\eta). \tag{6.48}$$

PROPOSITION 6.2.4 *For $N \geq 0$, consider the STINR process Ψ'. Then*

$$\frac{1}{\mathsf{Z}'_{(i)}} = 1 + A_{i-1} + \Sigma_i \tag{6.49}$$

where the random variable A_{i-1} is distributed as the sum of $i-1$ independent copies of A_1, with the characteristic function

$$\mathbb{E}[\mathrm{e}^{-\zeta A_{i-1}}] = (\bar{f}_\alpha(\zeta))^{i-1}, \tag{6.50}$$

and Σ_i is distributed as the sum of i independent copies of Σ_1, with the characteristic function

$$\mathbb{E}[\mathrm{e}^{-\zeta \Sigma_i}] = (\bar{g}_\alpha(\zeta))^{-i}. \tag{6.51}$$

Moreover, the random variables A_{i-1} and Σ_i are independent.

Remark 6.3 By observing that

$$\frac{1}{\Sigma_i} = \frac{1/\mathsf{L}_{(i)}}{1/\mathsf{L}_{(i+1)} + 1/\mathsf{L}_{(i+2)} + \cdots}, \tag{6.52}$$

Proposition 6.2.4 gives an alternative way (in case $N = 0$) to evaluate the coverage probability by the kth strongest signal with cancellation of all stronger signals, which can be compared with Example 6.4. In particular, we can derive an equivalent approach to $\mathcal{P}^{\{k\},k}(\tau) = \mathbb{P}\{1/\Sigma_i > \tau\} = \mathbb{P}\{\Sigma_i < 1/\tau\}$, previously evaluated in Eq. (6.24)

Similarly, the expression

$$\frac{(1 + A_{i-1})}{\Sigma_i} = \frac{1/\mathsf{L}_{(1)} + \cdots + 1/\mathsf{L}_{(i)}}{1/\mathsf{L}_{(i+1)} + 1/\mathsf{L}_{(i+2)} + \cdots}, \tag{6.53}$$

relates Proposition 6.2.4 to the coverage by the combination of the k strongest signals as in Example 6.6 with $\epsilon = 0$, when $N = 0$. In particular, one can derive an equivalent

approach to $\mathcal{P}_{SC}^{(k)}(\tau) = \mathbb{P}\{(1 + A_{i-1}))/\Sigma_i > \tau\}$ previously expressed via Eqs. (5.37), (6.27), and (6.31).

Proposition 6.2.4 leads also to a Laplace transform.

COROLLARY 6.2.5 *The inverse of the ith strongest STIR ($N = 0$) value, $1/\mathsf{Z}'_{(i)}$, has the Laplace transform*

$$\mathbb{E}[\mathrm{e}^{-\zeta/\mathsf{Z}'_{(i)}}] = \mathrm{e}^{-\zeta}(\bar{f}_\alpha(\zeta))^{i-1}(\bar{g}_\alpha(\zeta))^{-i}. \tag{6.54}$$

The Laplace transform of the inverse of the strongest SIR value,

$$\mathbb{E}[\mathrm{e}^{-\zeta\mathsf{Z}_{(1)}}] = \mathbb{E}[\mathrm{e}^{-\zeta\mathsf{Z}'_{(1)}/(1-\mathsf{Z}'_{(1)})}] = \mathrm{e}^{\zeta}\mathbb{E}[\mathrm{e}^{-\zeta/(1-\mathsf{Z}'_{(1)}))}],$$

has been previously studied (Błaszczyszyn et al. 2013).

Poisson-Dirichlet representation also offers another way for calculating the joint density of the order statistics of the STIR process previously expressed in Proposition 6.1.4 and Corollary 6.1.5. We define two-parameter family of functions

$$\rho_{\alpha',\theta'}(s) \stackrel{\triangle}{=} \mathbb{P}(V_{(1)} < 1/s) \tag{6.55}$$

where $V_{(1)}$ is the largest value of the PD(α', θ') process. This function can be computed with

$$\rho_{\alpha',\theta'}(s) \stackrel{\triangle}{=} \sum_{n=0}^{\infty} \frac{(-1)^n c_{n,\delta',\theta'}}{n!} h_{n,\delta',\theta'}(s), \tag{6.56}$$

where

$$c_{n,\delta',\theta'} = \prod_{i=1}^{n} \frac{\Gamma(\theta' + 1 + (i-1)\alpha')}{(\Gamma((1-\alpha')\Gamma(\theta' + i\alpha')},$$

$$h_{n,\delta',\theta'}(s) = \int_{\Delta_n} \prod_{i=1}^{n} \frac{\mathbf{1}_{[1,\infty)}(st'_i)}{t_i'^{\alpha'+1}} \left(1 - \sum_{j=1}^{n} t'_j\right)^{\theta'+\alpha' n-1} \mathrm{d}t'_1 \ldots \mathrm{d}t'_n, \tag{6.57}$$

and for $n = 1, 2, \ldots$, with $h_{n,\delta',\theta'}(s) = 0$ whenever $n > s$, which makes the right-hand side of Eq. (6.56) actually a finite sum.

Remark 6.4 When $\alpha' = \delta$ and $\theta' = 0$, then $V_{(1)}$ is equal in distribution to the strongest STIR value $\mathsf{Z}'_{(1)}$ and so Eq. (6.56) can be compared with Corollary 5.3.7 for $k = 1$ and $N = 0$. However, in the next result, PD(α', θ') is needed with two non-zero parameters to express the joint density of the order statistics of the STIR process.

PROPOSITION 6.2.6 *For the STIR process Ψ' ($N = 0$) and for each $m = 1, 2, \ldots$, the joint probability density of $\{\mathsf{Z}'_{(1)}, \ldots, \mathsf{Z}'_{(m)}\}$ is given by*

$$f_{m,\alpha}(t'_1, \ldots, t'_m) \stackrel{\triangle}{=} c_{m,\delta,0} \left(\prod_{i=1}^{m} t_i'^{-(\delta+1)}\right) \left(1 - \sum_{j=1}^{m} t'_j\right)^{2m/\alpha-1}$$
$$\times \rho_{\delta,2m/\alpha}\left(\frac{1 - \sum_{j=1}^{m} t'_j}{t'_m}\right) \mathbf{1}(t'_1, \ldots, t'_m)\mathbf{1}(t'_1 > \cdots > t'_n). \tag{6.58}$$

Finally, a mathematical fact is that the interference I, as defined by Eq. (6.2), and the noise N can be, at least theoretically, recovered from the STINR process.

PROPOSITION 6.2.7 *For the STINR process ($N \geq 0$),*

$$\frac{N}{I'} = \frac{1}{\sum_{i=1}^{\infty} \mathsf{Z}'_{(i)}} - 1, \tag{6.59}$$

and

$$N + I' = (L/a)^{-1/\delta}, \tag{6.60}$$

where the limit

$$L \stackrel{\triangle}{=} \lim_{i \to \infty} i(\mathsf{Z}'_{(i)})^{\delta}, \tag{6.61}$$

both exist almost surely and for all S means with $S \geq 1$.

The first statement of this result is just algebra, whereas the second can be proved using the same arguments used by Pitman and Yor (Pitman & Yor 1997, Proposition 10).

6.3 The Meta Distribution of the SIR

6.3.1 Motivation and definition

In the sequel, we neglect the noise at the user terminals, so that the SINR is replaced by the SIR for purposes of analysis. The SIR analysis in the previous sections gives information about the SIR distribution at the typical location. In the usual ergodic setting, this is the average of the coverage probabilities of all the users in the network. As such, it provides limited information about the performance seen by a specific user. For example, if $\mathcal{P}(\tau) = 0.8$, it could be that half the users are covered (that is, have SIR exceeding the threshold τ) with probability 0.75 and the other half are covered with probability 0.85. Or it could be that 20% of the users are covered with probability 0.20 whereas the remaining 80% are covered with probability 0.95. Clearly, the user experience will be quite different in the two cases, but the standard coverage probability metric $\mathcal{P}(\tau)$ does not capture the difference.

To obtain more fine-grained information, we start by defining the conditional SIR distribution given the base station point process.

DEFINITION 6.3.1 (Per-user conditional SIR distribution) *Let $\mathcal{U}$ denote the locations of the users in the network and for each $u \in \mathcal{U}$, let $\mathsf{x}_u \in \Phi$ denote its serving base station. For each $u \in \mathcal{U}$, u is said to be covered if $\mathrm{SIR}_u > \tau$ for a fixed threshold τ. The* per-user conditional SIR distribution *is the random variable defined as*

$$\mathcal{P}_s^{(u)}(\tau) \stackrel{\triangle}{=} \mathbb{P}\{\mathrm{SIR}_u > \tau \mid \Phi\} = \mathbb{P}\left\{ \frac{H_{\mathsf{x}_u} \ell(\mathsf{x}_u - u)}{\sum_{\mathsf{x} \in \Phi \setminus \{\mathsf{x}_u\}} H_{\mathsf{x}} \ell(\mathsf{x} - u)} > \tau \;\middle|\; \Phi \right\}, \quad u \in \mathcal{U}. \tag{6.62}$$

An illustration of $\mathcal{P}_s^{(u)}(\tau)$ is shown in Fig. 6.9.

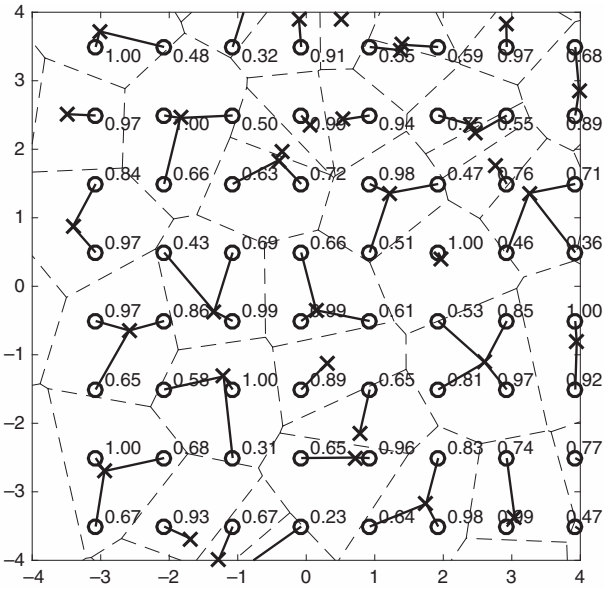

Figure 6.9 Realization of a Poisson cellular network of base stations Φ, marked by $\times$, with $\lambda = 1/2$. Users are located on a square lattice, marked by $\circ$. The number next to each user is its SIR distribution given Φ, that is, the distribution over the fading only, defined in Eq. (6.62), for $\tau = -3$ dB, Rayleigh fading, and path loss $\ell(x) = \|x\|^{-4}$. As expected, users close to a base station have a high reliability, whereas those near the Voronoi vertices suffer from a rather low one. The empirical average over the users is 0.68, which, as expected, is close to the standard coverage probability $\mathcal{P}(-3 \text{ dB}) = 0.70$ for these network parameters.

The standard coverage probability can then be expressed as the mean $\mathcal{P}(\tau) = \mathbb{E}[\mathcal{P}_s^{(u)}(\tau)]$. Owing to averaging with respect to the base station point process, it does not depend on u. In fact, all the random variables $\mathcal{P}_s^{(u)}(\tau)$ are identical in distribution, and we can simply write $\mathcal{P}_s(\tau)$ for a generic conditional SIR distribution. Moreover, because τ is assumed fixed and given, we can further simplify our notation and simply write $\mathcal{P}_s$ in place of $\mathcal{P}_s(\tau)$.

To capture the variability of the user experiences, we would like to find higher moments of $\mathcal{P}_s$, in particular the variance, and, ideally, the entire distribution, which we call the *meta distribution of the SIR*, because it is the distribution of a conditional distribution.

DEFINITION 6.3.2 (Meta distribution of the SIR) *The meta distribution is the two-parameter CCDF*

$$\bar{F}_{\mathcal{P}_s}(\tau, x) \stackrel{\Delta}{=} \mathbb{P}\{\mathcal{P}_s(\tau) > x\}, \quad \tau \in \mathbb{R}^+, \ x \in [0, 1]. \tag{6.63}$$

In this section, whenever possible without ambiguity, we drop the subscript $\mathcal{P}s$ and write $\bar{F}(\tau, x)$ for the meta distribution. The meta distribution gives detailed information about the network performance. For example, $\bar{F}(10, 0.8) = 0.5$ means that 50% of the users achieve an SIR of 10 dB with probability at least 0.8.

In the rest of this chapter we focus on networks where Φ forms a Poisson process of intensity λ, with iid Rayleigh fading on all links between base stations and users, and the singular path loss model in Eq. (4.2). A direct calculation of the meta distribution appears to be infeasible, but the moments $M_b \stackrel{\triangle}{=} \mathbb{E}[\mathcal{P}_s(\tau)^b]$ can be calculated, as we show in Section 6.3.2.

6.3.2 Moments of the conditional coverage probability

The moments of the conditional coverage probability (or the conditional SIR CCDF) $\mathcal{P}_s(\tau) = \mathbb{P}\{\mathrm{SIR}_u > \tau \mid \Phi\}$ for a given threshold τ are given in the Lemma 6.3.3.

LEMMA 6.3.3 (Moments of meta distribution) *The moments of the conditional coverage probability for Poisson cellular networks with Rayleigh fading are given by*

$$M_b = \frac{1}{{}_2F_1(b, -\delta; 1-\delta; -\tau)}, \quad b \in \mathbb{C}. \tag{6.64}$$

Proof Let $\mathsf{x}_0 = \arg\min\{\mathsf{x} \in \Phi \colon \|\mathsf{x}\|\}$ be the serving base station of (the user at) the typical location. Given the base station process Φ, the conditional coverage probabiliy is

$$\begin{aligned}\mathcal{P}_s(\tau) &= \mathbb{P}\Big\{H_{\mathsf{x}_0} > \tau \|\mathsf{x}_0\|^\alpha \sum_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} H_{\mathsf{x}} \|\mathsf{x}\|^{-\alpha} \;\Big|\; \Phi\Big\} \\ &= \prod_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} \frac{1}{1+\tau(\|\mathsf{x}_0\|/\|\mathsf{x}\|)^\alpha}.\end{aligned}$$

The bth moment follows as

$$M_b = \mathbb{E} \prod_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} \frac{1}{(1+\tau(\|\mathsf{x}_0\|/\|\mathsf{x}\|)^\alpha)^b}. \tag{6.65}$$

Instead of calculating this expectation in two steps as usual (first condition on $\|\mathsf{x}_0\|$ then take the expectation with respect to it), we use the pgfl of the RDP. Because Eq. (6.65) depends on the base station locations only through the relative distances, we can directly apply the pgfl of the RDP from Eq. (3.29) and obtain

$$M_b = \left\{ 1 + 2 \int_0^1 \left[1 - \frac{1}{(1+\tau r^\alpha)^b} \right] r^{-3} \, \mathrm{d}r \right\}^{-1}, \tag{6.66}$$

which can be expressed as Eq. (6.64). □

For $b \in \mathbb{N}$, the moment M_b equals the joint probability that the user is able to successfully receive b transmissions from the serving base station at a fixed bit rate (see the discussion in Section 5.1.1), which was calculated by Zhang and Haenggi (Zhang & Haenggi 2014, Theorem 2).

6.3.3 Exact expression, bounds, and beta approximation

Using the Gil-Pelaez inversion theorem, we obtain an exact integral expression for the meta distribution.

THEOREM 6.3.4 (Meta distribution) *The SIR meta distribution for Poisson cellular networks is given by*

$$\bar{F}(\tau,x) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\Im(\mathrm{e}^{-\mathrm{j}t\log x}M_{\mathrm{j}t})}{t}\,\mathrm{d}t, \quad \mathrm{j}=\sqrt{-1}. \tag{6.67}$$

Proof Let $X \stackrel{\triangle}{=} \log \mathcal{P}_{\mathrm{s}}(\tau)$. The characteristic function of X is

$$\varphi_X(t) \stackrel{\triangle}{=} \mathbb{E}[\mathrm{e}^{\mathrm{j}tX}] = \mathbb{E}[\mathcal{P}_s(\tau)^{\mathrm{j}t}] = M_{\mathrm{j}t}, \quad t\in\mathbb{R},$$

where $M_{\mathrm{j}t}$ is given by Eq. (6.64). Then by the Gil-Pelaez theorem (Gil-Pelaez 1951), the CCDF of X is given by

$$\bar{F}_X(x) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \frac{\Im(\mathrm{e}^{-\mathrm{j}tx}M_{\mathrm{j}t})}{t}\,\mathrm{d}t, \tag{6.68}$$

and the result follows since $\mathbb{P}\{\mathcal{P}_{\mathrm{s}}(\tau) > x\} = \mathbb{P}\{\log \mathcal{P}_{\mathrm{s}}(\tau) > \log x\}$. □

The complete meta distribution for $\alpha = 4$ is shown in Fig. 6.10. Interestingly, there appears to be a critical value τ_{c} of τ where the probability density $f(\tau,0) = -\partial\bar{F}(\tau,x)/\partial x|_{x=0}$ changes from 0 to ∞. Hence for $\tau > \tau_{\mathrm{c}}$, there is a rapidly increasing fraction of users with very low reliability.

Bounds

Although the meta distribution can be efficiently calculated numerically, it is often preferable to work with closed-form bounds. Because the moments are in simple form,

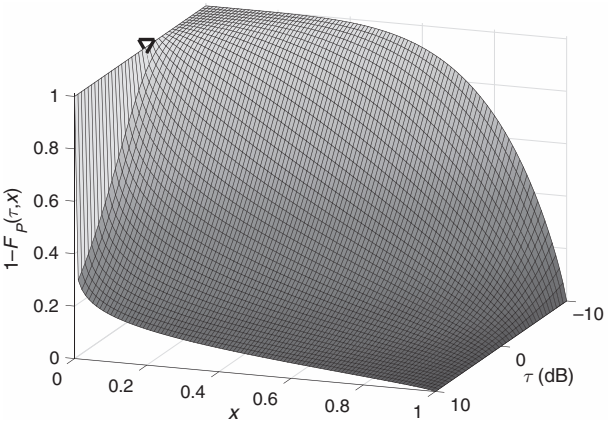

Figure 6.10 Meta distribution for $\alpha = 4$. The triangle marks the critical value $\tau_{\mathrm{c}} = -0.8$ dB, where the slope $\partial\bar{F}(\tau,x)/\partial x$ at $x = 0$ changes from 0 to $-\infty$.

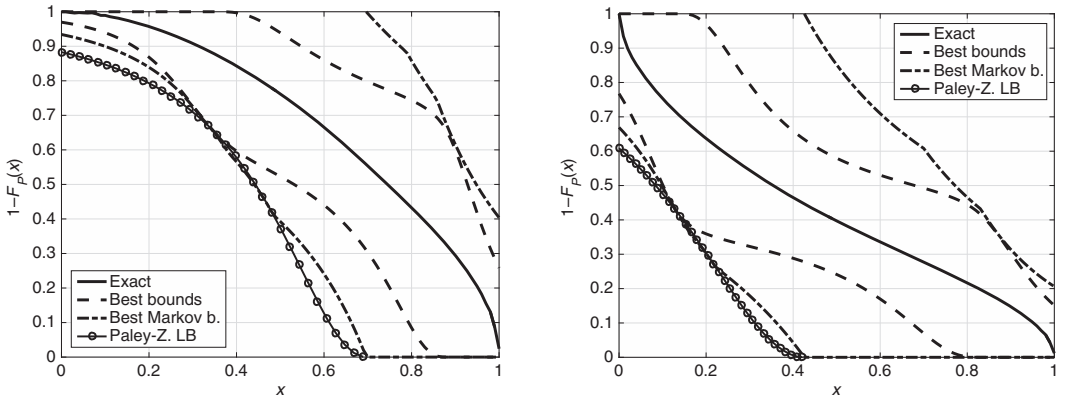

Figure 6.11 Meta distribution for $\tau = -3$ dB (*left*) and $\tau = 3$ dB (*right*), both for $\alpha = 4$, and its bounds. The best bounds are those obtained from Rázc et al. (2006) given the first four moments, and the best Markov bounds (best Markov b.) is obtained by choosing for each x the tightest bound from Eq. (6.69) for $b = 1, 2, 3, 4$. Lastly, the Paley-Zygmund lower bound (Paley-Z. LB) (Eq. [6.70]) is also shown.

standard bounding techniques (Markov, Chebyshev, and Paley-Zygmund) can readily be applied. The Markov bounds are

$$1 - \frac{\mathbb{E}[(1 - \mathcal{P}_s(\tau))^b]}{(1-x)^b} < \bar{F}(\tau, x) \leq \frac{M_b}{x^b}, \quad b > 0,\ x \in [0,1], \tag{6.69}$$

and the Paley-Zygmund lower bound is

$$\bar{F}(\tau, xM_1) \geq \frac{(1-x)^2}{1 - M_1^{p(1-\delta)} + (1-x)^2}, \quad x \in [0,1]. \tag{6.70}$$

In addition, the method from Rázc, Tari, & Telek (2006) can be used to find the best lower and upper bounds given the first k integer moments. Fig. 6.11 shows the meta distributions and its bounds for $\tau = \pm 3$ dB.

Approximation with the beta distribution

Since the meta distribution is supported on $[0, 1]$, a natural candidate for a simple approximate distribution is the beta distribution (see Appendix C, Section C.2.8). The probability density of a beta distributed random variable X with mean μ is

$$f_X(x) = \frac{x^{\frac{\mu(\beta+1)-1}{1-\mu}}(1-x)^{\beta-1}}{\mathrm{B}(\mu\beta/(1-\mu), \beta)}, \quad 0 \leq x \leq 1, \tag{6.71}$$

where B is the beta function:

$$\mathrm{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}, \quad \alpha > 0,\ \beta > 0.$$

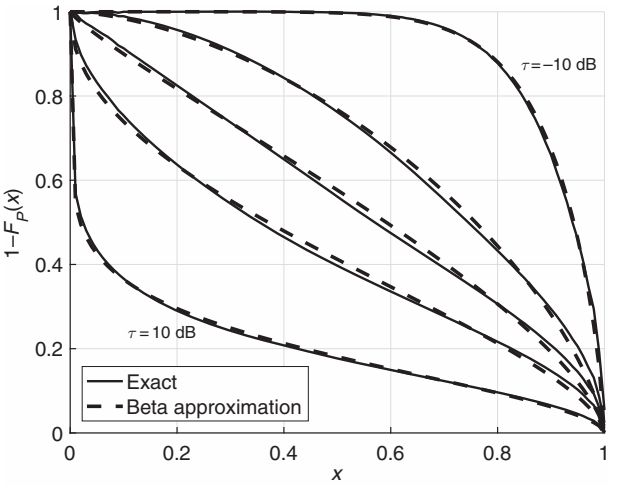

Figure 6.12 Meta distribution for $\alpha = 4$ for $\tau \in \{-10, -3, 0, 3, 10\}$ dB (from top to bottom) and the corresponding beta distribution.

The variance is given by

$$\operatorname{var} X = \frac{\mu(1-\mu)^2}{\beta + 1 - \mu}.$$

Matching mean and variance yields $\mu = M_1$ and

$$\beta = \frac{(\mu - M_2)(1-\mu)}{M_2 - \mu^2}.$$

Fig. 6.12 shows the exact CCDF and the beta approximation for $\tau \in \{-10, -3, 0, 3, 10\}$ dB for $\alpha = 4$. It turns out that the approximation is quite accurate, in particular for small and large values of τ.

As pointed out before, there appears to be a critical value of τ that marks a phase transition in the derivative of the meta CCDF at $x = 0$. Expressed in terms of the probability density $f(\tau, x) = -\partial \bar{F}(\tau, x)/\partial x$, this means that

$$f(\tau, 0) = \begin{cases} 0 & \tau < \tau_c \\ \infty & \tau > \tau_c \end{cases}.$$

From Fig. 6.12, τ_c is in between -3 and 0 dB, perhaps closer to 0 dB. This is in accordance with the value -0.8 dB, empirically found in the numerical evaluation shown in Fig. 6.10. The beta approximation exhibits the same phase transition, and its critical τ can be easily bound by numerically solving for the τ for which the exponent of x in (6.71) is zero, that is, solving

$$\frac{\mu(\beta+1) - 1}{1-\mu} = 0$$

for τ. This yields $\tau_c = -0.81$ dB for $\alpha = 4$ and $\tau_c = -3.37$ dB for $\alpha = 3$. The corresponding values of the (mean) coverage probability are $\mathcal{P}(\tau_c) = 0.60$ for $\alpha = 4$ and $\mathcal{P}(\tau_c) = 0.54$ for $\alpha = 3$.

The meta distribution analysis has been extended to device-to-device networks underlying cellular networks (Salehi, Mohammadi, & Haenggi 2017), cellular networks with power control in the uplink and downlink (Wang, Haenggi, & Tan 2017), millimeter wave communication (Deng & Haenggi 2017), and successive interference cancellation (Wang, Cui, Haenggi, & Tan 2017). Also, the meta distribution has been used to determine the *spatial outage capacity*, which is the maximum density of links in a network that satisfy a certain outage constraint (Kalamkar & Haenggi 2017).

6.4 Bibliographic Notes

The book on point process theory by Daley & Vere-Jones (2008) and the standard text on stochastic geometry (Chiu et al. 2013) give accessible introductions to factorial moment measures. Keeler, Błaszczyszyn, & Karray (2013) presented *k*-coverage results for a single-tier Poisson model, which was developed further by Błaszczyszyn & Keeler (2015) to give a factorial moment framework with various results, including generalized versions of previous SIC results in the zero-noise case, originally derived by Zhang & Haenggi (2013). The two-parameter Poisson-Dirichlet process is a well-studied object in probability, and its key results were derived by Pitman & Yor (1997). Observing the connection between this process and the SINR process, Keeler & Błaszczyszyn (2014) adapted the results by Pitman & Yor (1997) as well as those by Handa (2009) for the network setting. Some variations of these results were independently derived (Behnad, Wang, & Akhtar 2015). Haenggi introduced the meta distribution in 2016 (Haenggi 2016).

7 Downlink SINR: Further Extensions

In Chapter 6 we derive numerous results for the Poisson network model with a singular path loss function. In this chapter we extend these results in different directions.

7.1 SINR Analysis with General Path Loss Models

In this section, we discuss the sensitivity of the general properties of the interference and SIR to the path loss model. First, we summarize the results for the standard (singular) path loss model $\ell(\mathsf{x}) = (\kappa\|\mathsf{x}\|)^{-\alpha}$. For simplicity, we take $\kappa = 1$ in this chapter. We also ignore noise and analyze only the SIR.

7.1.1 Interference and SIR for the singular path loss model

The statistical properties of the interference power and SIR are governed by two sources of randomness, namely the point process of base station locations and the fading. As a result, the value of the CCDFs $F_I(x)$ and $F_{\mathrm{SIR}}(x)$ for $0 < x < \infty$ are determined by both the base station point process and the fading. However, the asymptotic behavior for $x \to 0$ or $x \to \infty$ may be determined by only one source of randomness. If this is the case, we obtain useful insight into what determines the main properties of interference power and SIR.

For the asymptotics, it is helpful to think about the main reasons for a quantity to become vanishingly small or extremely large.

Interference power

In a Poisson field of interference with singular path loss model, it is well known that the mean (and thus the higher moments) do not exist because an interferer has a relatively high probability to be very close and thus very strong. For path loss exponents $\alpha \leq 2$, the interference distribution is degenerate because $I = \infty$ almost surely.

In the cellular setting, this is less clear because the strongest interferer is the second-closest point of the Poisson point process. We have the mean interference power given by

$$\begin{aligned}
\mathbb{E}[I] &\stackrel{\triangle}{=} \mathbb{E} \sum_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} H_{\mathsf{x}} \|\mathsf{x}\|^{-\alpha} \\
&= \mathbb{E}[\mathbb{E}(I \mid \mathsf{x}_0)] \\
&\stackrel{\text{(a)}}{=} \mathbb{E}\left[\frac{2\pi\lambda}{\alpha-2}\|\mathsf{x}_0\|^{2-\alpha}\right] \\
&= \int_0^\infty \frac{(2\pi\lambda)^2}{\alpha-2} x^{3-\alpha} \mathrm{e}^{-\lambda\pi x^2} \mathrm{d}x \\
&= -(\lambda\pi)^{\alpha/2}\Gamma(1-\alpha/2) \\
&= (\lambda\pi)^{\alpha/2}\frac{2\Gamma(2-\alpha/2)}{\alpha-2}, \quad 2<\alpha<4.
\end{aligned} \tag{7.1}$$

Note that (a) follows from Campbell's formula applied to a Poisson point process with intensity function $\lambda(\mathsf{x}) = \lambda\mathbf{1}(\|\mathsf{x}\| > \|\mathsf{x}_0\|)$. For $\alpha \le 2$ and $\alpha \ge 4$, the mean is infinite. In fact, the condition $\alpha < 4$ for a finite mean is also the condition for the mean interference from the nearest interferer to be finite (Haenggi 2012, Section 5.1.4), that is, if I_1 is the interference from the nearest interferer, $\mathbb{E}I_1 < \infty \Leftrightarrow \alpha < 4$.

It is also interesting to note that the mean interference is not proportional to λ. It grows more quickly with λ because the interferers not only get denser with λ but they also get closer to the origin since $\mathbb{E}\|\mathsf{x}_0\| \propto \lambda^{-1/2}$.

The Laplace transform can be calculated in integral form by considering the interference from all nodes other (further) than $\|\mathsf{x}_0\|$ and deconditioning with respect to $\|\mathsf{x}_0\|$. From

$$\begin{aligned}
\mathcal{L}_I(s \mid \|\mathsf{x}_0\|) &= \exp\left(-\int_{\|\mathsf{x}_0\|}^\infty \frac{s}{s+r^\alpha} 2\lambda\pi r \mathrm{d}r\right) \\
&= \exp\left(-\lambda\pi\|\mathsf{x}_0\|^2 ({}_2F_1(1,-\delta;1-\delta;-s\|\mathsf{x}_0\|^{-\alpha})-1)\right),
\end{aligned}$$

we obtain

$$\mathcal{L}_I(s) = \int_0^\infty \exp\left(-\lambda\pi x^2 {}_2F_1(1,-\delta;1-\delta;-sx^{-\alpha})\right) 2\lambda\pi x \,\mathrm{d}x. \tag{7.2}$$

This expression cannot be simplified further but is easily evaluated numerically. The mean in Eq. (7.1) can be retrieved from

$$\mathbb{E}[I] = -\frac{\mathrm{d}}{\mathrm{d}s}\log\mathcal{L}_I(s)\Big|_{s=0} = -(\lambda\pi)^{\alpha/2}\Gamma(1-\alpha/2), \quad 2<\alpha<4.$$

The variance, however, is infinite for all α because

$$\operatorname{var} I = \frac{\mathrm{d}^2}{\mathrm{d}s^2}\log\mathcal{L}_I(s)\Big|_{s=0} = \infty.$$

The form of the Laplace transform also indicates that the interference does not have a stable distribution, as it does in a Poisson field of interferers (without restrictions on the location of the interferers).

SIR and coverage probability

First let us focus just on the signal power $S = H_{\mathsf{x}_0}\|\mathsf{x}_0\|^{-\alpha}$. Its distribution can be expressed in integral form (and for $\alpha = 4$, in closed form). The tail $\mathbb{P}(S > \tau)$ is easily calculated using the fact that $F_{\|\mathsf{x}_0\|}(x) \sim \lambda\pi x^2$, $x \to 0$:

$$\mathbb{P}(S > \tau) = \mathbb{P}(\|\mathsf{x}_0\| < (H_{\mathsf{x}_0}/\tau)^{\delta/2}) \sim \lambda\pi\mathbb{E}[H_{\mathsf{x}_0}^{\delta}]\theta^{-\delta} \tag{7.3}$$

This shows that $\mathbb{E}\,S = \infty$ for $\alpha \geq 2$ and that the reason for S to get really large is the proximity of the serving base station, and not the fading. The fading affects the tail only through its δth moment. In fact Eq. (7.3) holds for general stationary point process models.

Let $\mathrm{SIR}_1 = S/I_1$, where $S = H_{\mathsf{x}_0}\|\mathsf{x}_0\|^{-\alpha}$ and $I_1 = H_{\mathsf{x}_1}\|\mathsf{x}_1\|^{-\alpha}$ is the interference from the nearest interferer only. Also let $\xi = (\|\mathsf{x}_0\|/\|\mathsf{x}_1\|)^{\alpha} < 1$ be the distance ratio raised to the αth power. The probability density of ξ is known (Zhang & Haenggi 2014, Lemma 3) to be $f_\xi(y) = \delta y^{\delta-1}\mathbf{1}(y \in [0,1])$. With that, we obtain

$$\begin{aligned}
\bar{F}_{\mathrm{SIR}_1}(\tau) &\stackrel{\triangle}{=} \mathbb{P}\left\{\frac{S}{I_1} > \tau\right\} && (7.4)\\
&= \mathbb{P}\left\{\frac{H_{\mathsf{x}_0}}{H_{\mathsf{x}_1}} > \tau\xi\right\}\\
&= \mathbb{E}\left[\frac{1}{1+\tau\xi}\right]\\
&= {}_2F_1(1,\delta;1+\delta;-\tau). && (7.5)
\end{aligned}$$

It is interesting to compare this result with the coverage probability $\bar{F}_{\mathrm{SIR}}(\tau)$ when all interferers are taken into account. Including δ explicitly as a parameter, we have the relationship

$$\bar{F}_{\mathrm{SIR}}(\tau,\delta) \equiv \frac{1}{\bar{F}_{\mathrm{SIR}_1}(\tau,-\delta)}.$$

For $\tau \to \infty$, $\bar{F}_{\mathrm{SIR}_1}(\tau) \sim \mathrm{sinc}(\delta)^{-1}\tau^{-\delta}$ (Temme 2003, Eq. [9]), hence $\bar{F}_{\mathrm{SIR}}(\tau) \sim \mathrm{sinc}(\delta)\tau^{-\delta}$ (see also Section 8.2). Hence both SIR and SIR_1 scale in the same way as S itself, which indicates that the main reason for the SIR to become large is a large signal power S and not a small interference power. Also, $\mathbb{E}\,\mathrm{SIR}_1 = \mathbb{E}\,\mathrm{SIR} = \infty$, and the ratio of the two CCDFs approaches the constant $\mathrm{sinc}^2(\delta)$.

As $\tau \to 0$, a Taylor series expansion yields

$$\frac{F_{\mathrm{SIR}}(\tau)}{F_{\mathrm{SIR}_1}(\tau)} \to \frac{1+\delta}{1-\delta}.$$

The two asymptotic ratios quantify how accurate a nearest-interferer approximation is. For $\delta \to 0$ (high α), the ratio approaches 1 for all τ, whereas for $\delta \to 1$ (α near 2), the ratio becomes infinite (or zero).

Laplace transform and coverage probability
In a network model where signal power and interference power are independent, and the signal is subject to Rayleigh fading, the Laplace transform of the interference power equals the coverage probability (CCDF of the SIR). For example, putting the intended transmitter at x and the receiver at o, we have

$$\mathbb{P}\{H_{\mathsf{x}}\ell(\mathsf{x}) > \tau I\} = \mathbb{E}\exp(-\tau I/\ell(\mathsf{x})) = \mathcal{L}_I(\tau/\ell(\mathsf{x})).$$

In a cellular network, this does not hold since signal and interference power are positively correlated.

More generally, the coverage probability is, in fact, the Laplace transform of another random variable—the interference-to-(average)-signal ratio (ISR), defined as follows:

$$\bar{\mathsf{ISR}} \triangleq \frac{I}{\mathbb{E}_{H_{\mathsf{x}}}(S)} = I/\ell(\mathsf{x}).$$

For exponential H_{x}, that is, for Rayleigh fading, the coverage probability is

$$\mathcal{P}(\tau) = \bar{F}_{\mathrm{SIR}}(\tau) = \mathbb{P}\left\{\frac{H_{\mathsf{x}}\ell(\mathsf{x})}{I} > \tau\right\} = \mathbb{P}\{H_{\mathsf{x}} > \tau I/\ell(\mathsf{x})\} = \mathbb{E}\mathrm{e}^{-\tau\bar{\mathsf{ISR}}} = \mathcal{L}_{\bar{\mathsf{ISR}}}(\tau).$$

This relationship holds in great generality whenever the desired signal is subject to Rayleigh fading. The ISR is also the key quantity in the approximative approach discussed in Section 8.2.

Even without fading, the Laplace transform of the ISR is a relevant quantity, since its region of convergence determines the asymptotic coverage probability for the no-fading case (Ganti & Haenggi 2016b).

7.1.2 Results for general path loss models

Bounded path loss

Some of the results in the previous section critically depend on the path loss model. In particular, the behavior of the tail of the SIR distribution changes significantly if a bounded path loss model is used. With bounded path loss, the only way the signal power can get (very) large is through favorable fading conditions. Hence the tail of the SIR follows the tail of the fading distribution (Haenggi & Ganti 2008, Section 5.4.2).

Also, in the bounded case, the interference power converges to a Gaussian random variable as the density increases (Ak, Inaltekin, & Poor 2016). This is not the case in the singular case, as Eq. (7.2) shows.

Multislope path loss

A multislope path loss model was proposed and analyzed in Zhang & Andrews (2015).

DEFINITION 7.1.1 (N-slope path loss model) *Let $N \in \mathbb{N}$, $\alpha = (\alpha_1, \ldots, \alpha_N) \in \mathbb{R}^N$ with $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_N$ and $r = (r_1, \ldots, r_N) \in \mathbb{R}^{N+1}$ with $0 = r_0 < r_1 < \cdots < r_N = \infty$. The N-slope path loss model is defined as*

$$\ell_N(\alpha, r; \mathsf{x}) \triangleq (\kappa_n\|\mathsf{x}\|)^{-\alpha_n}\mathbf{1}\{\mathsf{x} \in b(o, r_n) \setminus b(o, r_{n-1})\},$$

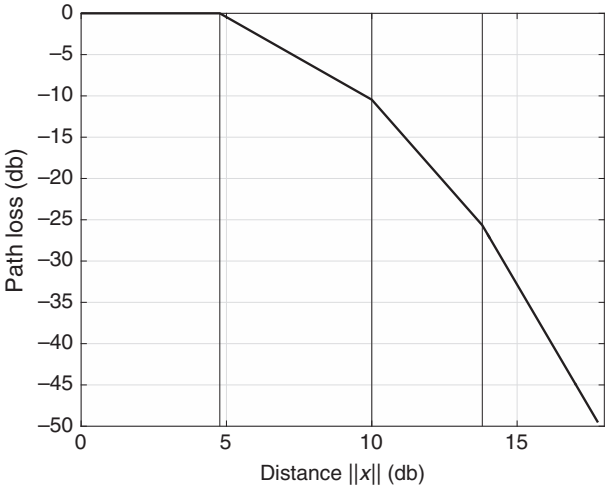

Figure 7.1 Multislope path loss function $\ell_4([0,2,4,6],[0,3,10,24,\infty];x)$. The vertical lines indicate the distances at which the slope changes.

where $K_1 \in \mathbb{R}^+$, $\kappa_1 = K_1^{-1/\alpha_1}$, *and*

$$\kappa_n = K_n^{-1/\alpha_n}, \quad K_n = K_1 \prod_{i=1}^{n} r_i^{\alpha_i - \alpha_{i-1}}, \quad n > 1,$$

to ensure continuity.

To keep the interference finite almost surely, $\alpha_N > 2$ is necessary.

The standard singular path loss law is retrieved by setting $N = 1$ and $\alpha_1 = \alpha$. A bounded path loss model is obtained by setting $\alpha_1 = 0$. Fig. 7.1 gives an example of a four-slope model.

The model is quite versatile but not fully tractable because the integral for the pgfl (see Eq. [3.30]) needs to be broken down according to the different pieces of the path loss function. Further, owing to the distances r, the network density has an impact on the SIR as soon as $N > 1$. That said, interesting observations are still possible. For example, the following order result holds for the coverage probability for arbitrary fading. Consider two path loss functions ℓ_N and ℓ_M, where we omit the dependence on the α and r vectors for brevity, and let $\mathcal{P}^N(\tau)$ and $\mathcal{P}^M(\tau)$ denote the corresponding coverage probabilities (SIR CCDFs). We have

$$(\forall x \geq 0)\, \ell_N(x) \geq \ell_M(x) \quad \Longrightarrow \quad (\forall \tau \geq 0)\, \mathcal{P}^N(\tau) \geq \mathcal{P}^M(\tau).$$

In other words, a uniformly better path loss function (higher path "loss" ℓ) results in uniformly higher coverage probability. This holds not just as an average over the point process, but for each realization.

Also, for any path loss function ℓ_N, the SIR coverage probability is uniformly decreasing as the network density λ increases. In contrast, if noise is included, there exists an optimum density that maximizes the coverage probability.

7.2 SINR Analysis for the Poisson Network with Advanced Signaling

The analysis in Chapter 5 implicitly assumes a single transmitting antenna and a single receive antenna on every link between a base station and a user terminal. We now extend this analysis to scenarios that are closer to those seen in real-world network deployments. We focus exclusively on the downlink, that is, transmissions from base stations to user terminals. In particular, we extend the analysis in two directions:

1. from a single transmit and single receive antenna (called *single-input single-output* or *SISO*) to multiple antennas at both the transmitter and the receiver (called *multiple-input multiple-output* or *MIMO*); and
2. from a single transmitting base station to multiple, cooperating, base stations transmitting jointly to the user terminal of interest. The latter is an example of a *coordinated multipoint* (*CoMP*) transmission scenario. Note that the transmission is called "coordinated" instead of just "joint" because the transmitting group of base stations may transmit different symbols, rather than all transmitting the same symbol.

It is worth noting that both these directions of research are highly active, with new results being added to the literature every week. In particular, SINR distributions at the user terminal have not been obtained for all the advanced signaling schemes that exist in the literature. This chapter aims to provide an overview of the stochastic geometry methods used to derive results for some basic signaling schemes. It is the authors' hope that the reader be able to apply these methods to investigate other signaling schemes on his or her own, and to follow such investigations when reported in the literature.

7.2.1 MIMO analysis

In MIMO transmission schemes, the user terminal is served by a single base station, using one or more than one transmit antenna, and the user terminal receives those transmissions with one or more than one receive antenna. (Note that this definition subsumes SISO as a special case of MIMO where there is a single transmit antenna and a single receive antenna.)

Serving base station selection

The serving base station for a *typical* user terminal is the base station whose "beacon" channel is received (without any special processing) at the user terminal with strongest average power. For a single-tier network, this is the scenario analyzed in Proposition 5.2.3. In other words, employing the notation of Proposition 5.2.3, such a deployment Φ of base stations with arbitrary iid shadow fading on the links from the base stations to the user terminal is equivalent to another deployment $\tilde{\Phi}$ of base stations without shadow fading, wherein the user terminal is served by its nearest base station in this equivalent deployment.

In a multitier HetNet, the candidate serving base station (from each accessible tier) for a given user terminal is chosen from among the base stations of that tier according to the

strongest average received power requirement, and then the actual serving base station for the user terminal is chosen from these candidate serving base stations as the one that is received with strongest average power (note that this average received power may be modified by a tier-dependent bias for the purposes of serving base station selection). At any rate, if we know that the serving base station comes from tier i, then it must be the base station in tier i that is nearest to the user terminal location. With the convention that the user is at the typical location (the origin), let this serving base station (from tier i) be located at x_i. If each base station in the kth tier transmits with total power P_k^{b} on the beacon channel, then the nearest base station from tier $j \neq i$, at location x_j, say, must be at distance $\|\mathsf{x}_j\|$ from the user such that

$$\theta_j P_j^{\mathrm{b}} \tilde{\ell}_j(\mathsf{x}_j) < \theta_i P_i^{\mathrm{b}} \tilde{\ell}_i(\mathsf{x}_i) \Leftrightarrow \tilde{\ell}_j(\mathsf{x}_j) < \beta_{ij} \tilde{\ell}_i(\mathsf{x}_i), \quad \beta_{ij} \equiv \frac{\theta_i}{\theta_j} \frac{P_i^{\mathrm{b}}}{P_j^{\mathrm{b}}}, \; j \neq i, \tag{7.6}$$

where θ_i and θ_j are the corresponding tier biases, and for each i, $\tilde{\ell}_i(\mathsf{x}) = (\kappa_i \|\mathsf{x}\|)^{-\alpha_i}$ is the path loss on the link to the user from a transmitter x in tier i. For future reference, we recall that the propagation process in each tier j of the equivalent base station deployment without shadow fading is

$$\tilde{\Theta}_j = \{1/\tilde{\ell}_j(\mathsf{x}) : \mathsf{x} \in \tilde{\Phi}_j\}.$$

Fading model

From the discussion in Section 7.2.1, it follows that the received power at the user terminal on a "traffic" channel contains the effect of only the nonshadow fading (e.g., Rayleigh fading) on the links to it from the serving base station and all the interfering base stations. However, we must describe in greater detail what a "link" between a base station and the user terminal is because each base station in tier i uses, say, $N_{\mathrm{t},i}$ transmit antennas, whereas the user terminal uses, say, N_{r} receive antennas, and there are therefore $N_{\mathrm{t},i} \times N_{\mathrm{r}}$ "links" between a base station in tier i and the user terminal. It is customary to define the MIMO "channel" between this base station and the user terminal as the $N_{\mathrm{t},i} \times N_{\mathrm{r}}$ matrix of baseband-equivalent complex "gains" of the (presumed frequency-nonselective) SISO channels between each of the $N_{\mathrm{t},i}$ transmit antennas and each of the N_{r} receive antennas.

If we have only a single transmit antenna at each base station and a single receive antenna at the user terminal, then the resulting SISO channel corresponds to the scenarios studied in Chapter 5. We continue to assume Rayleigh fading on each SISO channel in the MIMO channel, and make the additional assumption that the fading coefficients on the $N_{\mathrm{t},i} \times N_{\mathrm{r}}$ SISO channels between the transmit antennas at a base station in tier i and the receive antennas at the user terminal are iid. Note also that the common distribution of these fading coefficients does not depend on i; that is, it is the same for all tiers.

The (matrix) MIMO channels between all interfering base stations and the user terminal may also be assumed independent (and therefore iid). However, depending upon the specific MIMO transmission scheme in use in the system, the MIMO channel between the serving base station and the user terminal may not be independent of the MIMO channel between an interfering base station and the user terminal. For example, the

independence assumption is justified across the serving and interfering base stations for the case of *zero-forcing beamforming* with full *channel state information* at the user of the MIMO channel from the serving base station (Dhillon, Kountouris, & Andrews 2013, Appendix A), but does not hold when the user terminal performs MRC (Tanbourgi, Dhillon, & Jondral 2015).

Distribution of received power

Regardless of the specifics of the MIMO transmission scheme, the received power at each receive antenna of the user terminal from a base station (serving or interfering) is obtained by summing the powers of the received signals (sometimes after an orthogonal linear transformation) on the SISO channels making up the MIMO (strictly speaking, multiple-input single-output) channel between the transmit antennas at this base station and the single receive antenna of interest at the user terminal.

With the assumption of iid Rayleigh fading on each of these SISO channels forming a MIMO channel, it is not surprising that the received power at the user terminal from a base station (serving or interfering) is Erlang distributed, that is, gamma distributed with unit scale/rate parameter and integer shape parameter (with the shape parameter dependent on the specifics of the MIMO transmission scheme, and different between the serving base station and an interfering base station). What is remarkable, however, is that the received power from the serving base station may be approximated as independent of the received power from any interfering base station, even when the assumption of independence of the MIMO channels to the user terminal from the serving base station and the interfering base station is not justified (Huang, Papadias, & Venkatesan 2012).

From the above discussion, we henceforth assume that the received power at a single user terminal antenna (located at the origin) from a base station (located at $\mathbf{x}$) equipped with N_{t} transmit antennas and transmitting with total power P distributed equally over $M \leq N_{\mathrm{t}}$ transmit antennas, where the SISO channels between these transmit antennas and the receive antenna are iid Rayleigh, is the scaled Erlang random variable

$$\frac{P}{M(\kappa\|\mathbf{x}\|)^{\alpha}}G,$$

where

$$G \equiv \sum_{m=1}^{M} H_m \sim \Gamma(M,1),$$

and $H_1, \ldots, H_M$ iid Exp(1) are the fading coefficients on the M iid SISO Rayleigh-fading channels between the M active transmit antennas at the base station and the single receive antenna at the user terminal.

Note that G has the CCDF

$$\mathbb{P}\{G > \tau\} = \mathrm{e}^{-\tau} \sum_{m=0}^{M-1} \frac{\tau^m}{m!}$$

$$= \sum_{m=0}^{M-1} \frac{(-1)^m}{m!} \frac{\mathrm{d}^m}{\mathrm{d}\phi^m} \mathrm{e}^{-\phi\tau} \Big|_{\phi=1}$$

$$= \sum_{m=0}^{M-1} \frac{(-1)^m}{m!} \frac{\mathrm{d}^m}{\mathrm{d}\phi^m} \mathbb{P}\{H > \phi\tau\} \Big|_{\phi=1}, \quad \tau \geq 0, \tag{7.7}$$

where $H \sim \text{Exp}(1)$ has the same distribution as $H_1, \ldots, H_M$. In Section 7.2.2, we see how to exploit Eq. (7.7) to derive the SINR distribution by reusing results from Chapter 5 for SISO channels.

7.2.2 Distribution of SINR

Recall that Proposition 5.2.3 yields the probability of coverage by the closest base station in a single-tier network model. We wish to extend this result to the multitier network model where the serving base station is selected from among the closest base stations in the tiers (one per tier) according to the criterion in Eq. (7.6).

To do so, let us first rewrite the proof of Proposition 5.2.3 in terms of the path losses from the base stations in the (sole) tier to the user location, as follows: We begin with the propagation loss process Θ and endow it with independent exponential marks $\{E_\mathsf{L}\}_{\mathsf{L}\in\Theta}$ representing Rayleigh fading, so for each $\mathsf{L} \in \Theta$, the random variable E_L is exponentially distributed with unit mean, and let H be equal in distribution to E_L. This new process $\{(\mathsf{L}, E_\mathsf{L})\}_{\mathsf{L}\in\Theta}$ is an independently marked Poisson point process with intensity measure Q, which coincides with the intensity measure of Θ because of the unit mean assumption on each E_L. We express the coverage probability of Eq. (5.16) as

$$\tilde{\mathcal{P}}(\tau) = \int_0^\infty \mathbb{P}\Big\{ E \geq s\tau(N + I_{(s,\infty)}) \Big\} f_{\mathsf{L}_{(1)}}(s)\, \mathrm{d}s = \int_0^\infty \mathcal{L}_N(s\tau)\mathcal{L}_{I_{(s,\infty)}}(s\tau) f_{\mathsf{L}_{(1)}}(s)\, \mathrm{d}s, \tag{7.8}$$

where $f_{\mathsf{L}_{(1)}}(s)$ is the probability density of $\mathsf{L}_{(1)}$, which is given by

$$f_{\mathsf{L}_{(1)}}(\mathrm{d}s) = -\frac{d}{\mathrm{d}s} e^{-Q(s)} = a\delta s^{\delta-1}, \tag{7.9}$$

because Θ is a Poisson process, and $I_{(s,\infty)}$ is the random variable representing conditional interference given $\mathsf{L}_{(1)} = s$. The random variable $I_{(s,\infty)}$ is equal in distribution to the random sum

$$\sum_{\mathsf{L}\in\Theta} (E_\mathsf{L}/\mathsf{L}),$$

and has a Laplace transform that can be explicitly evaluated:

$$\mathcal{L}_{I_{(s,\infty)}}(\xi) = \exp\left(-\int_s^\infty [1 - \mathcal{L}_E(\xi/v)] Q(\mathrm{d}v) \right) = \exp\left(-a\delta \frac{\xi s^\delta}{s} \frac{{}_2F_1(1, 1-\delta; 2-\delta; -\xi/s)}{(1-\delta)} \right). \tag{7.10}$$

With this independence assumption between the received powers from the serving base station and the interfering base stations (and bearing in mind that this independence

assumption is only an approximation for certain MIMO transmission schemes), the distribution of SINR at the user terminal may be computed on the same lines as in Proposition 5.2.3, with the difference that in the term $\mathbb{P}\{H \geq s\tau(N+I_{(s,\infty)})\}$ in Eq. (7.8), H is no longer exponential with unit mean but Erlang with unit scale/rate parameter.

To be specific, the generalization of Proposition 5.2.3 to the case of coverage probability under fading (with a single receive antenna at the user terminal and multiple transmitting antennas at both the serving and interfering base stations) in a J-tier HetNet is as follows:

$$\tilde{\mathcal{P}}(\tau) = \sum_{i=1}^{J} \tilde{\mathcal{P}}_i(\tau), \tag{7.11}$$

where for each $i = 1, \ldots, J$, the probability of coverage with the serving base station being (the nearest base station) from tier i is given by

$$\begin{aligned}
&\tilde{\mathcal{P}}_i(\tau) \\
&= \mathbb{P}\left\{ \frac{P_i G_i^*}{M_i^* \tilde{\mathsf{L}}_{i,(1)}} > \tau \left(N + \frac{P_i}{M_i} \sum_{\mathsf{x} \in \tilde{\Phi}_i : \tilde{\ell}_i(\mathsf{x}) < 1/\tilde{\mathsf{L}}_{i,(1)}} G_{\mathsf{x}} \tilde{\ell}_i(\mathsf{x}) + \sum_{\substack{j=1 \\ j \neq i}}^{J} \frac{P_j}{M_j} \sum_{\mathsf{x} \in \tilde{\Phi}_j} G_{\mathsf{x}} \tilde{\ell}_j(\mathsf{x}) \right), \right. \qquad (7.12) \\
&\qquad \left. (\forall j \neq i)\, \tilde{\mathsf{L}}_{j,(1)} > \beta_{ij} \tilde{\mathsf{L}}_{i,(1)} \right\} \\
&= \int_0^\infty \mathbb{P}\left\{ \frac{P_i G_i^*}{M_i^*} > \tau s \left(N + \frac{P_i}{M_i} \sum_{\mathsf{x} \in \tilde{\Phi}_i : \tilde{\ell}_i(\mathsf{x}) < 1/\tilde{\mathsf{L}}_{i,(1)}} G_{\mathsf{x}} \tilde{\ell}_i(\mathsf{x}) + \sum_{\substack{j=1 \\ j \neq i}}^{J} \frac{P_j}{M_j} \sum_{\mathsf{x} \in \tilde{\Phi}_j : \tilde{\ell}_j(\mathsf{x}) < \beta_{ij}/\tilde{\mathsf{L}}_{i,(1)}} G_{\mathsf{x}} \tilde{\ell}_j(\mathsf{x}) \right) \right\} \\
&\qquad \times \exp\left[- \sum_{\substack{j=1 \\ j \neq i}}^{J} Q_j([0, \beta_j s)) \right] f_{\tilde{\mathsf{L}}_{i,(1)}}(s)\, \mathrm{d}s, \qquad (7.13)
\end{aligned}$$

where for each tier $j = 1, \ldots, J$,

$$Q_j([0, \nu)) = a_j \nu^{\delta_j}, \quad \nu \geq 0,\ a_j = \frac{\pi \lambda_j \, \mathbb{E}[H_{j,\mathsf{x}}^{\delta_j}]}{\kappa_j^2}, \quad \delta_j = \frac{2}{\alpha_j},$$

the minimum link loss

$$\tilde{\mathsf{L}}_{j,(1)} = \min\{1/\tilde{\ell}_j(\mathsf{x}) : \mathsf{x} \in \tilde{\Phi}_j\}$$

has the probability density function in Eq. (7.9) with a and δ replaced by a_j and δ_j respectively, and we assume that for each tier i, every base station in tier i, whether serving or interfering, transmits with the same total power P_i, but the serving base station distributes this transmit power equally over M_i^* transmit antennas (so $G_i^* \sim \Gamma(M_i^*, 1)$), whereas every interfering base station distributes this transmit power equally over M_i transmit antennas (so $G_{\mathsf{x}} \sim \Gamma(M_j, 1)$ for every $\mathsf{x} \in \tilde{\Phi}_j$ unless x is the serving base station from $\tilde{\Phi}_i$).

For each tier $j = 1, \ldots, J$, the total interference power from this tier received at the (single) receive antenna at the user terminal, conditioned on the event that this user is served by (the nearest base station x_i from) tier i with $\tilde{\ell}_i(\mathsf{x}_i) = s$, is

$$I_{j,(\beta_{ij}s,\infty)} = \sum_{\mathsf{x}\in\tilde{\Phi}_j:\tilde{\ell}_j(\mathsf{x})<\beta_{ij}s} G_{\mathsf{x}}\tilde{\ell}_j(\mathsf{x}), \quad G_{\mathsf{x}} \sim \Gamma(M_j, 1) \text{ iid}, \quad j = 1, \ldots, J. \tag{7.14}$$

Then Eq. (7.13) can be written as

$$\begin{aligned}&\tilde{\mathcal{P}}_i(\tau)\\ &= \int_0^\infty \mathbb{P}\left\{G_i^* > s\tau\left(\frac{M_i^*}{P_i}N + \sum_{j=1}^J \frac{M_i^*}{M_j}\frac{P_j}{P_i}I_{j,(\beta_{ij}s,\infty)}\right)\right\} \exp\left\{-\sum_{j=1}^J a_j\beta_{ij}^{\delta_j}s^{\delta_j}\right\} \frac{2a_i}{\alpha_i}s^{-1+\delta_i}\,\mathrm{d}s.\end{aligned} \tag{7.15}$$

From Eq. (7.7), the first term of the integrand in Eq. (7.15) can be written as

$$\begin{aligned}&\mathbb{P}\left\{G_i^* > s\tau\left(\frac{M_i^*}{P_i}N + \sum_{j=1}^J \frac{M_i^*}{M_j}\frac{P_j}{P_i}I_{j,(\beta_{ij}s,\infty)}\right)\right\}\\ &\quad= \sum_{m=0}^{M_i^*-1}\frac{(-1)^m}{m!}\frac{\mathrm{d}^m}{\mathrm{d}\phi^m}\mathbb{P}\left\{H^* > \phi s\tau\left(\frac{M_i^*}{P_i}N + \sum_{j=1}^J \frac{M_i^*}{M_j}\frac{P_j}{P_i}I_{j,(\beta_{ij}s,\infty)}\right)\right\}\Bigg|_{\phi=1}\\ &\quad= \sum_{m=0}^{M_i^*-1}\frac{(-1)^m}{m!}\frac{\mathrm{d}^m}{\mathrm{d}\phi^m}\left[\mathcal{L}_N\left(\frac{M_i^*}{P_i}\phi s\tau\right)\prod_{j=1}^J \mathcal{L}_{I_{j,(\beta_{ij}s,\infty)}}\left(\frac{M_i^*}{M_j}\frac{P_j}{P_i}\phi s\tau\right)\right]\Bigg|_{\phi=1},\end{aligned} \tag{7.16}$$

where $H^* \sim$ Exp(1).

Now, for any $j = 1, \ldots, J$ and any $s_j > 0$, the Laplace transform of $I_{j,(s_j,\infty)}$ can be evaluated from Eq. (5.17) and Eq. (7.14):

$$\begin{aligned}\mathcal{L}_{I_{j,(s_j,\infty)}}(\xi) &= \exp\left\{-\int_{s_j}^\infty \left[1 - \mathcal{L}_{G_j}\left(\frac{\xi}{\nu}\right)\right] Q_j(\mathrm{d}\nu)\right\}\\ &= \exp\left\{-\int_{s_j}^\infty \left[1 - \mathcal{L}_H\left(\frac{\xi}{\nu}\right)^{M_j}\right] Q_j(\mathrm{d}\nu)\right\},\end{aligned} \tag{7.17}$$

where $G_j \sim \Gamma(M_j, 1)$, $H \sim$ Exp(1).

The right side of Eq. (7.17) can be evaluated as follows (see also Tanbourgi et al. 2015, Lemma 3):

$$\begin{aligned}&\mathcal{L}_{I_{j,(s_j,\infty)}}(\xi)\\ &\quad= \exp\left\{-a_j s_j^{\delta_j}\left[\frac{M_j\xi}{s_j}\,\frac{{}_2F_1(M_j+1, 1-\delta_j; 2-\delta_j; -\xi/s_j)}{(1-\delta_j)} + \frac{1}{(1+\xi/s_j)^{M_j}} - 1\right]\right\}.\end{aligned} \tag{7.18}$$

Analytic expressions for the derivatives of this Laplace transform that are required in Eq. (7.16) can be obtained by applying Faà di Bruno's formula for the higher-order derivative of a composition of two functions, where the outer function is the exponential and the inner function is the argument of the exponential in the right side of Eq. (7.18):

$$\frac{\mathrm{d}^n}{\mathrm{d}x^n} f(g(x)) = \sum_{(m_1,\ldots,\, m_n)\in\mathbb{Z}_+^n : \sum_{j=1}^n j m_j = n} \frac{n!}{\prod_{j=1}^n m_j!} f^{(m_1+\cdots+m_n)}(g(x)) \prod_{j=1}^n \left(\frac{g^{(j)}(x)}{j!}\right)^{m_j}.$$

For the special case of a single-tier network where the user terminal has a single receive antenna ($N_\mathrm{r} = 1$), the transmitting base stations use *maximal ratio transmit* beamforming, and thermal noise may be neglected at the user terminal receiver, an elegant expression for the probability of outage may be derived (Li, Zhang, & Letaief 2014, Theorem 1).

All of this analysis was for a single receive antenna at the user terminal. If the user terminal uses MRC on the outputs of its receive antennas, then the calculation of coverage probability becomes the calculation of the CCDF of the sum of the SINRs at the outputs of the individual receive antennas. However, there are more effective ways to exploit the presence of multiple receive antennas, as we see in Section 7.2.3.

7.2.3 CoMP analysis

In CoMP transmission scenarios, multiple base stations transmit jointly to the user terminal of interest. These base stations may not all transmit the same symbols to the user. Further, even if they all transmit the same symbols, they may use different precoding schemes for their transmissions.

Serving base station set selection

In CoMP, a set of serving base stations is chosen as opposed to a single serving base station. The serving base station set selection criterion is a generalization of that for a single serving base station in MIMO: the set of serving base stations is the set (of some chosen size) yielding the greatest sum of received average powers (for the "beacon" channel, without any special processing) at the user terminal. The received average powers may have tier-dependent bias terms applied when computing the above sum. Note that in a multitier HetNet, the serving set could include base stations from multiple tiers, and more than a single base station from a given tier. The two-tier HetNet scenario analyzed in Section 6.1.5 is an example of a CoMP scheme where the jointly transmitting base stations transmit the same symbols to the user, albeit with only a single transmit antenna each (and the user terminal also has only one receive antenna).

Distribution of SINR

For a multitier HetNet where all base stations have only one transmit antenna, the user terminal has only one receive antenna, and the serving set base stations all transmit the same symbols to the user without precoding, we have the generalization of the scenario analyzed in Proposition 5.2.3 to multiple transmitting base stations. For the case where

the path loss exponents are the same for all tiers, the coverage probability can be derived in closed form (Nigam, Minero, & Haenggi 2014, Theorem 1).

A more sophisticated CoMP transmission scheme called *interference alignment* (IA) also selects a set (called a *cluster*) of, say, K base stations as those with the K-highest received average powers at the user terminal of interest.

Consider a single-tier deployment Φ. In the notation of Section 7.2.1, the base stations in the cluster for a given user terminal are the K nearest base stations to the user terminal in the equivalent deployment Θ. However, unlike the earlier CoMP transmission scenario where all the base stations in the serving set transmit jointly to the user of interest, in IA, the user terminal is served by a single base station, namely the one that is nearest to it. Each base station has N_{t} transmit antennas, while the user terminal has N_{r} receive antennas.

The $K-1$ other base stations in the cluster simultaneously serve $K-1$ other users, but use precoding vectors such that, assuming perfect knowledge of all MIMO channels (between all K base stations in the cluster and the user terminal of interest) at both the transmitters (the K base stations) and the receiver (the user terminal of interest), the received signal at the user terminal of interest can be passed through a filter that eliminates the interference from the other $K-1$ in-cluster base stations. This is possible if the number of streams transmitted by the serving base station is $d \leq \min\{N_{\mathrm{t}}, N_{\mathrm{r}}\}/2$ and $N_{\mathrm{t}} + N_{\mathrm{r}} \geq (K+1)d$ (Mungara, Morales-Jimenez, & Lozano 2015).

The cluster-based IA transmission scheme turns the MIMO channel between the serving base station and the user terminal into a set of d parallel and independent SISO channels. Each such SISO channel has the user terminal receiving from the base station that is nearest to it, and with no interference contribution from the nest $K-1$ nearest base stations. In other words, on each of these equivalent SISO channels, the scenario at the output of the user terminal's demodulator (after filtering) is the same as that analyzed in Proposition 5.2.3, with the difference that the total interference power comes from all base stations farther away than the Kth-nearest one. Conditioned on the distance to the serving (nearest) base station and the Kth-nearest base station, we can use Eq. (5.16) to get the coverage probability, where now the quantity r in Eq. (5.17) is replaced by the distance to the Kth-nearest base station. For the unconditional coverage probability, we have to average over the joint distribution of the distances to the nearest and Kth-nearest base stations, which is also known for a homogeneous Poisson point process.

7.3 Multi-Link SINR Analysis: Area Spectral Efficiency and Energy Efficiency

All the analysis up to this point has been for the SINR at the typical location. In other words, until now we have viewed the performance of the cellular network from the perspective of a user at the typical location of such a network. In the present chapter, we shall view the network from the perspective of the operator of the network. Whereas the individual users are only interested in the distribution of SINR at their respective terminals, the network operator has to take a wider view and try to support as many of its users as possible with as high a data rate as possible. As we shall see, this leads to different performance metrics for the network, and their analysis forms the core of this section.

7.3.1 Link-centric vs. cell-centric perspective

The analysis of SINR distribution at the terminal of a user at the typical location (which we may call the typical user experience) is for a single link, namely the (down) link from the serving base station to this user terminal.[1] On the other hand, the operator of the network needs to provide the largest number of subscribers with the "best" possible user experience. We shall shortly define exactly what we mean by this but for now it is sufficient to realize that the network performance metric of interest to the network operator is a network-level rather than a link-level metric.

For the purposes of the discussion here, we assume that from the perspective of the network operator, network-level metrics may be replaced by cell-level metrics. This implicitly assumes that (a) user distributions and traffic loads are identically distributed across cells and (b) anything that happens in one cell does not influence what happens in any other cell. Of course, both of these assumptions are only approximately correct, but if user distribution over the network may be assumed approximately homogeneous, and we ignore the effects of user handover between cells, or transmit power coordination across cells to mitigate interference from neighboring cells, then we may focus only on a typical cell of the network for the purposes of calculating performance metrics that are of interest to the network operator.

To summarize: for any user of the network, the performance metric of interest is the distribution of SINR at the typical location. For the operator of the network, on the other hand, the relevant perspective is at the level of the typical cell (the performance metric of interest is defined in Section 7.3.2). This distinction may be made rigorous in the language of stochastic geometry as follows: the typical location is usually taken to be the origin, which makes the cell containing the typical location the zero-cell (Section 4.1.6). Thus, the network performance metric of interest to the typical location is computed on the zero-cell, whereas the network performance metric of interest to the network operator is computed on the typical cell. Recall that the zero-cell and typical cell do not have the same distribution – for example, the area of the zero-cell is greater than that of the typical cell (see Section 4.1.6).

7.3.2 Performance metrics of interest to operators

Now that we have established the importance of the cell-centric viewpoint to the network operator for the analysis of network performance, it is time to examine exactly what network performance metrics the network operator should focus on as a reliable indicator of the quality of its network.

Throughput vs. instantaneous rate

Until now all the analysis for the typical location has focused on the distribution of SINR at a single snapshot of the deployment, that is, the instantaneous SINR. However,

[1] Strictly speaking, this scenario should be called the *single receiver terminal* because in the case of CoMP, there are multiple simultaneously transmitting base stations. However, because the analysis is still at the terminal of one user, we shall abuse notation and include the CoMP analysis of Section 7.2.3 under that for a single link.

the key metric for the network operator is neither the maximum instantaneous SINR (or data rate) to any specific user in the typical cell, nor the distribution of instantaneous data rates across the users in the typical cell, but the distribution of throughput over the links to the users in the typical cell, where by the term *throughput* (on the link to a given user terminal) we mean the rate of successful data transmission to that user terminal (unit: bits/s). Further, the rate to the user is not the instantaneous rate, but the time-averaged rate (averaged over a large number of transmissions or, equivalently, averaged over a large number of sessions to the user from the cell base station).

Clearly, the distribution of throughput across the users in the typical cell depends on the specific scheduling policy employed by the cell base station. This makes the analysis of the distribution of throughput across the users intractable because even the distribution of distance between the cell base station and a randomly selected user in the typical cell is unknown. We therefore focus on the *system throughput*, namely the sum of the throughputs to all users in the typical cell.

Spectral efficiency

The throughput to a given user is dependent on the modulation and coding scheme used for transmission, the transmit power, and the bandwidth of the link. It is convenient to normalize the throughput on a link by the bandwidth, and then use the resulting quantity instead, which is called the *spectral efficiency* on that link (unit: bits/s/Hz). Further, to remove the dependence on the scheduling policy, we define the spectral efficiency on the link to a given user (location) as the bandwidth-normalized throughput on that link, assuming that this user is the *only* user in the typical cell, that is, that this user is assigned the full resources (maximum transmit power, the highest modulation and coding scheme, etc.) available in that cell (Karray 2010). Note that "bits/s/Hz" is dimensionally identical to "bits," and the spectral efficiency on the link to a given user can be interpreted as the maximum number of bits that can be pumped over that link with a single use of the link, that is, a single transmission over that link.

Cell spectral efficiency

The *cell spectral efficiency* is the expected value of the spectral efficiency on the link to a user whose location is uniformly distributed over the typical cell. With this definition, we see that the cell spectral efficiency is also the bandwidth-normalized system throughput under round-robin scheduling. We use the cell spectral efficiency from now on as a tractable system throughput metric. Unfortunately, even the cell spectral efficiency cannot be computed exactly because, as stated in Section 7.3.2, we do not have an expression for the distribution of the distance from the base station to a user whose location is uniformly distributed over the typical cell, although some approximations have been suggested (Haenggi 2017, Yu et al. 2015).

Area spectral efficiency

The *area spectral efficiency* is just the cell spectral efficiency normalized by the area of the typical cell. With our interpretation of the cell spectral efficiency as the system throughput, the area spectral efficiency can be interpreted as the system throughput per unit area (unit: bits/s/Hz/m^2).

Energy efficiency

The *energy efficiency* is defined as the system throughput (in the typical cell) normalized by the power consumed by the cell base station (note that the power consumed is greater than the transmit power). For tractability, the system throughput is replaced by the cell spectral efficiency multiplied by the bandwidth.

From these definitions, it is easy to see that the spectral efficiency is the main measure of network-level performance, and the basis for several other metrics such as cell spectral efficiency, area spectral efficiency, and energy efficiency. Next, we discuss how to calculate the spectral efficiency on the link to a given user in the typical cell.

7.3.3 Spectral efficiency as ergodic capacity

Consider a user located at $\mathbf{x}$ in the typical cell (with the cell base station assumed to be at the origin). From the discussion in Section 7.3.2 of the spectral efficiency on the link to this user being the bandwidth-normalized throughput on the link, and the definition of throughput as the (long-term average) rate of successful data transmission on the link, the spectral efficiency may be expressed as the bandwidth-normalized *ergodic capacity* of this link:

$$\mathrm{SE}_{\mathbf{x}} = \mathbb{E}[\log_2(1 + \mathrm{SINR}_{\mathbf{x}})] \text{ bits/s/Hz.} \tag{7.19}$$

This expression for ergodic capacity on a link was derived for an AWGN channel (that is, with SNR instead of SINR), assuming one transmit and one receive antenna, iid fading across all transmissions on the link, and knowledge of the fading coefficient at the receiver.

The extension of the original formula for ergodic capacity on an AWGN channel to Eq. (7.19) for the interference-limited scenario requires the assumption that the (complex equivalent baseband) total interference (signal, not power) at the user terminal may be treated as (complex circularly symmetric) AWGN. This last assumption is clearly false, as can be seen from the expression of the characteristic function of the total interference signal from all interferers farther than $\|\mathbf{x}\|$ located at the points of a homogeneous Poisson process (Renzo & Guan 2014, Theorem 1).

However, if the transmitter and receiver use a codebook and decoder designed for AWGN, and the receiver simply ignores the distribution of the interference, then it is a remarkable fact that the achievable spectral efficiency is the same as if the interference were indeed AWGN (Mungara et al. 2015). In other words, Eq. (7.19) can be used to compute the spectral efficiency on a link so long as the transmitter and receiver use a codebook and decoder designed for AWGN.

Eq. (7.19) is an upper bound on the practically achievable spectral efficiency since it cannot be assumed that the transmitter knows the fading states of all channels between the other transmitters and the receiver. If, instead, the fading in the interferer's channels is ignored in the analysis, a tight lower bound on the achievable spectral efficiency is obtained (George et al., 2017).

8 Extensions to Non-Poisson Models

8.1 Non-Poisson Point Processes

8.1.1 Motivation

In Chapters 5–7 we see how Poisson network models often give tractable results, and how random propagation effects such as fading and shadowing can render a network to appear more Poisson. Nevertheless, the Poisson model may still not be appropriate for certain network layouts. For example, although the Poisson process statistically does not exhibit either repulsion or clustering, it is possible for the clusters or voids of points to exist in network layouts.

8.1.2 Appropriate point processes

In Chapter 3 we see just a small selection of the possible spatial point processes. To choose appropriate point processes for network models, researchers have statistically analyzed networks located in various cities to demonstrate that other point processes may be more appropriate, although of course such analyses can be only as good as the data. But when developing stochastic geometry models of cellular networks, the number of suitable and tractable point processes quickly becomes limited. In terms of the SINR for a single user, there are two important and necessary features of a potential point process. One is the knowledge of the Palm distribution, as covered in Section 3.2.4, which immediately limits the choice of possible point processes. The other is being able to write down a calculable expression (at least, numerically) of the Laplace functional of the point process.

In short, although there is a rich range of possible point processes, the bulk of them unfortunately lack the necessary properties that make them tractable in our setting, which partly explains why the clear majority of network models are based on the Poisson point process. Consequently, we restrict ourselves to results for point processes from two general families that exhibit repulsion and clustering, respectively, determinantal processes and a specific shot noise Cox process.

8.1.3 Choice of the base station and propagation effects

In Section 5.1.3, we see the two different model assumptions for how a user at the typical location (the origin) chooses the serving base station under the Poisson model. In this

chapter, we assume that the serving base station is chosen so it minimizes its distance to the typical location. Again, we write $\mathsf{x}^* \in \Phi$ to denote the closest station to the origin, that is, $|\mathsf{x}^*| \leq \min_{\mathsf{x}' \in \Phi} |\mathsf{x}'|$.

For the non-Poisson models we cover here, we assume that the random propagation effects are Rayleigh fading, so they are represented by iid exponential variables; as we have seen, there is a natural connection between this model and the Laplace functional.

8.1.4 Determinantal models

Determinantal point processes have been proposed to model the locations of base stations in cellular networks with a level of repulsion or regularity between the base stations. The repulsion of determinantal processes has been studied, resulting in methods to quantify it (Biscio & Lavancier 2016). The Laplace functional of a general determinantal process can be written down as an infinite series of integrals involving the kernel (or rather, the factorial moment densities) of the process, which simplifies to a closed-form expression in the Poisson case. Furthermore, the reduced Palm distribution of a determinantal process is given by another determinantal process with a modified kernel, which gives a tool similar to Slivnyak's theorem for Poisson processes. These two key properties, alongside further convenient properties for statistically fitting and simulating such processes, have motivated the use of determinantal to model network layouts that exhibit a degree of repulsion between base stations.

8.1.5 Ginibre point process

The most tractable determinantal process is arguably the Ginibre point process, often abbreviated as GPP, and was the first such process to be used as a model for base station locations with repulsion.

The Ginibre process is usually defined on $\mathbb{C}$ instead of $\mathbb{R}^2$. The product densities of the standard Ginibre process are given by

$$\varrho^{(n)}(x_1, \ldots, x_2) = \det[K](x_1, \ldots, x_n),$$

where the kernel function K is the Gaussian kernel

$$K(x, y) = \pi^{-1} \mathrm{e}^{-(|x|^2 + |y|^2)/2} \mathrm{e}^{x\bar{y}}, \quad x, y \in \mathbb{C},$$

with respect to the Lebesgue measure on $\mathbb{C}$. It follows that $\mathbb{E}\Phi(B) = \int_B \varrho^{(1)}(x)\mathrm{d}x = \pi^{-1}|B|$, that is, the density of the Ginibre process is π^{-1}. Moreover, the Ginibre process is motion invariant, and thus its second moment density $\varrho^{(2)}(x, y)$ only depends on $|x-y|$. It is given by

$$\varrho^{(2)}(x, y) = \pi^{-2}(1 - \mathrm{e}^{-|x-y|^2}),$$

which is consistent with the fact that the Ginibre process is repulsive (at all distances).

The standard Ginibre process can be generalized to the β-Ginibre process using independent thinning and rescaling. For $0 < \beta \leq 1$, the β-Ginibre process is obtained by retaining each point of the standard Ginibre process independently with probability

β and subsequent rescaling by the factor $\sqrt{\beta}$ to maintain the original density of the Ginibre process. As $\beta \to 0$, the β-Ginibre process converges weakly to the Poisson point process of intensity π^{-1}, so the parameter β can be used to interpolate between the standard Ginibre process (or 1-Ginibre process) and the Poisson process.

Next, to achieve arbitrary densities $\lambda = c\pi^{-1}$, the β-Ginibre process can be scaled, which results in the kernel

$$K_{\beta,c}(x,y) = c\pi^{-1}\mathrm{e}^{-c(|x|^2+|y|^2)/(2\beta)}\mathrm{e}^{(c/\beta)x\bar{y}}, \quad x,y \in \mathbb{C},$$

again with respect to the Lebesgue measure. Ripley's K function for this point process is

$$K_{\mathrm{R}}(r) = \pi r^2 - \frac{\beta\pi}{c}\left(1 - \mathrm{e}^{-cr^2/\beta}\right).$$

It is apparent that as β decreases to 0, the gap to the K function of the Poisson process vanishes.

A key result that makes the β-Ginibre process relatively tractable is due to Kostlan (Hough et al. 2009):

PROPOSITION 8.1.1 *Let $\Phi_c = \{\mathsf{x}_i\}$ be a scaled β-Ginibre process. For $k \in \mathbb{N}$, let Q_k be a random variable with probability density*

$$f_{Q_k}(x) = \frac{x^{k-1}\mathrm{e}^{-cx/\beta}}{\left(\frac{\beta}{c}\right)^k \Gamma(k)}.$$

That is, Q_k is gamma distributed with parameters k and β/c, such that $\mathbb{E}Q_k = k\beta/c$. Assume Q_k to be independent from Q_j for all $k \neq j$. Then the point process $\{\|x_i\|^2\}$ has the same distribution as the point process $\{\mathsf{y}_i\}$ obtained by retaining each point from $\{Q_k\}$ with probability β, independent from everything else.

Hence, in the case of the 1-Ginibre process, the set of square distances $\{\|x_i\|^2\}$ is distributed like a set of independent gamma random variables. For $\beta < 1$, some of the gamma random variables need to be removed. The Palm measure (at o) satisfies a similar property:

PROPOSITION 8.1.2 *The Palm measure of a scaled β-Ginibre process Φ_c is the law of the point process obtained by adding the origin to Φ_c and removing the point x with $\|\mathsf{x}\|^2 = Q_1$ (which only exists with probability β).*

If the point x with $\|\mathsf{x}\|^2 = Q_1$ has been eliminated from Φ_c, then no point is removed. Again this is consistent with the fact that as $\beta \to 0$, the β-Ginibre process converges to a Poisson process, because in this case, the point whose modulus corresponds to Q_1 is not present in the Ginibre process *with high probability*, hence the only difference of the Palm measure is the point at o.

Coverage probability expression

We consider the typical location in a cellular network where base stations form a β-Ginibre process of intensity c/π and transmit at unit power.

PROPOSITION 8.1.3 *We assume a path loss function $\ell(x) = \|x\|^{-\alpha}$ and Rayleigh fading with unit mean. Then for β-Ginibre process, the coverage probability is*

$$\mathbb{P}\{\,\mathrm{SINR}(\mathsf{x}^*, o) > \tau\,\} = \beta \int_0^\infty \mathrm{e}^{-s}\mathrm{e}^{-\theta\sigma^2(\beta s/c)^{\alpha/2}} M(\tau, s, \alpha, \beta) S(\tau, s, \alpha, \beta)\mathrm{d}s, \tag{8.1}$$

where

$$M(\tau, s, \alpha, \beta) = \prod_{k=1}^{\infty} \left(\int_s^\infty \frac{v^{k-1}\mathrm{e}^{-v}}{\Gamma(k)} \frac{\beta}{1 + \theta(s/v)^{\alpha/2}} \mathrm{d}v + 1 - \beta \right) \tag{8.2}$$

and

$$S(\tau, s, \alpha, \beta) = \sum_{i=1}^{\infty} s^{i-1} \left(\int_s^\infty v^{i-1}\mathrm{e}^{-v} \frac{\beta}{1 + \theta(s/v)^{\alpha/2}} \mathrm{d}v + (1-\beta)\Gamma(i) \right)^{-1}. \tag{8.3}$$

The proof, based on Proposition 8.1.1, was presented for $\beta = 1$ in Miyoshi & Shirai (2014a) and for general β in Miyoshi & Shirai (2014b), and, using a different approach, in Deng, Zhou, & Haenggi (2015a). This expression is somewhat tedious to evaluate due to the three nested integrals and infinite sums/products, but it is one of the few cases besides the Poisson point process where an exact expression of the coverage probability can be derived.

8.1.6 General determinantal point process

We present some results for general determinantal point process, which are mostly based on the results of Li et al. (2015) and Shirai & Takahashi (2003) who later statistically fitted some determinantal processes to real network deployments (Li et al. 2014). We present only a selection of these results and note others are available in Li et al. (2014).

Existence and examples

To guarantee the determinantal process is stationary, it is natural to assume that the kernel has the form

$$K(x, y) = K_0(x - y), \quad x, y \in \mathbb{R}^2,$$

where K_0 is called the *covariance function*. More generally, depending on the particular determinantal process, conditions on its kernel guarantee its existence as a mathematical object. These conditions can be given through the covariance function of the process.

Instead of the covariance function, sometimes one needs to work with the *spectral density* of a determinantal process, which for a stationary determinantal process is defined as the Fourier transform of its kernel, namely

$$\varphi(x) = \int_{\mathbb{R}^2} K_0(t)\mathrm{e}^{-2\pi i x t}\mathrm{d}t, \qquad x \in \mathbb{R}^2. \tag{8.4}$$

Example 8.1 A stationary point process is a *Gauss determinantal point process* if it has a covariance function given by

$$K_0(x) = \lambda e^{-\|x\|^2/\alpha^2}, \qquad x \in \mathbb{R}^2, \tag{8.5}$$

where λ, as in the Poisson case, is the density of points, while the parameter α controls the repulsiveness of the points. The Gauss determinantal process exists if the condition $\lambda \leq 1/(\pi\alpha^2)$ is satisfied.

Example 8.2 A stationary point process is a *Cauchy determinantal point process* if it has a covariance function given by

$$K_0(x) = \frac{\lambda}{(1 + \|x\|^2/\alpha^2)^{\nu+1}}, \qquad x \in \mathbb{R}^2, \tag{8.6}$$

where λ is the density of points, while α is the scale parameter and ν is the shape parameter. Both parameters α and ν influence the repulsiveness of the points. The Gauss determinantal process exists if the condition $\lambda \leq \nu/(\pi\alpha^2)$ is satisfied.

Example 8.3 A stationary point process is a *generalized Gamma determinantal point process* if has a spectral density given by

$$\varphi(x) = \frac{\lambda\nu\alpha^2}{2\pi\Gamma(2/\nu)} e^{-|\alpha x\|^\nu}, \qquad x \in \mathbb{R}^2, \tag{8.7}$$

where λ is the density of points, while the parameters α and ν again influence the repulsiveness of the points. The generalized Gamma determinantal process exists if the condition $\lambda \leq 2\pi\Gamma(2/\nu)/(\nu\alpha^2)$ is satisfied.

More examples of determinantal point processes can be found in Biscio & Lavancier (2016) and Lavancier et al. (2015).

Laplace functional

The Laplace functional of a determinantal process can be expressed as an infinite expansion by using its kernel (Li et al. 2015, Lemma 2).

LEMMA 8.1.4 *For a determinantal process Φ on $\mathbb{R}^2$ with a kernel K that ensures its existence, its Laplace functional is*

$$\mathcal{L}_\Phi(f) = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int_{(\mathbb{R}^2)^n} \det[K(x_i, x_j)]_{1\leq i,j\leq n} \prod_{i=1}^{n} [1 - e^{-f(x_i)}] dx_1 \dots dx_n, \tag{8.8}$$

where f is any nonnegative (measurable) function on $\mathbb{R}^2$ with compact (that is, bounded and closed) support.

In certain cases, we can lift the requirement that the support of f needs to be bounded, but we skip such details here (see Li et al. 2015, Lemma 2).

Reduced Palm distribution

The reduced Palm distributions of determintantal processes are known (Shirai & Takahashi 2003, Theorem 1.7).

LEMMA 8.1.5 *For a determinantal process Φ on $\mathbb{R}^2$ with a kernel K that ensures its existence, under the reduced Palm measure at $x_o \in R^2$, the process Φ coincides with another determinantal process with kernel $K^!_{x_o}$ such that for $K(x_o, x_o) > 0$ (almost) all x_o, the kernel is*

$$K^!_{x_o}(x, y) = \frac{1}{K(x_o, x_o)} \det \begin{bmatrix} K(x, y) & K(x, x_o) \\ K(x_o, y) & K(x_o, x_o) \end{bmatrix}.$$

Nearest-neighbor distance distribution

For general determinantal point processes, one can write down an expression for the nearest-neighbor distance distribution (Li et al. 2015, Lemma 6).

LEMMA 8.1.6 *For any determinantal point process Φ on $\mathbb{R}^2$ with a kernel K, its nearest-neighbor distance distribution is given by*

$$G(r) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} \int\limits_{(B(o,r))^n} \det[K^!_o(x_i, x_j)]_{1 \le i,j \le n} \prod_{i=1}^{n} \mathrm{d}x_1 \ldots \mathrm{d}x_n, \tag{8.9}$$

where

$$K^!_o(x, y) = \frac{1}{K(o, o)} \det \begin{bmatrix} K(x, y) & K(x, o) \\ K(o, y) & K(o, o) \end{bmatrix}.$$

Coverage probability expression

In the zero-noise case, we present the coverage probability of (a user located at) the typical location that connects to its nearest base station located at x_0. We assume a non-increasing path loss function ℓ, for which we use a slight abuse of notation, writing both $\ell(x)$ and $\ell(r)$, where $x \in \mathbb{R}^2$, $r = \|x_o\|$ and $x_o \in \mathbb{R}^2$.

PROPOSITION 8.1.7 *We assume a path loss function ℓ and Rayleigh fading with unit mean. Then for a stationary determinantal point process Φ on $\mathbb{R}^2$ with a kernel K, the coverage probability is*

$$\mathbb{P}\{\,\mathrm{SIR}(\mathsf{x}^*, o) > \tau\,\} = 2\pi\lambda \times \int_0^{\infty} \left(\sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int\limits_{(\mathbb{R}^2)^n} \det[K^!_{x_o}(x_i, x_j)]_{1 \le i,j \le n} \prod_{i=1}^{n} \left[1 - \frac{\mathbf{1}_{\|x_i\| \ge r}}{1 + \tau \ell(x_i)/\ell(r)} \right] \mathrm{d}x_1 \ldots \mathrm{d}x_n \right) r \mathrm{d}r, \tag{8.10}$$

This coverage probability result was originally derived by Li et al. (2015). Despite its look, the above expression has been numerically evaluated using quasi-Monte Carlo techniques (Li et al. 2015). The coverage probability results for the β-Ginibre model in Proposition 8.1.3 can be obtained from the general results in Proposition 8.1.7.

8.1.7 Cox point processes

As covered in Chapter 7, Cox processes form a considerable generalization of the Poisson process by replacing its nonrandom measure with a random one that is independent of the underlying Poisson process. Cox processes exhibit clustering among the points, which intuitively makes sense, because if a point of the process exists at one location, then this is probably due to the random intensity measure being large at this location, so it is likely that the measure is large in the nearby vicinity.

Being based on the Poisson process, the Laplace functional and the Palm distribution of Cox processes can be tractable in certain situations, depending on the random measure. Consequently, Cox processes serve as a potential model for cellular networks that exhibits a clustering of base stations. But there is relatively little work on the using Cox processes to model base station locations. One notable exception is the Neyman-Scott cluster process (Section 8.1.8). A particular subclass of Cox processes are formed from permanental point processes, but, as mentioned in Chapter 7, these processes appear not to be as tractable as determinantal processes.

8.1.8 Neyman-Scott cluster processes

We give results for the Neyman-Scott (cluster) point process, which we can interpret as shot noise Cox point process, introduced in Section 3.5.4. The Neyman-Scott point process is formed by first considering a homogeneous Poisson process Φ_c on $\mathbb{R}^2$ with density λ_c. The points of this point process are referred to as the cluster centers and can be used to model, for example, macro base stations. Conditioned on the point process Φ_c, the Neyman-Scott process is then an inhomogeneous Poisson point process Φ_c with intensity function

$$\lambda_{\mathrm{NS}}(y) = \lambda_m \sum_{\mathsf{x}\in\Phi_c} f(x-y), \quad y \in \mathbb{R}^2,$$

where $\lambda_m > 0$ is a parameter and f is some continuous density function. Different choices of the function f lead to specific point processes.

Example 8.4 A *Matérn cluster process* is obtained by setting

$$f(y) = \frac{\mathbf{1}_{\|y\|\le r}}{\pi r^2},$$

where $r > 0$. The nearest-neighbor distance distribution for the Matern cluster process is derived in Afshang, Saha, & Dhillon (2017b).

Example 8.5 A *Thomas process* is obtained by setting

$$f(y) = \frac{\mathrm{e}^{-\|\mathsf{y}\|^2(2\nu^2)}}{\nu\sqrt{2\pi}},$$

where $\nu > 0$. The nearest-neighbor distance distribution is known for the Thomas process (Afshang, Saha, & Dhillon 2017a).

The reduced Palm distribution for the Neyman-Scott process and other shot noise Cox processes can be written down (Møller 2003).

Coverage probability expression

In the zero-noise case, we present the coverage probability of the typical location that connects to its nearest base station at a distance R, which is a continuous random variable with probability density $g_R(r)$. Again, we assume a nonincreasing path loss function ℓ, and write both $\ell(x)$ and $\ell(r)$, where $x \in \mathbb{R}^2$ and $r = \|x\|$.

PROPOSITION 8.1.8 *We assume a path loss function ℓ and Rayleigh fading with unit mean. Then for a stationary Neyman-Scott cluster process Φ on $\mathbb{R}^2$ the coverage probability is*

$$\mathbb{P}\{\,\mathrm{SIR}(\mathbf{x}^*, o) > \tau\,\} = \int_0^\infty \exp\left(-\lambda_c \int_{\mathbb{R}^2} \left[1 - \frac{1}{1+\tau\ell(x)/\ell(r)} \exp\left[-\lambda_m H(x,r,\tau)\right]\right]\right) \mathrm{d}x g_R(r)\mathrm{d}r\,, \quad (8.11)$$

where

$$H(x,r,\tau) = \int_{\mathbb{R}^2} \frac{\tau\ell(y)/\ell(r)}{1+\tau\ell(y)/\ell(r)} f(x-y)\mathrm{d}y.$$

We refer the reader to the paper by Suryaprakash, Møller, & Fettweis (2015) for the proof and more details on the Neyman-Scott process in the network setting.

8.2 Approximate SIR Analysis for General Networks

8.2.1 Motivation

Analyzing the SIR distribution of non-Poisson deployments is a daunting task. Because a general exact analytical approach does not exist, there is a need for good approximations or bounds. A recently developed method called ASAPPP (approximate SIR analysis based on the Poisson point process) is based on the observation that empirical SIR CCDFs for different point processes (while keeping the other network parameters the same) appear to be essentially shifted versions of each other. This is apparent in Fig. 8.2 (Section 8.2.2) for both urban and rural deployments, as well as for the triangular lattice. As a result it appears that the SIR CCDF can be obtained approximately by shifting the curve for the Poisson point process by a relatively small amount.

The extreme case of a regular deployment is the triangular lattice, where the Voronoi cells form hexagons. Fig. 8.1 compares the SIR distributions (CCDFs) for the Poisson point process and the triangular lattice for the case of nearest-base-station association. It is apparent that there is approximately a constant (horizontal) gap of 3.4 dB between the two. This means that for all practical[1] (stationary) deployments of base stations, the

[1] Here, "practical" means repulsive. As stated later (at the end of Section 8.2.4), the curve of the SIR distribution for a stationary cluster process may be lower than the Poisson point process curve.

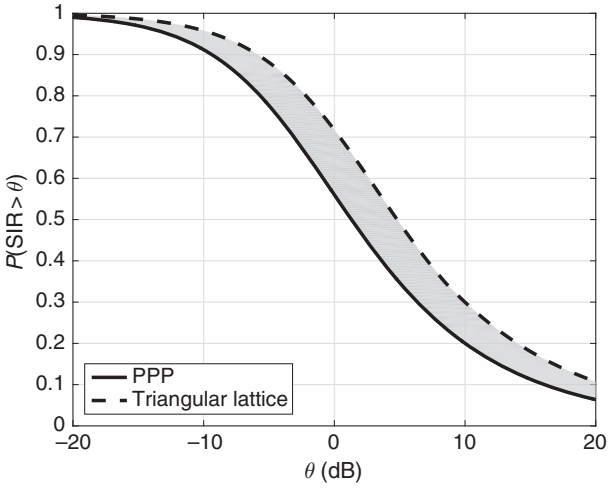

Figure 8.1 SIR CCDF for the Poisson point process and the triangular lattice for $\alpha = 4$ and Rayleigh fading. The gap between the two curves appears to be approximately constant.

SIR distribution falls into the gray band, and the Poisson point process curve is never off by more than 3.4 dB.

8.2.2 Accuracy of the SIR distributions compared with real networks

In this section, we explore how accurately the results for Poisson model represent those of actual deployments of cellular networks. We assume that in real networks (at least for coverage-oriented deployments), base stations are placed more regularly than in a Poisson process; as a consequence, one may believe that the results obtained for the Poisson process are pessimistic. Fig. 8.2 shows the SIR CCDF for the Poisson process, for the (randomly translated) triangular (honeycomb) lattice (which is considered as the optimum deployment) and for two actual base station deployments taken from an urban and a rural region in the United Kingdom,[2] for Rayleigh fading and a path loss exponent of $\alpha = 4$. The Poisson point process curve for this case (with users associating with the nearest base station) is given by Eq. (5.19):

$$\mathcal{P}_{\mathrm{PPP}}(\tau) = \frac{1}{1 + \sqrt{\tau} \arctan \sqrt{\tau}},$$

whereas the other curves are obtained by simulation. It can be seen that the curve for the Poisson process constitutes a lower bound that is reasonably tight. Similar observations were made in Andrews et al. (2011). Lee, Shih, and Chen conducted a study for the case of (instantaneous) max-SIR base station assocation (Lee, Shih, & Chen, 2013). In this case, the CCDFs approach 1 much more quickly as $\tau \to 0$ due to the diversity gain obtained by being able to connect to the instantaneously best base station.

[2] The data used here are the locations of Vodafone base stations in the 900 MHz band, obtained from the Ofcom web page.

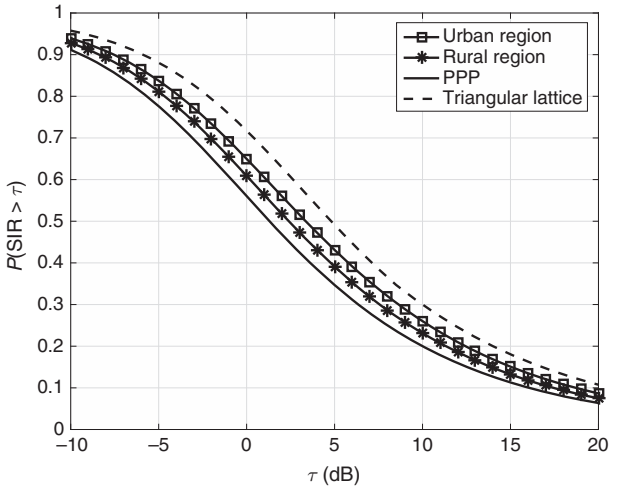

Figure 8.2 SIR CCDF for the Poisson point process, two actual deployments of about 150 base stations, and the triangular lattice for $\alpha = 4$ and Rayleigh fading.

Also, the curve for the urban deployment is about in the middle of the ones for the Poisson point process and the lattice, and it appears that the different curves are very similar in shape and hence essentially shifted versions of each other. This observation was the motivation to develop an approximate SIR analysis for non-Poisson networks based on the idea of shifting the CCDF of the Poisson process. This ASAPPP method is the topic of Section 8.2.3.

The problem of fitting point process to actual base station data has been studied (Guo & Haenggi 2013; Riihijärvi & Mähönen 2012). The specific case of the β-Ginibre process is discussed in Deng et al. (2015a). The values of β giving the best fit for the base station data used in Fig. 8.2 are $\beta = 0.23$ for the rural region and $\beta = 0.93$ for the urban region. This is consistent with the figure because a higher value of β means the point process is more regular, resulting in a better CCDF.

8.2.3 ASAPPP

Let $\mathcal{P}_{\text{PPP}}(\tau) = \bar{F}_{\text{SIR}}(\tau)$ denote the SIR CCDF (or coverage probability) for a Poisson cellular network with nearest-base-station association. Recall that for Rayleigh fading and the standard singular path loss $\ell(\mathbf{x}) = (\kappa\|\mathbf{x}\|)^{-\alpha}$, the coverage probability can be expressed in terms of the gauss hypergeometric function ${}_2F_1$ by Eq. (5.19):

$$\mathcal{P}_{\text{PPP}}(\tau) = \frac{1}{{}_2F_1(1, -\delta; 1-\delta; -\tau)}, \tag{8.12}$$

where $\delta \stackrel{\triangle}{=} 2/\alpha$. For $\alpha = 4$ this simplifies to

$$\mathcal{P}_{\text{PPP}}(\tau) = \frac{1}{1 + \sqrt{\tau}\arctan\sqrt{\tau}}.$$

The SIR CCDF for another stationary and ergodic point process can be written as

$$\mathcal{P}(\tau) = \mathcal{P}_{\text{PPP}}(\tau/G(\tau)),$$

where $G(\tau)$ is the (horizontal) gap between the two curves at τ, that is, the amount by which the Poisson process curve needs to be shifted to the right (in decibels) to obtain the other curve. Formally, the gap is given by

$$G(\tau) \stackrel{\triangle}{=} \frac{\bar{F}^{-1}_{\text{SIR}}(\mathcal{P}_{\text{PPP}}(\tau))}{\tau}, \tag{8.13}$$

where $\bar{F}^{-1}_{\text{SIR}}$ is the inverse of the SIR CCDF of the non-Poisson point process.

Visually, as illustrated in Fig. 8.1, it appears that $G(\tau)$ is essentially a constant; that is, the SIR CCDF for other stationary and ergodic point processes can be approximated closely as

$$\mathcal{P}(\tau) \approx \mathcal{P}_{\text{PPP}}(\tau/G), \tag{8.14}$$

where G is the gain relative to the Poisson process. To support this claim, we explore the limits

$$G_0 \stackrel{\triangle}{=} \lim_{\tau\to 0} G(\tau); \quad G_\infty \stackrel{\triangle}{=} \lim_{\tau\to\infty} G(\tau).$$

If we find that $G_0 \approx G_\infty$, we can be confident that the shifted curve is indeed a good approximation.

Determining G_0

Rayleigh fading

For Rayleigh fading, and with the usual convention of putting the typical location at o and its serving (nearest) base station at x_0, the asymptotic gain G_0 is closely related to the mean interference-to-signal ratio (MISR), which is the mean of the interference-to-(average)-signal ratio:

$$\bar{\mathsf{ISR}} \stackrel{\triangle}{=} \frac{I}{\mathbb{E}_{H_{\mathsf{x}_0}}[S]}.$$

$\mathbb{E}_{H_{\mathsf{x}_0}}[S] = \mathbb{E}[S \mid \Phi]$ is the mean received signal power, averaged only over the fading. Using the ISR, the coverage probability can be written as

$$\begin{aligned}\mathcal{P}(\tau) &= \mathbb{P}\{H_{\mathsf{x}_0}\bar{S} > \tau I)\\ &= \mathbb{P}\{H_{\mathsf{x}_0} > \tau\, \bar{\mathsf{ISR}})\\ &= \mathbb{E}\bar{F}_{H_{\mathsf{x}_0}}(\tau\, \bar{\mathsf{ISR}}).\end{aligned}$$

For Rayleigh fading, $\bar{F}_{H_{\mathsf{x}_0}}(x) \sim 1 - x$, $x \to 0$, hence

$$\mathcal{P}(\tau) \sim 1 - \mathbb{E}[\tau\, \bar{\mathsf{ISR}}] = 1 - \tau\, \mathsf{MISR}, \quad \tau \to 0.$$

From Eq. (8.14), it follows that the asymptotic gain (relative to the Poisson process deployment of base stations) is given by

$$G_0 = \frac{\mathsf{MISR}_{\mathrm{PPP}}}{\mathsf{MISR}} \tag{8.15}$$

because from the asymptotic Poisson process CCDF $1 - \tau\,\mathsf{MISR}_{\mathrm{PPP}}$, replacing τ by τ/G_0, we obtain $1 - \tau\,\mathsf{MISR}$. The MISR does not depend on the fading because, assuming x_0 is the serving (nearest) base station,

$$\mathsf{MISR} = \mathbb{E}\left[\frac{I}{\mathbb{E}_{H_{\mathsf{x}_0}}[S]}\right] = \mathbb{E}\left[\frac{\sum_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} H_{\mathsf{x}}\|\mathsf{x}\|^{-\alpha}}{\mathsf{x}_0^{-\alpha}}\right] = \mathbb{E}\sum_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} \frac{\|\mathsf{x}\|^{-\alpha}}{\|\mathsf{x}_0\|^{-\alpha}}.$$

The MISR for the Poisson point process for the standard path loss law can be calculated easily from the relative distance process, introduced in Def. 3.4.5. Letting x_0 denote the closest point of the Poisson point process to the origin, we have $\mathbb{E}_{H_{\mathsf{x}_0}}[S] = \|\mathsf{x}_0\|^{-\alpha}$ and thus

$$\begin{aligned}
\mathsf{MISR}_{\mathrm{PPP}} &= \mathbb{E}\sum_{\mathsf{x}\in\Phi\setminus\{\mathsf{x}_0\}} \frac{\|\mathsf{x}_0\|^{\alpha}}{\|\mathsf{x}\|^{\alpha}} \\
&= \sum_{\mathsf{y}\in\mathcal{R}} \mathbb{E}(\mathsf{y}^{\alpha}) \\
&\overset{\text{(a)}}{=} \int_0^1 r^{\alpha} 2r^{-3}\,\mathrm{d}r \\
&= \frac{2}{\alpha-2},
\end{aligned} \tag{8.16}$$

where (a) follows from Campbell's formula in Eq. (3.9) and the intensity measure of the RDP of a Poisson process in Eq. (3.27).

The MISR for the Poisson point process and the square and triangular lattices is shown in Fig. 8.3 (*left*). For other point processes, the MISR may have to be obtained by simulation, which can be done efficiently because fading is not involved.

General fading and diversity

For general fading models and transmission schemes with diversity, the gain G_0 is obtained as the ratio of the corresponding moments of the ISR. We first define the generalized MISR.

DEFINITION 8.2.1 (Generalized MISR) *The* generalized MISR *with parameter n is defined as*

$$\mathsf{MISR}_n \triangleq (\mathbb{E}(\bar{\mathsf{ISR}}^n))^{1/n}.$$

The generalized MISR can be obtained by taking the corresponding derivative of the Laplace transform $\mathbb{E}(\mathrm{e}^{-s\bar{\mathsf{ISR}}})$ at $s = 0$. In case of the Poisson process with Rayleigh fading, the Laplace transform is known and equals the coverage probability (8.12), thus

$$\mathsf{MISR}_n^n = \mathbb{E}(\mathsf{I\bar{S}R}^n) = (-1)^n \frac{\mathrm{d}}{\mathrm{d}\tau^n} \mathcal{P}(\tau)\big|_{\tau=0}. \tag{8.17}$$

For the Poisson process with arbitrary fading, the generalized MISR can be calculated explicitly.

THEOREM 8.2.2 (Generalized MISR and lower bound for Poisson point process) *For a Poisson cellular network with arbitrary fading,*

$$\mathbb{E}(\mathsf{I\bar{S}R}^n) = \sum_{k=1}^{n} k!\, B_{n,k}\left(\frac{\delta}{1-\delta}, \ldots, \frac{\delta \mathbb{E}[H_{\mathsf{x}_0}^{n-k+1}]}{n-k+1-\delta}\right), \tag{8.18}$$

where $B_{n,k}$ are the (incomplete) Bell polynomials. For $n > 1$, the generalized MISR is lower bounded as

$$\mathsf{MISR}_n \geq \left[\left(\frac{\delta}{1-\delta}\right)^n n! + \frac{\delta \mathbb{E}[H_{\mathsf{x}_0}^n]}{n-\delta}\right]^{1/n}. \tag{8.19}$$

For $n = 2$, equality holds, and for $\delta \to 0$ and $\delta \to 1$, the lower bound is asymptotically tight.

We can also derive two simpler asymptotically tight bounds for the generalized MISR:

$$\mathsf{MISR}_n \gtrsim \left(\frac{\delta}{n} \mathbb{E}[H_{\mathsf{x}_0}^n]\right)^{1/n}, \quad \delta \to 0 \tag{8.20}$$

$$\mathsf{MISR}_n \gtrsim \frac{\delta (n!)^{1/n}}{1-\delta} = \mathsf{MISR}_1 (n!)^{1/n}, \quad \delta \to 1. \tag{8.21}$$

For Rayleigh fading, Eq. (8.20) yields $\mathsf{MISR}_n \sim (\delta \Gamma(n))^{1/n}$, $\delta \to 0$.

The term $\mathsf{MISR}_1 (n!)^{1/n}$ is dominant for $\alpha \leq 4$ even if the fading is severe (Rayleigh fading). For less severe fading, the term with $\mathbb{E}[H_{\mathsf{x}_0}^n]$ is less relevant; it only becomes dominant for unrealistically high path loss exponents ($\delta \ll 1$).

The second moment of the ISR follows from Eq. (8.18) as

$$\mathbb{E}(\mathsf{I\bar{S}R}^2) = 2\,\mathsf{MISR}_1^2 + \frac{\delta \mathbb{E}[H_{\mathsf{x}_0}^2]}{2-\delta},$$

and the third moment is

$$\mathbb{E}(\mathsf{I\bar{S}R}^3) = 6\,\mathsf{MISR}_1^3 + \frac{6\delta^2 \mathbb{E}[H_{\mathsf{x}_0}^2]}{(1-\delta)(2-\delta)} + \frac{\delta \mathbb{E}[H_{\mathsf{x}_0}^3]}{3-\delta}.$$

COROLLARY 8.2.3 *For the Poisson process with Rayleigh fading and $\alpha \leq 4$,*

$$\mathsf{MISR}_n \sim \frac{n}{e}\,\mathsf{MISR}_1 = \frac{n}{e}\frac{\delta}{1-\delta}, \quad n \to \infty.$$

Proof For the Poisson process with Rayleigh fading and $\delta \geq 1/2$, it follows from Eq. (8.18) that

$$\mathsf{MISR}_n \sim \left(\frac{\delta}{1-\delta}\right)(n!)^{1/n}, \quad n \to \infty,$$

since the dominant term in Eq. (8.18) for large n is the one with $\delta^n/(1-\delta)^n$, which increases geometrically (or stays constant) with n for $\delta \geq 1/2$. For the factorial term, $\log((n!)^{1/n}) \sim \log n - 1$, hence we obtain $\mathsf{MISR}_n \sim e^{\log n - 1}\,\mathsf{MISR}_1$. □

Remark. Using Stirling's formula $n! \sim \sqrt{2\pi n}(n/e)^n$, this asymptotic result can be sharpened slightly.

Equipped with the results from Theorem 8.2.2, we can now discuss the gain G_0 for general fading.[3] If $F_{H_{\mathsf{x}_0}}(x) \sim c_m x^m$, $x \to 0$, then, for $\tau \to 0$, we have $\mathcal{P}(\tau) \sim 1 - c_m\,\mathbb{E}[(\tau\,\mathsf{I\bar{S}R})^m]$, hence

$$G_0^{(m)} = \left(\frac{\mathbb{E}[\mathsf{I\bar{S}R}_{\mathrm{PPP}}^m]}{\mathbb{E}[\mathsf{I\bar{S}R}^m]}\right)^{1/m} = \frac{\mathsf{MISR}_{m,\mathrm{PPP}}}{\mathsf{MISR}_m}. \tag{8.22}$$

The ASAPPP approximation follows as

$$\mathcal{P}(\tau) \approx \mathcal{P}_{\mathrm{PPP}}^{(m)}(\tau/G_0^{(m)}),$$

where $\mathcal{P}_{\mathrm{PPP}}^{(m)}$ is the coverage probability for the Poisson process with fading parameter m, which is not known in closed form. However, we have the exact MISR_m from Eq. (8.18) and the lower bound $\mathsf{MISR}_m \gtrsim \mathsf{MISR}_1(m!)^{1/m}$.

For Nakagami-m fading, the preconstant is $c_m = m^{m-1}/\Gamma(m)$, and we have

$$\begin{aligned}\mathcal{P}_{\mathrm{PPP}}^{(m)}(\tau) &\sim 1 - c_m\mathbb{E}[(\tau\,\mathsf{I\bar{S}R})^m] \\ &\lesssim 1 - \frac{m^{m-1}}{\Gamma(m)}\mathsf{MISR}_1 m!\,\tau^m \\ &= 1 - \mathsf{MISR}_1(m\tau)^m,\end{aligned}$$

where $\lesssim$ indicates an upper bound with asymptotic equality. Adding the second term in the lower bound and noting that

$$\mathbb{E}[H_{\mathsf{x}_0}^m] = \frac{\Gamma(2m)}{\Gamma(m)m^m}$$

yields the slightly sharper result

$$\mathcal{P}_{\mathrm{PPP}}^{(m)}(\tau) \lesssim 1 - \tau^m\left[\left(\frac{m\delta}{1-\delta}\right)^m + \frac{\delta}{m-\delta}\frac{\Gamma(2m)}{\Gamma(m)m^m}\right].$$

The gain for general fading is applicable to arbitrary transmission techniques that provide the same amount of diversity, not just to compare different base station deployments.

Determining G_∞

Here we define the *expected fading-to-interference ratio (EFIR)* and explore its connection to the gain G_∞. We see that the EFIR plays a similar role for $\tau \to \infty$ as the MISR does for $\tau \to 0$.

[3] General fading refers to a fading distribution that satisfies $F_{H_{\mathsf{x}_0}}(x) = \Theta(x^m)$, $x \to 0$, for arbitrary $m \in \mathbb{N}$.

Definition and EFIR for Poisson base station deployment

DEFINITION 8.2.4 (Expected fading-to-interference ratio [EFIR]) *For a point process Φ, let $I_\infty = \sum_{\mathsf{x}\in\Phi} H_{\mathsf{x}}\|\mathsf{x}\|^{-\alpha}$ and let H be a fading random variable independent of all (H_{x}). The EFIR is defined as*

$$\mathsf{EFIR} \stackrel{\triangle}{=} \left(\lambda\pi\mathbb{E}_o^!\left[\left(\frac{H}{I_\infty}\right)^\delta\right]\right)^{1/\delta}, \tag{8.23}$$

where $\mathbb{E}_o^!$ is the expectation with respect to the reduced Palm measure of Φ.

Here we use I_∞ for the interference term because the interference here is the total received power from all points in Φ, in contrast to the interference I, which stems from $\Phi^!$.

Remark. For the Poisson base station deployment, the EFIR does not depend on λ, because $\mathbb{E}_o^!(I_\infty^{-\delta}) \propto 1/\lambda$. To see this, let $\Phi' \stackrel{\triangle}{=} c\Phi$ be a scaled version of Φ. Then

$$I_c \stackrel{\triangle}{=} \sum_{x\in\Phi'} H_{\mathsf{x}}\|\mathsf{x}\|^{-\alpha} = c^{-\alpha}\sum_{\mathsf{x}\in\Phi} H_{\mathsf{x}}\|\mathsf{x}\|^{-\alpha}$$

and thus $I_c^{-\delta} = c^2 I^{-\delta}$. Multiplying by the intensities, $\lambda_c I_c^{-\delta} = \lambda I^{-\delta}$ because $\lambda/\lambda_c = c^2$. The same argument applies to all point processes for which changing the intensity by a factor c^{-2} is equivalent in distribution to scaling the process by c; that is, for point processes $\Phi(\lambda)$ where $c\Phi(1) \stackrel{\mathrm{d}}{=} \Phi(c^{-2})$. This excludes hard-core processes with fixed hard-core distance but includes lattices and hard-core processes whose hard-core distance scales with $\lambda^{-1/2}$.

LEMMA 8.2.5 (EFIR for the Poisson base station deployment) *For the Poisson base station deployment, with arbitrary fading,*

$$\mathsf{EFIR}_{\mathrm{PPP}} = (\operatorname{sinc}\delta)^{1/\delta}. \tag{8.24}$$

Proof The term $\mathbb{E}_o^!(I_\infty^{-\delta})$ in Eq. (8.23) can be calculated by taking the expectation of the following identity which follows from the definition of the gamma function $\Gamma(x)$.

$$I_\infty^{-\delta} \equiv \frac{1}{\Gamma(\delta)}\int_0^\infty e^{-sI_\infty}s^{-1+\delta}\mathrm{d}s.$$

Hence

$$\mathbb{E}_o^!(I_\infty^{-\delta}) = \frac{1}{\Gamma(\delta)}\int_0^\infty \mathcal{L}_{o,I_\infty}^!(s)s^{-1+\delta}\mathrm{d}s. \tag{8.25}$$

From Slivnyak's theorem in Eq. (3.21), $\mathbb{E}_o^! \equiv \mathbb{E}$ for the Poisson point process, so we can replace $\mathcal{L}_{o,I_\infty}^!(s)$ by the unconditioned Laplace transform $\mathcal{L}_{I_\infty}(s)$, which is well known for the Poisson point process and given by Haenggi & Ganti (2008)

$$\mathcal{L}_{I_\infty}(s) = \exp(-\lambda\pi\mathbb{E}[H^\delta]\Gamma(1-\delta)s^\delta).$$

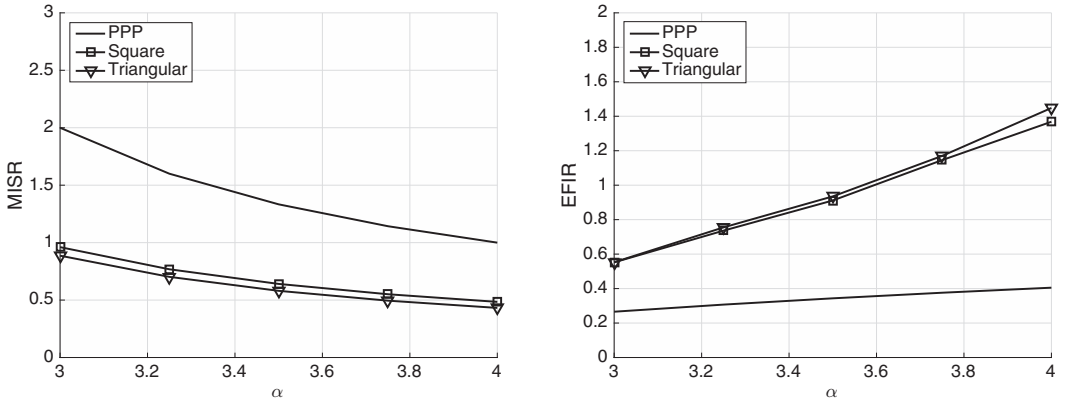

Figure 8.3 MISR (*left*) and EFIR (*right*) for the Poisson process and the square and triangular lattices as a function of α. The curves for Poisson process are from Eqs. (8.16) and (8.24), respectively, whereas those for the lattices are obtained by simulation.

From Eq. (8.25), we have

$$\mathbb{E}[I_\infty^{-\delta}] = \frac{1}{\Gamma(\delta)} \int_0^\infty e^{-\lambda\pi\mathbb{E}[H^\delta]\Gamma(1-\delta)s^\delta} s^{-1+\delta} \mathrm{d}s$$

$$= \frac{1}{\lambda\pi\mathbb{E}[H^\delta]\Gamma(1-\delta)\Gamma(1+\delta)} = \frac{\operatorname{sinc}\delta}{\lambda\pi\mathbb{E}(H^\delta)}.$$

So $\lambda\pi\mathbb{E}_o^!(I_\infty^{-\delta})\mathbb{E}[H^\delta] = \operatorname{sinc}\delta$, and the result follows. □

Remarkably, $\mathsf{EFIR}_{\mathrm{PPP}}$ only depends on the path loss exponent. It can be closely approximated by $\mathsf{EFIR}_{\mathrm{PPP}} \approx 1 - \delta$.

The EFIR for the Poisson process and the square and triangular lattices is shown in Fig. 8.3 (*right*).

The tail of the SIR distribution

Let $f(R, \Phi^!)$ be a positive function of the distance $R = \|\mathsf{x}_0\|$ and the point process $\Phi^! = \Phi \setminus \{\mathsf{x}_0\}$. The average $\mathbb{E}[f(R, \Phi^!)]$ can in principle be evaluated using the joint distribution of R and Φ, which is, however, known only for very few point processes. Thus we introduce an alternative representation of $f(R, \Phi^!)$ that is easier to work with.

The indicator variable $\mathbf{1}(\Phi(b(o, \|\mathsf{x}\|)) = 0)$, $\mathsf{x} \in \Phi$, equals 1 only when $\mathsf{x} = \mathsf{x}_0$. Hence it follows that

$$f(R, \Phi^!) \equiv \sum_{\mathsf{x}\in\Phi} f(\|\mathsf{x}\|, \Phi \setminus \{\mathsf{x}\})\mathbf{1}(\Phi(b(o, \|\mathsf{x}\|)) = 0). \tag{8.26}$$

This representation of $f(R, \Phi^!)$ permits the computation of the expectation of $f(R, \Phi^!)$ using the Campbell-Mecke theorem (Theorem 3.32).

Next we use the representation in Eq. (8.26) to analyze the tail asymptotics of the CCDF $\bar{F}_{\mathrm{SIR}}$ of the SIR (or, equivalently, the coverage probability $\mathcal{P}(\tau)$).

THEOREM 8.2.6 *For all simple stationary and ergodic base station point processes Φ, satisfying condition (B) in Theorem 2.1 in Miyoshi & Shirai (2017) where a user at the typical location is served by the nearest base station,*

$$\mathcal{P}(\tau) \sim \left(\frac{\tau}{\mathsf{EFIR}}\right)^{-\delta}, \quad \tau \to \infty.$$

Proof The coverage probability is given by $\mathcal{P}(\tau) = \mathbb{E}\bar{F}_H(\tau R^\alpha I)$. Using the representation given in Eq. (8.26), it follows from the Campbell-Mecke theorem that the coverage probability equals

$$\mathbb{E}\sum_{\mathsf{x}\in\Phi}\bar{F}_H\left(\tau\|\mathsf{x}\|^\alpha\sum_{\mathsf{y}\in\Phi\setminus\{\mathsf{x}\}}H_{\mathsf{y}}\|\mathsf{y}\|^{-\alpha}\right)\mathbf{1}\big(\Phi(b(o,\|\mathsf{x}\|))=0\big)$$
$$=\lambda\int_{\mathbb{R}^2}\mathbb{E}_o^!\left[\bar{F}_H\Big(\tau\|\mathsf{x}\|^\alpha\sum_{y\in\Phi_{\mathsf{x}}}H_y\|y\|^{-\alpha}\Big)\mathbf{1}(\Phi_x(b(o,\|\mathsf{x}\|))=0)\right]\mathrm{d}\mathsf{x},$$

where $\Phi_{\mathsf{x}} \triangleq \{\mathsf{y}\in\Phi\colon \mathsf{y}+\mathsf{x}\}$ is a translated version of Φ. Substituting $\mathsf{x}\tau^{\delta/2}\mapsto\mathsf{x}$,

$$\mathcal{P}(\tau)=\lambda\tau^{-\delta}\int_{\mathbb{R}^2}\mathbb{E}_o^!\left[\bar{F}_H\Big(\|\mathsf{x}\|^\alpha\sum_{\mathsf{y}\in\Phi_{\mathsf{x}\tau^{-\delta/2}}}H_{\mathsf{y}}\|\mathsf{y}\|^{-\alpha}\Big)\mathbf{1}(\Phi_x(b(o,\|\mathsf{x}\|\tau^{-\delta/2}))=0)\right]\mathrm{d}\mathsf{x}$$
$$\overset{(a)}{\sim}\lambda\tau^{-\delta}\int_{\mathbb{R}^2}\mathbb{E}_o^!\bar{F}\left(\|\mathsf{x}\|^\alpha I_\infty\right)\mathrm{d}\mathsf{x},\quad \tau\to\infty \tag{8.27}$$
$$\overset{(b)}{=}\lambda\tau^{-\delta}\mathbb{E}_o^![I_\infty^{-\delta}]\int_{\mathbb{R}^2}\bar{F}_H\left(\|\mathsf{x}\|^\alpha\right)\mathrm{d}\mathsf{x},\quad \tau\to\infty,$$

where (a) follows because $\tau^{-\delta/2}\to 0$ and hence $\mathbf{1}\{\Phi_x(b(o,\|\mathsf{x}\|\tau^{-\delta/2}))=0\}\to 1$, provided some uniform integrability condition holds allowing one to move the limit under the reduced Palm expectation (see Theorem 2.1 in Miyoshi & Shirai [2017] for more details). The equality in (b) follows by using the substitution $\mathsf{x}I^{1/\alpha}\to\mathsf{x}$. Changing into polar coordinates, the integral can be written as

$$\int_{\mathbb{R}^2}\bar{F}_H\left(\|\mathsf{x}\|^\alpha\right)\mathrm{d}\mathsf{x}=\pi\delta\int_0^\infty r^{\delta-1}\bar{F}_H(r)\mathrm{d}r\overset{(a)}{=}\pi\mathbb{E}[H^\delta],$$

where (a) follows because $H\geq 0$. Because $\mathbb{E}[H]=1$ and $\delta<1$, it follows that $\mathbb{E}[H^\delta]<\infty$. □

Condition (B) in Theorem 2.1 in Miyoshi & Shirai (2017) is a very mild one. In fact, no stationary and ergodic point process is known that violates the condition.

For Rayleigh fading, from the definition of the coverage probability and Theorem 8.2.6,

$$\mathcal{P}(\tau)=\mathcal{L}_{\bar{\mathsf{ISR}}}(\tau)\sim\left(\frac{\tau}{\mathsf{EFIR}}\right)^{-\delta},\quad\tau\to\infty.$$

Hence the Laplace transform of $\bar{\mathsf{ISR}}$ behaves as $\Theta(\tau^{-\delta})$ for large τ. Hence using the Tauberian theorem (Feller 1970, p. 445), we can infer that

$$\mathbb{P}(\mathsf{I\bar{S}R} < x) \sim x^{\delta} \frac{\mathsf{EFIR}^{\delta}}{\Gamma(1+\delta)}, \quad x \to 0. \tag{8.28}$$

From Theorem 8.2.6, the gain G_{∞} immediately follows.

COROLLARY 8.2.7 (Asymptotic gain at $\tau \to \infty$) *For an arbitrary simple stationary point process Φ with EFIR given in Def. 8.2.4, the asymptotic gain at $\tau \to \infty$ relative to the Poisson point process is*

$$G_{\infty} = \frac{\mathsf{EFIR}}{\mathsf{EFIR}_{\mathrm{PPP}}} = \left(\frac{\lambda \pi \mathbb{E}_o^!\left[I_{\infty}^{-\delta}\right] \mathbb{E}[H^{\delta}]}{\operatorname{sinc} \delta} \right)^{1/\delta}.$$

The Laplace transform of the interference in Equation (8.25) for general point processes can be expressed as

$$\begin{aligned} \mathcal{L}_{o,I_{\infty}}^{!}(s) &= \mathbb{E}_o^! \left(e^{-s \sum_{\mathsf{x} \in \Phi} H_{\mathsf{x}} \|\mathsf{x}\|^{-\alpha}} \right) \\ &= \mathbb{E}_o^! \prod_{\mathsf{x} \in \Phi} \mathcal{L}_H(s\|\mathsf{x}\|^{-\alpha}) = \mathcal{G}_o^![\mathcal{L}_H(s\|\cdot\|^{-\alpha})], \end{aligned}$$

where $\mathcal{G}_o^!$ is the pfgl with respect to the reduced Palm measure and $\mathcal{L}_H$ is the Laplace transform of the fading distribution.

COROLLARY 8.2.8 (Rayleigh fading) *With Rayleigh fading, the EFIR simplifies to*

$$\mathsf{EFIR} = \left(\lambda \int_{\mathbb{R}^2} \mathcal{G}_o^![\Delta(\mathsf{x}, \cdot)] \, \mathrm{d}\mathsf{x} \right)^{1/\delta},$$

where

$$\Delta(\mathsf{x}, \mathsf{y}) = \frac{1}{1 + \|\mathsf{x}\|^{\alpha} \|\mathsf{y}\|^{-\alpha}}.$$

Proof With Rayleigh fading, the power fading coefficients are exponential, so $\bar{F}_H(x) = \exp(-x)$. From Eq. (8.27), we have

$$\begin{aligned} \mathcal{P}(\tau) &\sim \lambda \tau^{-\delta} \int_{\mathbb{R}^2} \mathbb{E}_o^! \bar{F}\left(\|\mathsf{x}\|^{\alpha} I\right) \, \mathrm{d}\mathsf{x} \\ &= \lambda \tau^{-\delta} \int_{\mathbb{R}^2} \mathbb{E}_o^! \prod_{\mathsf{y} \in \Phi} \frac{1}{1 + \|\mathsf{x}\|^{\alpha} \|\mathsf{y}\|^{-\alpha}} \, \mathrm{d}\mathsf{x}, \end{aligned}$$

and the result follows from the definition of the pgfl with respect to the reduced Palm measure. □

Tail of received signal strength

While Theorem 8.2.6 shows that $\mathcal{P}(\tau) = \Theta(\tau^{-\delta}), \tau \to \infty$, it is not clear whether the scaling is mainly due to the received signal strength or to the interference. Intuitively, because an infinite network is considered, the event that the interference is small is negligible and hence for large τ, the event $S/I > \tau$ is mainly determined by the random variable S. This is in fact true, as is shown in the next lemma.

LEMMA 8.2.9 *For all stationary point processes and arbitrary fading, the tail of the CCDF of the desired signal strength S is*

$$\mathbb{P}\{S > \tau\} \sim \lambda\pi\mathbb{E}[H^{\delta}]\tau^{-\delta}, \quad \tau \to \infty.$$

Proof The CDF of the distance R to the nearest base station is $F_R(x) \sim \lambda\pi x^2$ for all stationary point processes (Chiu et al. 2013). Hence

$$\begin{aligned}\mathbb{P}\{S > \tau\} &= \mathbb{P}(R < (H/\tau)^{\delta/2}) \\ &\sim \lambda\pi\mathbb{E}[(H/\tau)^{\delta}].\end{aligned}$$

□

So the tail of the received signal power S is of the same order $\Theta(\tau^{-\delta})$, and the interference and the fading affect only the preconstant. In the Poisson case with Rayleigh fading,

$$\mathcal{P}(\tau) \sim \lambda\pi\Gamma(1+\delta)\tau^{-\delta}, \quad \tau \to \infty.$$

The same holds near $\tau = 0$. If for the fading CDF, $F_H(x) \sim ax^m$, $x \to 0$,

$$\mathbb{P}\{S < \tau\} = \mathbb{E}F_H(\tau R^{\alpha}) \sim a\tau^m\mathbb{E}(R^{m\alpha}), \quad \tau \to 0.$$

For the Poisson base station deployment,

$$\mathbb{P}\{S < \tau\} \sim \frac{\Gamma(1+m\alpha/2)}{(\lambda\pi)^{m\alpha/2}}\tau^m, \quad \tau \to 0.$$

So on both ends of the SIR distribution, the interference only affects the preconstant.

8.2.4 Why is ASAPPP so effective?

The ASAPPP method is very effective as an approximation of the SIR distribution because it requires a single parameter only: the gain or shift relative to the Poisson point process. As can be seen from Fig. 8.4, the gain G_0 is rather insensitive to α, whereas G_∞

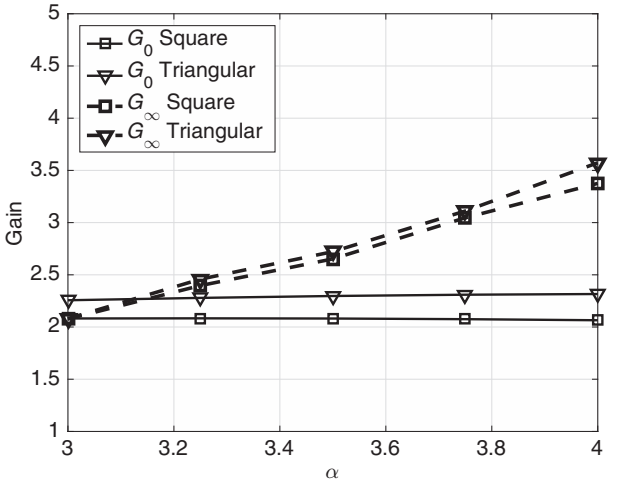

Figure 8.4 Gains G_0 and G_∞ for the square and triangular lattices with Rayleigh fading as a function of α.

increases slightly. Because $G_0 \approx G_\infty$, it can be expected that the gap $G(\tau)$ as defined in Eq. (8.13) does not vary much over the entire range of τ. Indeed, $\mathcal{P}_{\text{PPP}}(\tau/G_0)$ provides an excellent approximation for all practical coverage probabilities. Furthermore, the gains are also insensitive to fading parameter m in Nakagami-m fading. For $m > 1$, the SIR CCDF for the Poisson process is not known in closed form.

For coverage-oriented deployments, where base stations are placed more regularly than a Poisson point process, $G_0 > 1$, which means that the results for the Poisson process give a lower bound for other point processes. In capacity-oriented deployments, the base stations may be more clustered than a Poisson process, which results in a gain G smaller than 1, that is, a negative shift in decibels in the SIR CCDF.

8.2.5 ASAPPP for HetNets

Wei and colleagues propose a HetNet extension of the ASAPPP method (Wei, Deng, Zhou, & Haenggi 2016). If the gains for the individual tiers are properly combined, a very accurate approximation of the SIR distribution can be obtained for HetNets composed of arbitrary stationary point processes (with different intensities, transmit powers, and path loss exponents) – as long as they are independent.

The general HetNet model considered here consists of K tiers of independent stationary and ergodic base station processes of arbitrary type, such as clustered, Poisson, determinantal, hard-core, and Gibbs point processes or lattices. Each tier has a density λ_k and a transmit power P_k, $k \in [K]$. A user at the typical location is served by the base station that provides the strongest power (on average), whereas all other base stations interfere. For ease of exposition, we assume that the path loss exponent is the same for all tiers. The analysis can be extended in a straightforward manner to the case of different path loss exponents (Wei et al. 2016).

THEOREM 8.2.10 (Approximate coverage probability for general HetNets) *Let* $T_\delta(\tau) \stackrel{\triangle}{=} {}_2F_1(1, -\delta; 1-\delta; -\tau)$ *and*

$$\hat{\mathcal{P}}(\tau) \stackrel{\triangle}{=} \sum_{k\in[K]} \frac{1}{T_\delta(\tau/G_k) + \sum_{i\in[K]\setminus\{k\}} \frac{\lambda_i}{\lambda_k}\left(\frac{P_i}{P_k}\right)^\delta T_\delta(\tau)}, \tag{8.29}$$

where G_k is the ASAPPP gain if only tier k existed. Then $\mathcal{P}(\tau) \approx \hat{\mathcal{P}}(\tau)$. Moreover, if $\min G_k \geq 1$, $\mathcal{P}(\tau) \gtrsim \hat{\mathcal{P}}(\tau)$, *where $\gtrsim$ denotes an asymptotic lower bound, that is,* $\exists t > 0$ *s.t.* $(\forall \tau < t)\, \mathcal{P}(\tau) \geq \hat{\mathcal{P}}(\tau)$. *Conversely, if* $\max G_k \leq 1$, $\mathcal{P}(\tau) \lesssim \hat{\mathcal{P}}(\tau)$, *where $\lesssim$ denotes an asymptotic upper bound, that is,* $\exists t > 0$ *s.t.* $(\forall \tau < t)\, \mathcal{P}(\tau) \leq \hat{\mathcal{P}}(\tau)$.

Proof Let x_0 be the serving base station, that is,

$$\mathsf{x}_0 = \arg\max \left\{ \mathsf{x} \in \bigcup_{k\in[K]} \Phi_k \colon P_{\text{tier}(\mathsf{x})}\ell(\mathsf{x}) \right\},$$

where tier(x) returns the tier that x belongs to. We calculate the joint probability that the SIR exceeds τ and the serving base station belongs to tier k. We define $\Phi^! \stackrel{\triangle}{=} \Phi\backslash\{\text{NP}(\Phi)\}$, where NP($\Phi$) is the point of Φ closest to the origin, $\mathcal{I}_{i,k}(x) \stackrel{\triangle}{=} \mathbf{1}(P_i\ell(x) \leq P_k\ell(\mathsf{x}_0))$, and we let Ψ_k, $k \in [K]$, be independent Poisson point processes of intensities λ_k. We have

$$
\begin{aligned}
&\mathbb{P}\{\mathrm{SIR} > \tau,\, x_0 \in \Phi_k\} \\
&= \mathbb{E}\left[\exp\left(-\tau \frac{\sum\limits_{\mathsf{x}\in\Phi_k^!} P_k \ell(\mathsf{x}) H_{\mathsf{x}} + \sum\limits_{i\in[K]\setminus\{k\}} \sum\limits_{\mathsf{y}\in\Phi_i} P_i \ell(\mathsf{y}) H_{\mathsf{y}}}{P_k \ell(\mathsf{x}_0)}\right) \mathbf{1}(\mathsf{x}_0 \in \Phi_k)\right] \\
&= \mathbb{E}\left[\prod_{\mathsf{x}\in\Phi_k^!} \left(1 + \frac{\tau \ell(\mathsf{x})}{\ell(\mathsf{x}_0)}\right)^{-1} \prod_{i\in[K]\setminus\{k\}} \prod_{\mathsf{y}\in\Phi_i} \left(1 + \frac{\tau P_i \ell(\mathsf{y})}{P_k \ell(\mathsf{x}_0)}\right)^{-1} \mathcal{I}_{i,k}(\mathsf{y})\right] \\
&\overset{(a)}{\sim} \mathbb{E}\left[\prod_{\mathsf{x}\in\Psi_k^!} \left(1 + \frac{\tau \ell(\mathsf{x})}{G_k \ell(\mathsf{x}_0)}\right)^{-1} \prod_{i\in[K]\setminus\{k\}} \prod_{\mathsf{y}\in\Phi_i} \left(1 + \frac{\tau P_i \ell(\mathsf{y})}{P_k \ell(\mathsf{x}_0)}\right)^{-1} \mathcal{I}_{i,k}(\mathsf{y})\right] \\
&\overset{(b)}{\approx} \mathbb{E}\left[\prod_{\mathsf{x}\in\Psi_k^!} \left(1 + \frac{\tau \ell(\mathsf{x})}{\ell(\mathsf{x}_0)}\right)^{-1} \prod_{i\in[K]\setminus\{k\}} \prod_{\mathsf{y}\in\Psi_i} \left(1 + \frac{\tau P_i \ell(\mathsf{y})}{P_k \ell(\mathsf{x}_0)}\right)^{-1} \mathcal{I}_{i,k}(\mathsf{y})\right] \\
&\overset{(c)}{=} \int_0^\infty \exp\left\{-2\pi\lambda_k \int_r^\infty \frac{t\,\mathrm{d}t}{1 + \frac{G_k t^\alpha}{\tau r^\alpha}}\right. \\
&\qquad \left. - \sum_{i\in[K]\setminus\{k\}} \pi\lambda_i \left[r^2 \left(\frac{P_i}{P_k}\right)^\delta + \int_{r\left(\frac{P_i}{P_k}\right)^{1/\alpha}}^\infty \frac{2t\,\mathrm{d}t}{1 + \frac{P_k t^\alpha}{\tau P_i r^\alpha}}\right]\right\} f_k(r)\,\mathrm{d}r \\
&= \int_0^\infty 2\pi\lambda_k r \exp\left[-\pi\lambda_k r^2 T_\delta(\tau/G_k) - \sum_{i\in[K]\setminus\{k\}} \pi\lambda_i r^2 \left(\frac{P_i}{P_k}\right)^\delta T_\delta(\tau)\right]\mathrm{d}r \\
&= \int_0^\infty \exp\left[-rT_\delta(\tau/G_k) - \sum_{i\in[K]\setminus\{k\}} \frac{\lambda_i}{\lambda_k} \left(\frac{P_i}{P_k}\right)^\delta rT_\delta(\tau)\right]\mathrm{d}r.
\end{aligned}
$$

Here $f_k(r) = 2\lambda_k \pi r \mathrm{e}^{-\lambda_k \pi r^2}$ is the probability density of $|\mathrm{NP}(\Psi_k)|$. In step (a), the single-tier ASAPPP method is used to replace Φ_k by the Poisson process Ψ_k and scaling τ by G_k, in (b), the interference from Φ_i is approximated by that of Ψ_i, and in (c), the pgfl of the Poisson point process is used. Summing over $[K]$ yields the final result of Eq. (8.29).

In step (b), if all (nonserving) tiers are more regular than the Poisson process, that is, if $G_k > 1$, then we obtain an asymptotic lower bound because the actual coverage probability is slightly higher. Conversely, if all tiers are more clustered than the Poisson process, that is, if $G_k < 1$, then we obtain an asymptotic upper bound. □

Remark 8.1 Because $\mathcal{P}_{\mathrm{PPP}}(\tau) = ({}_2F_1(1, -\delta; 1-\delta; -\tau))^{-1} = 1/T_\delta(\tau)$ (see Eq. [8.12]) is the SIR CCDF of a (single-tier) Poisson process, we may write Eq. (8.29) as

$$
\hat{\mathcal{P}}(\tau) = \mathcal{P}_{\mathrm{PPP}}(\tau) \sum_{k\in[K]} \frac{1}{\frac{\mathcal{P}_{\mathrm{PPP}}(\tau)}{\mathcal{P}^{(k)}(\tau)} + \sum_{i\in[K]\setminus\{k\}} \frac{\lambda_i}{\lambda_k}\left(\frac{P_i}{P_k}\right)^\delta},
$$

where $\mathcal{P}^{(k)}(\tau) = 1/T_\delta(\tau/G_k)$ is the approximate coverage probability if only tier k existed. Hence if all tiers are Poisson point processes, that is, $G_k \equiv 1$, $\hat{\mathcal{P}}(\tau) = \mathcal{P}_{\text{PPP}}(\tau)$, and the approximation yields the exact result.

Alternatively, letting

$$w_k \stackrel{\triangle}{=} \frac{\lambda_k P_k^\delta}{\sum_{i\in[K]} \lambda_i P_i^\delta},$$

we obtain Eq. (8.29) in the form

$$\hat{\mathcal{P}}(\tau) = \sum\nolimits_{k\in[K]} \frac{w_k}{w_k T_\delta(\tau/G_k) + (1-w_k) T_\delta(\tau)}. \tag{8.30}$$

Because $\sum_{k\in[K]} w_k = 1$, the sequence (w_k) can be interpreted as a probability mass function. For the case where the locations of base stations in the tiers are assumed to be points of independent homogeneous PPPs (called the homogeneous independent Poisson model), w_k is the probability that a user at the typical location is associated with a base station of tier k.

Because the HetNet architecture does not provide a diversity gain relative to single-tier networks, the ASAPPP approach is applicable; that is, there exists an SIR shift parameter that, if applied to the Poisson point process, yields an accurate approximation for the SIR CCDF of the HetNet. Because the HetNet consists of several tiers, each with its own gain, the question is how to calculate the overall gain, henceforth called the *effective gain*, of the HetNet, such that the SIR distribution for the HetNet is approximated as

$$\bar{F}_{\text{SIR}}(\tau) \approx \bar{F}_{\text{SIR}}^{\text{PPP}}(\tau/G_{\text{eff}}).$$

COROLLARY 8.2.11 (Effective gain for HetNet) *The effective gain is given by*

$$G_{\text{eff}} = 1 + \sum_{k\in[K]} w_k^2 (G_k - 1), \tag{8.31}$$

where

$$w_k = \frac{\lambda_k P_k^\delta}{\sum_{i\in[K]} \lambda_i P_i^\delta}.$$

Proof Because $T_\delta(\tau/G)$ is a convex function of $G \in \mathbb{R}^+$, we have from Eq. (8.30)

$$\hat{\mathcal{P}}(\tau) \le \sum_{k\in[K]} \frac{w_k}{T_\delta\left(\frac{\tau}{w_k G_k + 1 - w_k}\right)}.$$

Next, because the w_k sum up to 1 and $1/T_\delta(\tau/G)$ is a concave function of G,

$$\begin{aligned}\hat{\mathcal{P}}(\tau) &\le \frac{1}{T_\delta\left(\frac{\tau}{\sum_{k\in[K]} w_k(w_k G_k + 1 - w_k)}\right)} \\ &= \mathcal{P}_{\text{PPP}}\left(\frac{\tau}{\sum_{k\in[K]} w_k(w_k G_k + 1 - w_k)}\right),\end{aligned}$$

and thus the effective gain is as given in Eq. (8.31). □

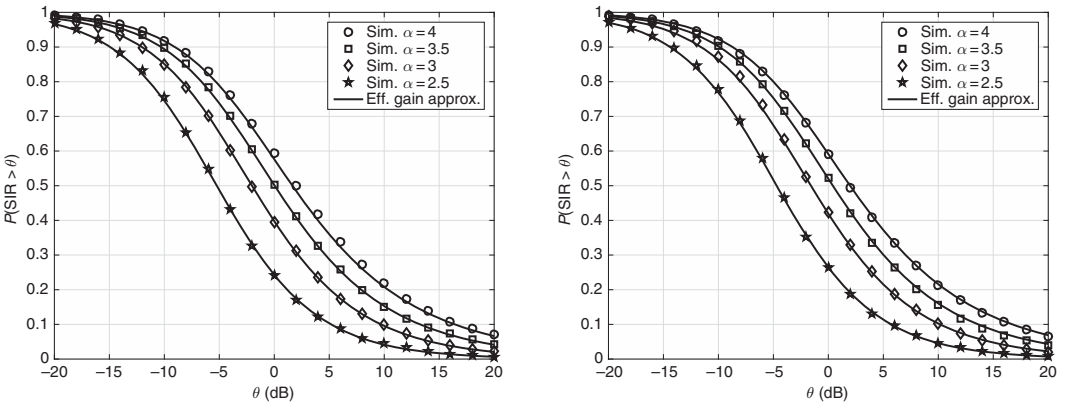

Figure 8.5 Success probabilities for three-tier HetNet models and effective gain approximation. (*Left*) The tiers are a 1-GPP with $\lambda_1 = 10^{-5}$ and $P_1 = 1$, a $1/2$-GPP with $\lambda_2 = 2\lambda_1$ and $P_2 = 1/5$ and a Poisson process with $\lambda_3 = 5\lambda_1$ and $P_3 = 1/25$. (*Right*) The tiers are a square lattice with $\lambda_1 = 10^{-5}$ and $P_1 = 1$, a 1-GPP with $\lambda_2 = 2\lambda_1$ and $P_2 = 1/5$ and a Poisson process with $\lambda_3 = 5\lambda_1$ and $P_3 = 1/25$.

Subtracting 1 from the per-tier gains G_k, the effective gain can be written compactly as a weighted sum of the per-tier gains: Let $\tilde{G}_k = G_k - 1$ and $\tilde{G}_{\mathsf{eff}} = G_{\mathsf{eff}} - 1$, we have

$$\tilde{G}_{\mathsf{eff}} = \sum_{k \in [K]} w_k^2 \tilde{G}_k.$$

Remark 8.2 The effective gain is not given simply as the weighted sum $\sum_k \tilde{G}_k \mathbb{P}(\{\text{serving tier is } k\})$. Because the superposition of many independent stationary point processes results (under mild conditions) in a uniform Poisson process, we necessarily have $\tilde{G}_{\mathsf{eff}} \to 0$ as $K \to \infty$, no matter what the tiers are.

For example, if all tiers are square lattices with the same intensity and transmit power, $\tilde{G}_k = 1$ and $w_k = 1/K$.

$$\tilde{G}_{\mathsf{eff}} = \sum_{k \in [K]} \frac{1}{K^2} = \frac{1}{K}.$$

Generally, for K identical tiers, $\tilde{G}_{\mathsf{eff}} = \tilde{G}/K$.

Because the per-tier gains are generally larger than 1, Eq. (8.29) yields a lower bound on the exact coverage probability and the effective gain approach yields an upper bound on Eq. (8.29), the two approximations together yield a pleasingly accurate simple expression for a very general network model. Fig. 8.5 shows simulations results and the effective gain approximation for two examples of three-tier HetNets. The approximation turns out to be very accurate over the entire range of τ.

8.3 Bibliographic Notes

For the first use of determinantal processes in a wireless network setting, Torrisi & Leonardi (2014) and Miyoshi & Shirai (2014a) independently proposed using Ginibre

point processes to model base station locations. Torrisi & Leonardi studied the total interference term with large deviations theory, whereas Miyoshi & Shirai derived a closed-form expression for the probability distribution of the SINR, with approximations later derived by Nagamatsu, Miyoshi, and Shirai (Nagamatsu et al. 2014). Miyoshi & Shirai (2014b) and Deng, Zhou, & Haenggi (2015a) then extended the SIR analysis to the β-Ginibre process, which was also statistically fitted by Deng, Zhou, & Haenggi (2015a) to real-world cellular network layouts. Nakata and Miyoshi considered also heterogeneous networks and Miyoshi and Shirai focused on the asymptotic properties of the distribution function of the SIR at zero and at the infinity. More generally, three different determinantal point processes (not including the Ginibre process) were fitted to cellular networks by Li et al. (2014), who then later presented coverage probability results for general determinatal point processes (Li et al. 2015). Generally speaking, the expressions involve infinite series of integrals, which need to be estimated numerically, but more tractable approximations have also been derived by Nagamatsu, Miyoshi, & Shirai (2014). Lavancier et al. (2015) showed that determinantal processes have convenient properties for statistical methods. Biscio & Lavancier (2016) proposed two methods to quantify the strength of the repulsion or regularity between the points.

To capture clustering base station locations, Suryaprakash and colleagues used a Neyman-Scott process, interpreted as a type of shot noise Cox process, to model locations of base stations, deriving expressions for the SINR distribution (Suryaprakash et al. 2015). Mathematical details on Neyman-Scott are found in Chiu et al. (2013) and Møller (2003), with the latter giving results on shot noise Cox processes in general. In the mathematics literature, Hough et al. (2009b) have studied determinantal processes, whereas Møller derived mathematical results for shot noise Cox (Møller 2003) and permanental (McCullagh & Møller 2006) processes.

The ASAPPP method originated in Guo & Haenggi (2015), where the horizontal shift is called *deployment gain*. Haenggi generalized it to other single-tier architectures (Haenggi 2014a, 2014b) and showed its accuracy over the entire SIR distribution (Ganti & Haenggi 2016a). The method was extended to multitier networks in Wei et al. (2016).

Concluding Remarks

Future wireless deployments will increasingly rely on HetNets to satisfy society's demands for high data rates coupled with high cell capacities. We argue that the design and deployment of such HetNets present many challenges, which, because of the much larger space of system deployment parameters relative to a single-tier network, cannot be feasibly addressed by the traditional approaches of measurement and simulation alone.

We show that techniques and results from stochastic geometry can relieve system designers from the burden of having to perform exhaustive simulations of all feasible combinations of deployment parameters, because they yield analytical results for key performance metrics such as coverage probability, thereby permitting (a) a uniform comparison across architectures and (b) quick rejection of certain combinations of deployment parameters without needing to simulate their performance. These theoretical results are illustrated with examples of their application to transmission scenarios specified in the LTE standard.

These analytical results for coverage probability have, until now, mostly been derived for independent Poisson deployments of base stations in tiers, a scenario for which there are tractable derivations of exact expressions. Although it is well known that real-world base station deployments do not behave like Poisson deployments, we show that the Poisson model is nonetheless fundamental to an analytical treatment of all deployments, for the following reasons: (a) the set of propagation losses to the typical location in an arbitrary network deployment converge asymptotically to that from a Poisson deployment of base stations and (b) long before this asymptotic limit is reached, the coverage results for a Poisson deployment can be employed to obtain very accurate approximations to the coverage results for various regular deployments. The present work is the first book-length treatment of these results.

We also provide the first book-level exposition of exact results on coverage for certain non-Poisson deployment models that are analytically tractable. Such models include the Ginibre point process and determinantal point processes. We also discuss the challenges of understanding and evaluating the complex analytical expressions so obtained.

For future work, in addition to stochastic geometry and point process theory covered in this book, other tools from probability may also be employed to analyze cellular networks. One such tool, for example, is the theory of large deviations, which is a

powerful but, at times, technical series of results for examining rare events. Researchers have started to use it to study the SIR in Poisson networks with a focus on studying the effects of atypical (but disruptive) network configurations.

We expect the book to be of interest to researchers in academia and industry and to anyone interested in the application of stochastic geometry to problems in communication. The investigation of HetNets via stochastic geometry is a topic of intense interest in the research community, and the number of publications increases by leaps and bounds every month. It is our hope that this book equips the reader to understand, utilize, and extend the results in the literature to better design next-generation wireless networks to fulfill the needs of humankind.

Appendix A Proof of Lemma 5.3.6

Owing to the propagation invariance expressed in Lemma 4.2.2, Remark 4.3, and the proof of Lemma 5.3.6, we assume, without loss of generality, that H are exponential random variables. Based on the definition in Eq. (5.24), we can rewrite the symmetric sums $\mathcal{S}_n$, with a conditioning argument, giving

$$\mathcal{S}_n(\tau) = \mathbb{E}\left[\sum_{\mathsf{x}_1,\ldots,\mathsf{x}_n \in \Phi}^{\neq} \mathbb{P}\{\, \mathrm{SINR}(\mathsf{x}_i) > \tau,\, i = 1,\ldots,n \mid \Phi \,\} \right], \tag{A.1}$$

where

$$\mathrm{SINR}(\mathsf{x}_i) = \frac{H_{\mathsf{x}_i}\ell(\mathsf{x}_i)}{N + I' - H_{\mathsf{x}_i}\ell(\mathsf{x}_i)}$$

and $\mathbb{P}\{\cdot \mid \Phi\}$ denotes the conditional probability given Φ with random marks $\{H_{\mathsf{x}}\}_{\mathsf{x}\in\Phi}$ for the propagation effects. We derive an expression for the right-hand side of Eq. (A.1) by first assuming exponential distribution of H with unit mean and $\kappa = 1$, and then we bring back the general assumptions appropriately by rescaling the Poisson intensity λ as explained in Example 4.17. To calculate the expectation of the random sum in $\mathcal{S}_n$, we employ the (higher-order) version of the Campbell-Mecke result (Eq. [3.32]), which, coupled with Slivnyak's theorem, gives

$$\mathcal{S}_n(\tau) = \frac{\lambda^n}{n!} \int\limits_{(\mathbb{R}^+)^n} \mathbb{P}\left(\bigcap_{i=1}^{n} \{\mathrm{SINR}(x_i) > \tau\} \right) \mathrm{d}x_1 \ldots \mathrm{d}x_n \,. \tag{A.2}$$

Symmetry and a change to polar coordinates $r_i = \|x_i\|$ (with a slight abuse of notation) lead to

$$\mathcal{S}_n(\tau) = \frac{(2\pi\lambda)^n}{n!} \int\limits_{(\mathbb{R}^+)^n} \mathbb{P}\left(\bigcap_{i=1}^{n} \{\mathrm{SINR}(r_i) > \tau\} \right) r_1 \mathrm{d}r_1 \ldots r_n \mathrm{d}r_n \,. \tag{A.3}$$

We define

$$\mathrm{STINR}(r_i) \triangleq \frac{H_i \ell(r_i)}{(N + I' + \sum_{j=1}^{n} H_i \ell(r_i))},$$

where $H_i \triangleq H_{x_i}$ and the random variable I' is the total interference or total received power

$$I' = \sum_{\mathsf{x}\in\Phi} [H_{\mathsf{x}}\ell(\mathsf{x})]\,,$$

with another slight abuse of notation, because we now write ℓ both as a function of a scalar $r \in \mathbb{R}^+$ and a point $x \in \mathbb{R}^2$. Then the relationship between SINR and STINR implies that we can write symmetric sums $\mathcal{S}_n$ as

$$\mathcal{S}_n(\tau) = \frac{(2\pi\lambda)^n}{n!} \int\limits_{(\mathbb{R}^+)^n} \mathbb{P}\left(\bigcap_{i=1}^{n}\{\mathrm{STINR}(r_i) > \tau'\}\right) r_1 \mathrm{d}r_1 \ldots r_n \mathrm{d}r_n\,, \tag{A.4}$$

where $\tau' = \tau/(1+\tau)$. The event in Expression (A.4) is equivalent to

$$\left\{ \frac{\min\left(H_1\ell(r_1), \ldots, H_n\ell(r_n)\right)}{N + I' + \sum_{j=1}^{n} H_j\ell(r_j)} > \tau' \right\}. \tag{A.5}$$

For integer $i \in [1, n]$, the exponential variable assumption on H allows us to define a new set of exponential random variables

$$\hat{E}_i \stackrel{\triangle}{=} H_i\ell(r_i),$$

with means $1/\mu_i = \ell(r_i)$, and we define

$$\hat{E}_M \stackrel{\triangle}{=} \min(\hat{E}_1, \hat{E}_2, \ldots, \hat{E}_n),$$

as another exponential variable with mean $1/\mu_M = 1/(\sum_{i=1}^{n}\mu_i)$. We introduce the random variable

$$\hat{D}_n \stackrel{\triangle}{=} \sum_{i=1}^{n} \hat{E}_i - n\hat{E}_M,$$

which has a mixed distribution. For example, with probability $\mu_1/(\sum_{i=2}^{n}\mu_i)$ the event $\{\hat{E}_M = \hat{E}_1\}$ occurs, and, so $\hat{D}_n = \sum_{i=1}^{n}\hat{E}_i - n\hat{E}_1$. The memoryless property of the exponential distribution implies that $\hat{E}_M$ and $\hat{D}_n$ are independent, and that the latter's distribution is

$$\mathbb{P}(\hat{D}_n \le d) = \frac{1}{\sum_{i=1}^{n}\mu_i}\left[\mu_1\mathbb{P}\left(\sum_{i=2}^{n}\hat{E}_i \le d\right) + \cdots + \mu_n\mathbb{P}\left(\sum_{i=1}^{n-1}\hat{E}_i \le d\right)\right]. \tag{A.6}$$

Using the reasoning behind Expression (A.5), observe that

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=1}^{n}\{\mathrm{SINR}(r_i) > \tau\}\right) &= \mathbb{P}\left(\bigcap_{i=1}^{n}\{\mathrm{STINR}(r_i) > \tau'\}\right) \\
&= \mathbb{P}\left\{\frac{\hat{E}_M}{N + I' + \hat{D}_n + n\hat{E}_M} > \tau'\right\} \\
&= \mathbb{P}\left\{\frac{\hat{E}_M}{N + I' + \hat{D}_n + (n-1)\hat{E}_M} > \bar{t}_1\right\} \\
&= \mathbb{P}\left\{\frac{\hat{E}_M}{N + I' + \hat{D}_n} > \bar{t}_n\right\}.
\end{aligned}$$

where $\bar{t}_n$ is given by Eq. (5.34) via induction, and we recall that $\bar{t}_n$ is a function of τ, but we drop its dependence from subsequent notation.

Because the random variables N, I', $\hat{D}_n$, and $\hat{E}_M$ are all independent of each other. We can take the expectation of the tail distribution of $\hat{E}_M$ to give

$$\mathbb{P}\left(\bigcap_{i=1}^{n}\{\text{SINR}(r_i) > \tau\}\right) = \mathcal{L}_N(\mu_M\bar{t}_n)\mathcal{L}_{I'}(\mu_M\bar{t}_n)\mathcal{L}_{\hat{D}_n}(\mu_M\bar{t}_n), \tag{A.7}$$

which is a product of three Laplace transforms. The first transform is by the definition

$$\mathcal{L}_N(\mu_M\bar{t}_n) = \exp\{-N\mu_M\bar{t}_n\}. \tag{A.8}$$

By Eq. (4.14), Example 4.9, and Eq. (5.30) using $\Gamma(1+\delta) = \delta\Gamma(\delta)$, the second transform is equal to

$$\mathcal{L}_{I'}(\mu_M\bar{t}_n) = \exp\left\{-\lambda(\mu_M\bar{t}_n)^{\delta}\pi C(\delta)/\kappa^2\right\}. \tag{A.9}$$

Given Expression (A.6), then the Laplace transform of a general exponential variable and the convolution theorem imply that the distribution of $\hat{D}_n$ has the transform

$$\mathcal{L}_{\hat{D}_n}(\mu_M\bar{t}_n) = \frac{\prod_{i=1}^{n}\mu_i}{\sum_{i=1}^{n}\mu_i}\left[\sum_{i=1}^{n}\frac{1}{\left(\prod_{j\neq i}[\mu_j+\mu_M\bar{t}_n]\right)}\right]. \tag{A.10}$$

We substitute the path loss function $\ell(r) = (\kappa r)^{-\alpha}$, set $\kappa = 1$, which we recover later, and some algebra gives

$$\mathbb{P}\left(\bigcap_{i=1}^{n}\{\text{SINR}(r_i) > \tau\}\right) = \frac{\left(\prod_{i=1}^{n} r_i^{\alpha}\mathrm{e}^{-N\bar{t}_n r_i^{\alpha}}\right)\mathrm{e}^{-\lambda(\bar{t}_n\sum_{i=1}^{n} r_i^{\alpha})^{\delta}\pi C(\delta)}}{\sum_{i=1}^{n} r_i^{\alpha}} \tag{A.11}$$

$$\times\sum_{i=1}^{n}\left(\frac{1}{\prod_{j\neq i}[r_j^{\alpha}+\bar{t}_n\sum_{i=k}^{n} r_k^{\alpha}]}\right). \tag{A.12}$$

But there is symmetry in the integration variables, which means for some integer $j \in [1,n]$ we can keep, say, the jth term in the above sum on the right-hand side and multiply the integral by n, giving

$$\mathcal{S}_n(\tau) = \frac{(2\pi\lambda)^n}{(n-1)!} \tag{A.13}$$

$$\times\int_0^{\infty}\cdots\int_0^{\infty}\frac{\left(\prod_{i=1}^{n} r_i^{\alpha+1}\mathrm{e}^{-N\bar{t}_n r_i^{\alpha}}\right)\mathrm{e}^{-\lambda(\bar{t}_n(\tau)\sum_{i=1}^{n} r_i^{\alpha})^{\delta}\pi C(\delta)}}{\left(\sum_{i=1}^{n} r_i^{\alpha}\right)\left(\prod_{i\neq j}[r_i^{\alpha}+\bar{t}_n\sum_{i=1}^{n} r_i^{\alpha}]\right)}\mathrm{d}r_1\ldots\mathrm{d}r_n. \tag{A.14}$$

Changing the integration variables $s_i \stackrel{\triangle}{=} r_i(\lambda \bar{t}_n^{\delta} \pi C(\delta))^{1/2}$, replacing λ by $a/(\pi\Gamma(1+\delta))$ with $a = \pi\lambda\mathbb{E}(H^{\delta})/\kappa^2$ to remove the exponential assumption and bring back the general fading distribution for H and the constant κ (as explained in Example 4.17) gives

$$\mathcal{S}_n(\tau) = \frac{2^n}{\bar{t}_n^{n\delta}(C(\delta))^n(n-1)!} \int_0^{\infty} \cdots \int_0^{\infty} \tag{A.15}$$

$$\frac{\mathrm{e}^{-(\sum_{i=1}^n s_i^{\alpha})^{\delta}} \left(\prod_{i=1}^n s_i^{\alpha+1} \mathrm{e}^{-N(a\Gamma(1-\delta))^{-1/\delta} s_i^{\alpha}}\right)}{\left(\sum_{i=1}^n s_i^{\alpha}\right)\left(\prod_{i=2}^n [H_i^{\alpha} + \bar{t}_n \sum_{k=1}^n s_k^{\alpha}]\right)} ds_1 \ldots ds_n. \tag{A.16}$$

The final steps involve a substitution of n-dimensional spherical-like variables

$$\begin{aligned}
s_1 &\stackrel{\triangle}{=} u[\sin\theta_1 \sin\theta_2 \ldots \sin\theta_{n-1}]^{\delta} \\
s_2 &\stackrel{\triangle}{=} u[\cos\theta_1 \sin\theta_2 \ldots \sin\theta_{n-1}]^{\delta} \\
s_3 &\stackrel{\triangle}{=} u[\cos\theta_2 \sin\theta_3 \ldots \sin\theta_{n-1}]^{\delta} \\
&\ldots \\
s_n &\stackrel{\triangle}{=} u[\cos\theta_{n-1}]^{\delta}.
\end{aligned}$$

Observe that $\sum_{i=1}^n s_i^{\alpha} = u^{\alpha}$ and $\prod_{i=1}^n s_i = u^n \left[\prod_{i=1}^n q_i\right]^{\delta}$, where

$$q_i = q_i(\theta_i, \ldots, \theta_{n-1}) \stackrel{\triangle}{=} (s_i/u)^{1/\delta}.$$

When $\alpha = 2$, so $\delta = 1$, our system of coordinates boils down to the regular n-dimensional spherical coordinates, whose Jacobian is $\bar{J}(u, \theta_1, \ldots, \theta_{n-1}) = u^{n-1} \prod_{i=1}^{n-1} \sin^{i-1}\theta_i$. By induction our coordinate system has the Jacobian

$$J(u, \theta_1, \ldots, \theta_{n-1}) = \delta^{n-1} \bar{J}(u, \theta_1, \ldots, \theta_n) \left[\prod_{i=1}^{n-1} \sin^i\theta_i \cos\theta_i\right]^{\delta-1},$$

which is positive over the integration domain of interest. Writing $z \stackrel{\triangle}{=} N(a\Gamma(1-\delta))^{-1/\delta}$, the integral in Expression (A.16) becomes

$$\int_0^{\infty} \int_{[0,\pi/2]^{n-1}} \frac{u^{n(\alpha+1)} \left[\prod_{i=1}^{n-1} \sin^i\theta_i \cos\theta_i\right]^{2(\alpha+1)/\alpha} \mathrm{e}^{-u^2} \mathrm{e}^{-zu^{\alpha}}}{u^{n\alpha} \prod_{i=2}^n [q_i^2 + \bar{t}_n]}$$

$$\times J(u, \theta_1, \ldots, \theta_n)\, du\, \mathrm{d}\theta_1 \ldots \mathrm{d}\theta_{n-1}$$

$$= \delta^{n-1} \int_0^\infty u^{2n-1} \mathrm{e}^{-u^2} \mathrm{e}^{-zu^\alpha} du$$

$$\times \int_{[0,\pi/2]^{n-1}} \frac{\prod_{i=1}^{n-1} \left[\sin^i \theta_i \cos \theta_i\right]^{4/\alpha+1} [\sin \theta_i]^{i-1}}{\prod_{i \neq j} [q_i^2 + \bar{t}_n]} \mathrm{d}\theta_1 \dots \mathrm{d}\theta_{n-1}.$$

The substitution $v_i = \sin^2 \theta_i$ makes the second integral over the $(n-1)$-dimensional cube equal to $2^{1-n} \bar{\mathcal{J}}_{n,\delta}(\bar{t}_n)$. To complete the proof, define η_i accordingly and take into consideration Expressions (A.16) and (5.34).

Appendix B Timeline of Cellular Technology Generations

B.1 3GPP and LTE

The scale and complexity of design, manufacture, and deployment of wireless cellular systems make it all but impossible for a single-company proprietary architecture to gain traction in the global telecommunications market. Instead, cellular network operators and equipment vendors from all over the world have joined to form standards bodies at the national and international level to facilitate the evolution of cellular communication systems. As of this writing, six national standards bodies from around the world[1] have combined to form a single international standards organization for wireless cellular communications, called the *Third Generation Partnership Project* or 3GPP. The 3GPP partnership was first created to further the universal mobile telecommunications system (UMTS) standard which defined the so-called third generation of cellular communications using code division multiple access on the air interface, but the name 3GPP was retained for the fourth generation (4G) using orthogonal frequency division multiple access. This fourth generation of cellular systems came to be known as the *long-term evolution* of the 3GPP standard and is simply abbreviated as 3GPP-LTE or just LTE.

B.2 Support for HetNets in LTE

3GPP standards are published as *releases*, with a typical interval of 12–18 months between releases. As the UMTS standard matured and the standards activity ramped up on LTE, the same release (Releases 8 and 9) contained standards specifications for both UMTS and LTE.

Release 9 provided end-to-end support for LTE small cells, but these were assumed to be femtocells deployed in homes (called *home enhanced node base stations* or HeNBs, *NodeB* being the technical term for a base station in 3GPP), with backhaul provided by the internet service provider to the subscriber's home. However, there was no widespread deployment of such an LTE HeNB network. This standard also

[1] They are the Association of Radio Industries and Businesses (ARIB) and the Telecommunication Technology Committee (TTC) from Japan, the China Communications Standards Association (CCSA) from China, the Alliance for Telecommunications Industry Solutions (ATIS) from North America, the European Telecommunications Standards Institute (ETSI) from Europe, and the Telecommunications Technology Association (Korea) (TTA) from South Korea.

ported the “high interference indicator" flag from UMTS, whereby a cell suffering high interference on some orthogonal frequency division multiplexed subcarriers could broadcast this flag so that its neighboring cells could reduce their activity (or the activity of their served users) when they detected the flag. This frequency-domain information exchange was the only form of intercell interference coordination (ICIC) supported by this standard.

Release 10 from 2011 is taken to be the first release that is exclusively for LTE, and the specification of LTE in 3GPP Release 10 is called *LTE-advanced*, or LTE-A . An excellent description of the LTE and LTE-A (Releases 8–10) standards is provided in Sesia, Toufik, & Baker (2011). The Release 10 standard includes not only several enhancements to the macrocellular LTE standard, but also support for cochannel macro-pico HetNets, where the pico deployment is also owned and provisioned by the operator of the macrocellular tier. A key feature is the introduction of almost-blank subframes together with the (near-) zero-power transmissions by the macro base stations during these almost-blank subframes to reduce interference to users (UEs, in 3GPP terminology) that have been induced to associate with pico base stations due to range expansion bias. This is called *enhanced ICIC* or eICIC to emphasize the evolution over the ICIC version in Release 9. Note that eICIC is a time-domain method.

Release 11 from 2012 (also called LTE-A) adds more capabilities to the eICIC version from Release 10, principally the support for enhanced algorithms at the receiver to minimize interference. This form of ICIC is called *further enhanced ICIC* or feICIC.

Release 12 (finalized in 2014, and continuing to be called LTE-A) expands on the features introduced in Release 11, with emphasis on MIMO and coordinated multipoint (CoMP) transmissions on the downlink, and support for direct device-to-device transmissions between user terminals in close proximity to one another, user mobility in HetNets, and HetNet self-optimization.

Release 13 (finalized in 2016) adds enhancements and improvements to Release 12, including support for machine-to-machine communications, operation of LTE on unlicensed bands, and simultaneous use of WiFi and LTE by a terminal, all to prepare for the adoption of LTE in the internet of things. The goals are to support low-power, low-capability terminals such as sensors and extend the battery life of such devices in the network.

Release 14 (standardization work ongoing) extends Release 13 to vehicle-to-vehicle communications, and adds support for public warning systems and emergency services. The eventual goal of Releases 13 and 14 is to prepare LTE for the upcoming fifth generation (5G) of wireless communication systems. Together, these two releases are referred to as LTE-A Pro.

Note that 5G is the first generation of cellular network that is not defined by a change in the physical layer access method, i.e., it retains orthogonal frequency division multiple access from 4G. Although there is at present no specification for 5G, it is intended to enhance the data rates, cell capacities, coverage, and throughput of 4G (LTE) to the user terminals deployed in 4G networks, and also extend the features and benefits of LTE to many new classes of devices, such as sensors, many of which may be low power and incapable of supporting advanced transmission schemes. The first 5G specifications are expected to be finalized in 2020.

Appendix C Some Useful Probability Distributions

In this appendix, we define the probability distributions used in the book. We do not, however, provide detailed properties of those distributions. Such details may be obtained from any of several standard references, for example, the texts by Papoulis (Papoulis 1991) or Grimmett and Stirzaker (Grimmett & Stirzaker 2001).

C.1 Discrete Distributions

C.1.1 Uniform distribution

Uniform distribution is what we mean when we talk of "randomly" selecting one object from a set of size n. If we label the objects in this set from 1 through n and let X denote the random variable representing the label of the selected object, then the selection of an object "at random" corresponds to choosing a label with the following distribution:

$$X \sim \text{Unif}\{1, \ldots, n\} \Leftrightarrow \mathbb{P}\{X = k\} = \frac{1}{n}, \quad k = 1, \ldots, n.$$

C.1.2 Bernoulli distribution

A random variable X that takes only one of the values 0 or 1 is said to have a Bernoulli distribution with parameter p:

$$\mathbb{P}\{X = 1\} = p = 1 - \mathbb{P}\{X = 0\}.$$

The Bernoulli distribution can be used to model the outcome of a single experiment (called a "trial" in probability parlance), with 1 denoting success and 0 failure, say. For example, it can model the outcome of a single toss of a biased coin, with 1 denoting heads, say, and 0, tails.

C.1.3 Binomial distribution

The sum of n iid Bernoulli random variables $X_1, \ldots, X_n$ (each with the same probability p of taking the value 1) is said to have the binomial distribution with parameters n and p:

$$X \sim \text{Bin}(n, p) \Leftrightarrow \mathbb{P}\{X = k\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.$$

It follows that the Bernoulli distribution with parameter p is equivalent to the binomial distribution with parameters 1 and p, $\text{Bin}(1,p)$. Thus $\text{Bin}(n,p)$ is the distribution of the number of successes in n independent Bernoulli trials. For the coin-tossing example, the $\text{Bin}(n,p)$ distribution models the number of heads in n independent coin tosses.

C.1.4 Poisson Distribution

The Poisson distribution with parameter λ, denoted $\text{Poiss}(\lambda)$, is the $\text{Bin}(n,\lambda/n)$ distribution where $p = \lambda/n$ for some fixed positive λ in the limit as $n \to \infty$:

$$X \sim \text{Poiss}(\lambda) \Leftrightarrow \mathbb{P}\{X = k\} = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}, \quad k = 0, 1, 2, \ldots.$$

C.1.5 Negative binomial distribution

The negative binomial distribution with parameters r and p, denoted $\text{NB}(r,p)$, is the distribution of the number of successes in a sequence of independent Bernoulli trials that is conducted as long as required to obtain exactly r failures (where the probability of failure on each trial is $1-p$):

$$X \sim \text{NB}(r,p) \Leftrightarrow \mathbb{P}\{X = k\} = \binom{r+k-1}{k}(1-p)^r p^k, \quad k = 0, 1, \ldots.$$

C.1.6 Generalized negative binomial distribution

The generalized negative binomial distribution or Pólya distribution with parameters $r > 0$ and p is just the negative binomial distribution generalized to the case where r is not an integer:

$$\mathbb{P}\{X = k\} = \frac{\Gamma(r+k)}{\Gamma(r)k!}(1-p)^r p^k, \quad k = 0, 1, \ldots.$$

C.2 Continuous Distributions

C.2.1 Uniform distribution

The continuous uniform distribution is the counterpart of the continuous discrete distribution. Specifically, the continuous distribution on the interval $[a,b]$ is one where the PDF is constant over this interval and zero everywhere else:

$$X \sim \text{Unif}[a,b] \Leftrightarrow f_X(x) = \begin{cases} 1/(b-a), & a \le x \le b, \\ 0, & \text{elsewhere.} \end{cases}$$

C.2.2 Normal or Gaussian distribution

A normal or Gaussian distribution with mean μ and variance σ^2 is one with the following PDF:

$$X \sim \mathcal{N}(\mu, \sigma^2) \Leftrightarrow f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right], \quad -\infty < x < \infty.$$

The $\mathcal{N}(0,1)$ distribution is called the *standard* or *unit* normal or Gaussian distribution.

C.2.3 Circularly symmetric complex Gaussian distribution

If X and Y are iid $\mathcal{N}(0, \sigma^2)$ for some $\sigma > 0$, then $Z = X + \mathrm{j}Y$ is said to have a circularly symmetric complex Gaussian distribution, written $\mathcal{CN}(0, 2\sigma^2)$. The $\mathcal{CN}(0,1)$ distribution is called the unit circularly symmetric complex Gaussian distribution.

C.2.4 Rayleigh distribution

The Rayleigh distribution with scale parameter $\sigma > 0$ is the distribution of the magnitude $|Z|$ of a circularly symmetric complex Gaussian random variable $Z \sim \mathcal{CN}(0, 2\sigma^2)$. In other words, the distribution of $R \equiv \sqrt{X^2 + Y^2}$, where X and Y are iid $\mathcal{N}(0, \sigma^2)$ is Rayleigh with parameter σ and has the PDF

$$R \sim \text{Rayleigh}(\sigma) \Leftrightarrow f_R(r) = \frac{r}{\sigma^2} \exp\left(-\frac{r^2}{2\sigma^2}\right), \quad r \geq 0.$$

When we talk of Rayleigh fading on a wireless link, we mean that the in-phase and quadrature components of the received signal are iid zero-mean Gaussian random variables, so that the magnitude of the equivalent complex baseband signal is Rayleigh distributed. The Rayleigh(1) distribution is called the unit Rayleigh distribution.

C.2.5 Exponential distribution

The exponential distribution with mean 1, called the unit exponential distribution and denoted Exp(1), is the distribution of the random variable $U = R^2/2$ where $R \sim$ Rayleigh(1), i.e., $U = (X^2 + Y^2)/2$, where X and Y are iid $\mathcal{N}(0,1)$:

$$U \sim \text{Exp}(1) \Leftrightarrow f_U(u) = \mathrm{e}^{-u}, \quad u \geq 0.$$

Thus, the unit exponential distribution is the distribution of the normalized power of a received complex baseband signal under unit Rayleigh fading. In other words, when the fading is Rayleigh, the magnitude of the signal is Rayleigh, and the power is exponential.

The exponential distribution with mean $\mu > 0$, denoted Exp(μ) is just the distribution of μU where $U \sim \text{Exp}(1)$:

$$V \sim \text{Exp}(\mu) \Leftrightarrow f_V(v) = \frac{1}{\mu} \exp\left(-\frac{v}{\mu}\right), \quad v \geq 0.$$

C.2.6 Erlang distribution

The Erlang distribution with shape parameter $k \in \{1, 2, \dots\}$ and scale parameter $\mu > 0$ (or equivalently rate parameter $\lambda = 1/\mu$), written Erlang(k, μ), is the distribution of the sum of k iid Exp(μ) random variables:

$$T \sim \text{Erlang}(k, \mu) \Leftrightarrow f_T(t) = \frac{t^{k-1} \exp(-t/\mu)}{\mu^k (k-1)!}, \quad t \geq 0.$$

Note that $\text{Erlang}(1, \mu) \equiv \text{Exp}(\mu)$.

C.2.7 Gamma distribution

The gamma distribution with shape parameter $k > 0$ and scale parameter $\mu > 0$, written $\text{Gamma}(k, \mu)$, is just the generalization of the $\text{Erlang}(k, \mu)$ distribution to the case where k is not an integer:

$$\Gamma \sim \text{Gamma}(k, \mu) \Leftrightarrow f_\Gamma(\gamma) = \frac{\gamma^{k-1} \exp(-\gamma/\mu)}{\mu^k \Gamma(k)}, \quad \gamma \geq 0,$$

where $\Gamma(\cdot)$ is the gamma function:

$$\Gamma(z) = \int_0^\infty t^{z-1} \, \mathrm{e}^{-t} \, \mathrm{d}t \, .$$

C.2.8 Beta distribution

If $U \sim \text{Gamma}(\alpha, \mu)$ and $V \sim \text{Gamma}(\beta, \mu)$ are independent, then $X = U/(U+V)$ is said to have the beta distribution with shape parameters α, β:

$$X \sim \text{Beta}(\alpha, \beta) \Leftrightarrow f_X(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\text{B}(\alpha, \beta)}, \quad 0 \leq x \leq 1,$$

where $\text{B}(\cdot, \cdot)$ is the beta function:

$$\text{B}(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}, \quad \alpha, \beta > 0.$$

Note that if $X \sim \text{Beta}(\alpha, \beta)$ then $1 - X \sim \text{Beta}(\beta, \alpha)$.

C.2.9 Nakagami distribution

The relation of the Nakagami distribution to the gamma distribution is the same as that of the Rayleigh distribution to the exponential distribution, viz., if the power (i.e., square of the magnitude) of the complex baseband signal has the gamma distribution, then the magnitude of the complex baseband signal is Nakagami distributed. More precisely, if $\Gamma \sim \text{Gamma}(k, \mu)$, then $\sqrt{\Gamma} \sim \text{Nakagami}(k, k\mu)$:

$$X \sim \text{Nakagami}(m, \omega) \Leftrightarrow f_X(x) = \frac{2m^m}{\Gamma(m)\omega^m} x^{2m-1} \exp\Big(-\frac{m}{\omega}x^2\Big), \quad x \geq 0, \, m > 0, \omega > 0.$$

Note that if $X \sim \text{Nakagami}(k, k\mu)$ with $k \in \{1, 2, \dots\}$ and $\mu > 0$, then $X^2 \sim \text{Erlang}(k, \mu)$. Also, $\text{Nakagami}(1, 2\sigma^2) \equiv \text{Rayleigh}(\sigma)$.

C.2.10 Log-normal distribution

If $X \sim \mathcal{N}(\mu, \sigma^2)$ for some μ and $\sigma > 0$, then $Y = \mathrm{e}^X$ is said to be log-normally distributed with parameters μ and σ^2, written $\mathcal{LN}(\mu, \sigma^2)$:

$$Y \sim \mathcal{LN}(\mu, \sigma^2) \Leftrightarrow f_Y(y) = \frac{1}{y\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\ln y - \mu)^2}{2\sigma^2}\right], \quad y > 0.$$

In the context of wireless communications, it is customary to use the exponent 10 instead of e and say that $X \sim \mathcal{LN}(\mu, \sigma^2)$ if $X_{\mathrm{dB}} \equiv 10 \log_{10} X \sim \mathcal{N}(\mu, \sigma^2)$, i.e., if $X \equiv \mathrm{e}^{X_{\mathrm{dB}}(\ln 10)/10}$, where $X_{\mathrm{dB}} \sim \mathcal{N}(\mu, \sigma^2)$. In this case, the units of μ and σ are also taken to be decibels.

C.2.11 Suzuki distribution

If $Y \sim \mathcal{LN}(\mu, \sigma^2)$ and if the conditional distribution of X given Y is Rayleigh(Y), then X is said to have the Suzuki distribution with parameters μ and σ^2, written Suzuki(μ, σ^2):

$$X \sim \text{Suzuki}(\mu, \sigma^2) \Leftrightarrow f_X(x) = \int_0^\infty \frac{x}{y^3\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left[\frac{x^2}{y^2} + \frac{(\ln y - \mu)^2}{\sigma^2}\right]\right\} \mathrm{d}y, \quad x > 0.$$

Note that X has the same distribution as HY, where $Y \sim \mathcal{LN}(\mu, \sigma^2)$ and $H \sim$ Rayleigh(1). The Suzuki distribution is used to model link loss accounting for both log-normal and Rayleigh fading on a link where the received power (in decibels) in the absence of any fading (i.e., incorporating only the effect of distance-dependent path loss) is μ, and the power (in decibels) of the locally averaged received signal (i.e., incorporating only the effect of shadow fading) is $\mathcal{N}(\mu, \sigma^2)$.

References

Abate, J. & Whitt, W. (1995), 'Numerical inversion of Laplace transforms of probability distributions', *ORSA Journal on Computing* **7**(1), 36–43.

Afshang, M., Saha, C., & Dhillon, H. S. (2017a), 'Nearest-neighbor and contact distance distributions for Thomas cluster process', *IEEE Wireless Communications Letters* **6**(1), 130–133.

Afshang, M., Saha, C., & Dhillon, H.S. (2017b), 'Nearest-neighbor and contact distance distributions for Matérn cluster process', *IEEE Communications Letters*, to appear.

Ak, S., Inaltekin, H., & Poor, H. V. (2016), 'Gaussian approximation for the downlink interference in heterogeneous cellular networks'. ArXiv, http://arxiv.org/abs/1601.06023.

Andrews, J. G. (2010), 'Understanding femtocell-overlaid cellular networks'. Keynote at the IEEE Femtocell Workshop. Available at http://users.ece.utexas.edu/~jandrews/pubs/AndrewsKeynote_Femnet2010.pdf.

Andrews, J. G. (2011), 'Can cellular networks handle 1000× the data?' Seminar at the University of Texas, Austin, Tx. Available at http://users.ece.utexas.edu/~bevans/courses/realtime/lectures/Andrews_Cellular1000x_Nov2011.pdf.

Andrews, J. G., Baccelli, F., & Ganti, R. K. (2011), 'A tractable approach to coverage and rate in cellular networks', *IEEE Transactions on Communications* **59**(11), 3122–3134.

Baccelli, F. & Błaszczyszyn, B. (2001), 'On a coverage process ranging from the Boolean model to the Poisson Voronoi tessellation, with applications to wireless communications', *Advances in Applied Probability* **33**, 293–323.

Baccelli, F. & Błaszczyszyn, B. (2009a), *Stochastic Geometry and Wireless Networks, Volume I — Theory*, Vol. 3, No. 3–4 of *Foundations and Trends in Networking*, Boston; Delft, The Netherlands: Now Publishers.

Baccelli, F. & Błaszczyszyn, B. (2009b), *Stochastic Geometry and Wireless Networks, Volume II — Applications*, Vol. 4, No. 1–2 of *Foundations and Trends in Networking*, Boston; Delft, The Netherlands: Now Publishers.

Baccelli, F., Błaszczyszyn, B., & Mühlethaler, P. (2003), A spatial reuse Aloha MAC protocol for multihop wireless mobile networks, *in* 'Proceedings of the Annual Allerton Conference on Communication, Control, and Computing'. Monticello, IL.

Baccelli, F., Błaszczyszyn, B., & Mühlethaler, P. (2006), 'An Aloha protocol for multihop mobile wireless networks', *IEEE Transactions on Information Theory* **52**(2), 421–436.

Baccelli, F., Błaszczyszyn, B., & Mühlethaler, P. (2009), 'Stochastic analysis of spatial and opportunistic Aloha', *IEEE Journal on Selected Areas in Communications, special issue on Stochastic Geometry and Random Graphs for Wireless Networks* **27**(7), 1105–1119.

Baccelli, F., Klein, M., Lebourges, M., & Zuyev, S. (1997), 'Stochastic geometry and architecture of communication networks', *Telecommunication Systems* **7**(1), 209–227.

Baccelli, F. & Zhang, X. (2015), A correlated shadowing model for urban wireless networks, *in* 'Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM)', IEEE, pp. 801–809.

Banani, S. A., Eckford, A. W., & Adve, R. S. (2014), The penalty for random deployment in hexagonal lattice networks with perturbed interferers, *in* 'Heterogeneous and Small Cell Workshop, IEEE Globecom Workshops, 2014', pp. 1272–1277.

Behnad, A., Wang, X., & Akhtar, A. M. (2015), 'Communication neighbors comparison in a Poisson field of nodes', *IEEE Communications Letters* **19**(11), 2025–2028.

Berman, A. & Plemmons, R. J. (1994), *Nonnegative Matrices in the Mathematical Sciences*, Philadelphia: Society for Industrial and Applied Mathematics.

Biscio, C. A. N. & Lavancier, F. (2016), 'Quantifying repulsiveness of determinantal point processes', *Bernoulli* **22**(4), 2001–2028.

Błaszczyszyn, B. & Karray, M. K. (2012), Linear-regression estimation of the propagation-loss parameters using mobiles' measurements in wireless cellular networks, *in* '2012 10th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt)', Paderborn, Germany, May 14–18, 2012, pp. 54–59.

Błaszczyszyn, B. & Karray, M. K. (2016), 'Spatial distribution of the SINR in Poisson cellular networks with sector antennas', *IEEE Transactions on Wireless Communications* **15**(1), 581–593.

Błaszczyszyn, B. & Keeler, H. P. (2013), Equivalence and comparison of heterogeneous cellular networks, *in 2013 IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC Workshops)*, London, September 08, 2013, pp. 153–157.

Błaszczyszyn, B. & Keeler, H. P. (2014), 'SINR in wireless networks and the two-parameter Poisson-Dirichlet process', *IEEE Wireless Communications Letters* **3**(5), 525–528.

Błaszczyszyn, B. & Keeler, H. P. (2015), 'Studying the SINR process of the typical user in Poisson networks by using its factorial moment measures', *IEEE Transactions on Information Theory* **61**(12), 6774–6794.

Blaszczyszyn, B. & Mühlethaler, P. (2015), 'Interference and SINR coverage in spatial non-slotted Aloha networks', *Annales des Telecommunications–Annals of Telecommunications* **70**(7), 345–358.

Błaszczyszyn, B., Jovanović, M., & Karray, M. K. (2014), How user throughput depends on the traffic demand in large cellular networks, *in 2014 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Hammamet, Tunisia, May 12–16, 2014, pp. 611–619.

Błaszczyszyn, B., Karray, M. K., & Keeler, H. P. (2013), Using Poisson processes to model lattice cellular networks, *in* 'Proceedings of IEEE INFOCOM 2013', pp. 773–781.

Błaszczyszyn, B., Karray, M. K., & Keeler, H. P. (2015), 'Wireless networks appear Poissonian due to strong shadowing', *IEEE Transactions on Wireless Communications* **14**(8), 4379–4390.

Błaszczyszyn, B., Karray, M. K., & Klepper, F.-X. (2010), Impact of the geometry, path-loss exponent and random shadowing on the mean interference factor in wireless cellular networks, *in Proceedings of the Third Joint IFIP Wireless and Mobile Networking Conference (WMNC) 2010*, pp. 1–6.

Błaszczyszyn, B., Keeler, H. P., & Mühlethaler, P. (2017), Optimizing spatial throughput in device-to-device networks, *in 2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Paris, May 15–19, 2017, pp. 1–6.

Brown, T. X. (2000), 'Cellular performance bounds via shotgun cellular systems', *IEEE Journal on Selected Areas in Communications*, **18**(11), 2443–2455.

Chiu, S. N., Stoyan, D., Kendall, W. S., & Mecke, J. (2013), *Stochastic Geometry and Its Applications*, 3rd ed. Somerset, NJ: John Wiley & Sons.

Cover, T. M. & Thomas, J. A. (1991), *Elements of Information Theory*, Somerset, NJ: John Wiley & Sons.

Daley, D. J. & Vere-Jones, D. (2008), *An Introduction to the Theory of Point Processes*, Vols I and II, 2nd ed. New York: Springer.

Deng, N. & Haenggi, M. (2017), 'A fine-grained analysis of millimeter-wave device-to-device networks', *IEEE Transactions on Communications* **65**(11), 4940–4954.

Deng, N., Zhou, W., & Haenggi, M. (2015a), 'The Ginibre point process as a model for wireless networks with repulsion', *IEEE Transactions on Wireless Communications* **14**(1), 107–121.

Deng, N., Zhou, W., & Haenggi, M. (2015b), 'Heterogeneous cellular network models with dependence', *IEEE Journal on Selected Areas in Communications* **33**(10), 2167–2181.

Dhillon, H., Ganti, R., & Andrews, J. G. (2011), A tractable framework for coverage and outage in heterogeneous cellular networks, *in* 'Information Theory and Applications Workshop (ITA), 2011', IEEE, pp. 1–6.

Dhillon, H., Ganti, R., Baccelli, F., & Andrews, J. (2012), 'Modeling and analysis of K-tier downlink heterogeneous cellular networks', *IEEE Journal on Selected Areas in Communications* **30**(3), 550–560.

Dhillon, H., Kountouris, M., & Andrews, J. G. (2013), 'Downlink MIMO HetNets: Modeling, ordering results and performance analysis', *IEEE Transactions on Wireless Communications* **12**(10), 5208–5222.

Dick, J., Kuo, F. Y., & Sloan, I. H. (2013), 'High-dimensional integration: the quasi-Monte Carlo way', *Acta Numerica* **22**, 133–288.

ElSawy, H., Hossain, E., & Haenggi, M. (2013), 'Stochastic geometry for modeling, analysis, and design of multi-tier and cognitive cellular wireless networks: A survey', *IEEE Communications Surveys & Tutorials* **15**(3), 996–1019.

ElSawy, H., Sultan-Salem, A., Alouini, M. S., & Win, M. Z. (2017), 'Modeling and analysis of cellular networks using stochastic geometry: A tutorial', *IEEE Communications Surveys & Tutorials* **19**(1), 167–203.

Feller, W. (1968), *An Introduction to Probability Theory and Its Applications*, Vol. 1, 3rd ed. New York: J. Wiley & Sons.

Feller, W. (1970), *An Introduction to Probability Theory and Its Applications*, Vol 2, 2nd ed. Somerset, NJ: John Wiley & Sons.

Ganti, R. K. & Haenggi, M. (2016a), 'Asymptotics and approximation of the SIR distribution in general cellular networks', *IEEE Transactions on Wireless Communications* **15**(3), 2130–2143.

Ganti, R. K. & Haenggi, M. (2016b), SIR asymptotics in Poisson cellular networks without fading and with partial fading, *in* 'IEEE International Conference on Communications (ICC16)', Kuala Lumpur, Malaysia.

Gentner, D. & Last, G. (2011), 'Palm pairs and the general mass-transport principle', *Mathematische Zeitschrift* **267**(3–4), 695–716.

George, G., Mungara, R. K., Lozano, A., & Haenggi, M. (2017) 'Ergodic Spectral Efficiency in MIMO Cellular Networks', *IEEE Transactions on Wireless Communications* **16**(5), 2835–2849.

Gil-Pelaez, J. (1951), 'Note on the Inversion Theorem', *Biometrika* **38**(3/4), 481–482.

Gradshteyn, I. S. & Ryzhik, I. M. (2007), *Table of Integrals, Series, and Products*, 7th ed. Boston: Academic Press.

Grimmett, G. & Stirzaker, D. (2001), *Probability and Random Processes*, Oxford, UK: Oxford University Press.

Guo, A. & Haenggi, M. (2013), 'Spatial stochastic models and metrics for the structure of base stations in cellular networks', *IEEE Transactions on Wireless Communications* **12**(11), 5800–5812.

Guo, A. & Haenggi, M. (2015), 'Asymptotic deployment gain: A simple approach to characterize the SINR distribution in general cellular networks', *IEEE Transactions on Communications* **63**(3), 962–976.

Haenggi, M. (2008), 'A geometric interpretation of fading in wireless networks: Theory and applications', *IEEE Transactions on Information Theory* **54**(12), 5500–5510.

Haenggi, M. (2012), *Stochastic Geometry for Wireless Networks*, Cambridge, UK: Cambridge University Press.

Haenggi, M. (2014a), ASAPPP: A simple approximative analysis framework for heterogeneous cellular networks. Keynote presentation at the 2014 Workshop on Heterogeneous and Small Cell Networks (HetSNets'14). Available at www.nd.edu/~mhaenggi/talks/hetsnets14.pdf.

Haenggi, M. (2014b), 'The mean interference-to-signal ratio and its key role in cellular and amorphous networks', *IEEE Wireless Communications Letters* **3**(6), 597–600.

Haenggi, M. (2016), 'The meta distribution of the SIR in Poisson Bipolar and cellular networks', *IEEE Transactions on Wireless Communications* **15**(4), 2577–2589.

Haenggi, M. (2017), 'User point processes in cellular networks', *IEEE Wireless Communications Letters* **6**(2), 258–261.

Haenggi, M. & Ganti, R. K. (2008), 'Interference in Large Wireless Networks', *Foundations and Trends in Networking* **3**(2), 127–248. Available at www.nd.edu/~mhaenggi/pubs/now.pdf.

Haenggi, M., Andrews, J. G., Baccelli, F., Dousse, O., & Franceschetti, M. (2009), 'Stochastic geometry and random graphs for the analysis and design of wireless networks', *IEEE Journal on Selected Areas in Communications* **27**(7), 1029–1046.

Handa, K. (2009), 'The two-parameter Poisson-Dirichlet point process', *Bernoulli* **15**(4), 1082–1116.

Hirsch, C., Jahnel, B., Keeler, P., & Patterson, R. I. (2016), 'Large deviation principles for connectable receivers in wireless networks'. *Advances in Applied Probability*, **48**(4), 1061–1094.

Hirsch, C., Jahnel, B., Keeler, P., & Patterson, R. I. (2017), 'Large deviations in relay-augmented wireless networks'. *Queueing Systems* (October), 1–39. https://doi.org/10.1007/s11134-017-9555-9

Hough, J. B., Krishnapur, M., Peres, Y., & Virág, B. (2009), *Zeros of Gaussian Analytic Functions and Determinantal Point Processes*, University Lecture Series 51, Providence, RI: American Mathematical Society.

Huang, H., Papadias, C. B., & Venkatesan, S. (2012), *MIMO Communication for Cellular Networks*. New York: Springer.

Jo, H.-S., Sang, Y. J., Xia, P., & Andrews, J. G. (2011), Outage probability for heterogeneous cellular networks with biased cell association, *in* 'Proceedings of 2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)', IEEE, pp. 1–5.

Kalamkar, S. S. & Haenggi, M. (2017), Spatial outage capacity of Poisson Bipolar networks, *in* 'IEEE International Conference on Communications (ICC'17)', Paris, France.

Karray, M. K. (2010), Spectral and energy efficiencies of OFDMA wireless cellular networks, *in* 'Wireless Days (WD), 2010 IFIP', pp. 1–5.

Keeler, H. P. (2014), 'Studying the SINR process in Poisson networks by using its factorial moment measures', MATLAB Central File Exchange. Available at www.mathworks.com.au/matlabcentral/fileexchange/45299-studying-the-sinr-process-in-poisson-networks-by-using-its-factorial-moment-measures.

Keeler, H. P. & Błaszczyszyn, B. (2014), 'SINR in wireless networks and the two-parameter Poisson-Dirichlet process', *IEEE Wireless Communications Letters* **3**(5), 525–528.

Keeler, H. P., Błaszczyszyn, B., & Karray, M. K. (2013), SINR-based k-coverage probability in cellular networks with arbitrary shadowing, in '2013 IEEE International Symposium on Information Theory', Istanbul, July 7–12, 2013, pp. 1167–1171.

Keeler, H. P., Ross, N., & Xia, A. (2016a), 'When do wireless network signals appear Poisson?'

Keeler, H. P., Ross, N., Xia, A., & Błaszczyszyn, B. (2016b), 'Stronger wireless signals appear more Poisson', *IEEE Wireless Communications Letters* **5**(6), 572–575.

Kingman, J. F. C. (1993), *Poisson Processes*, Oxford, UK: Oxford University Press.

Kleinrock, L. (1996), 'Nomadicity: Anytime, anywhere in a disconnected world', *Mobile Networks and Applications* **1**, 351–357.

Kuo, F. & Sloan, I. (2005), 'Lifting the curse of dimensionality', *Notices of the AMS* **52**(11), 1320–1328.

Last, G. & Penrose, M. (2017), *Lectures on the Poisson Process*, Cambridge, UK: Cambridge University Press.

Lavancier, F., Møller, J., & Rubak, E. (2015), 'Determinantal point process models and statistical inference', *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **77**(4), 853–877.

Lee, C.-H. & Haenggi, M. (2012), 'Interference and outage in Poisson cognitive networks', *IEEE Transactions on Wireless Communications* **11**(4), 1392–1401.

Lee, C.-H., Shih, C.-Y., & Chen, Y.-S. (2013), 'Stochastic geometry-based models for modeling cellular networks in urban areas', *Wireless Networks* **19**(6), 1063–1072.

Lee, J., Zhang, X., & Baccelli, F. (2016), Shadowing and coverage in Poisson buildings, *in* 'IEEE INFOCOM 2016-The 35th Annual IEEE International Conference on Computer Communications', pp. 1–9.

Li, Y., Baccelli, F., Dhillon, H. S., & Andrews, J. G. (2014), Fitting determinantal point processes to macro base station deployments, *in* 'IEEE Global Communications Conference (GLOBECOM), 2014', pp. 3641–3646.

Li, Y., Baccelli, F., Dhillon, H. S., & Andrews, J. G. (2015), 'Statistical modeling and probabilistic analysis of cellular networks with determinantal point processes', *IEEE Transactions on Communications* **63**(9), 3405–3422.

Li, C., Zhang, J., & Letaief, K. (2014), 'Throughput and energy efficiency analysis of small cell networks with multi-antenna base stations', *IEEE Transactions on Wireless Communications* **13**(5), 2505–2517.

Lu, W. & Di Renzo, M. (2015), Stochastic geometry modeling of cellular networks: Analysis, simulation and experimental validation, *in* 'Proceedings of the 18th ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems', pp. 179–188.

MacDonald, V. H. (1979), 'Advanced mobile phone system: The cellular concept', *Bell System Technical Journal* **58**(1), 15–41.

Madhusudhanan, P., Restrepo, J. G., Liu, Y., Brown, T. X., & Baker, K. R. (2011), Multi-tier network performance analysis using a shotgun cellular system, *in* 'Global Telecommunications Conference (GLOBECOM 2011)', pp. 1–6.

Madhusudhanan, P., Restrepo, J. G., Liu, Y., Brown, T. X., & Baker, K. R. (2014), 'Downlink performance analysis for a generalized shotgun cellular system', *IEEE Transactions on Wireless Communications* **13**(12), 6684–6696.

Matérn, B. (1986), *Spatial Variation*, Springer Lecture Notes in Statistics. 2nd ed. New York: Springer-Verlag.

McCullagh, P. & Møller, J. (2006), 'The permanental process', *Advances in Applied Probability* **38**(4), 873–888.

Miyoshi, N. & Shirai, T. (2014a), 'A cellular network model with Ginibre configured base stations', *Advances in Applied Probability* **46**(3), 832–845.

Miyoshi, N. & Shirai, T. (2014b), Cellular networks with α-Ginibre configurated base stations, *in* M. Wakayama et al., eds. 'The Impact of Applications on Mathematics', New York: Springer, pp. 211–226.

Miyoshi, N. & Shirai, T. (2017), 'Tail asymptotics of signal-to-interference ratio distribution in spatial cellular network models', arXiv: https://arxiv.org/abs/1703.05024.

Møller, J. (2003), 'Shot noise Cox processes', *Advances in Applied Probability* **35**(3), 614–640.

Mukherjee, S. (2011), Downlink SINR distribution in a heterogeneous cellular wireless network with max-SINR connectivity, *in* '49th Annual Allerton Conference on Communication, Control, and Computing (Monticello, IL), 2011', pp. 1649–1656.

Mukherjee, S. (2012), 'Distribution of downlink SINR in heterogeneous cellular networks', *IEEE Journal on Selected Areas in Communications* **30**(3), 575–585.

Mukherjee, S. (2014), *Analytical Modeling of Heterogeneous Cellular Networks – Geometry, Coverage, and Capacity*, Cambridge: Cambridge University Press.

Mungara, R. K., Morales-Jimenez, D., & Lozano, A. (2015), 'System-level performance of interference alignment', *IEEE Transactions on Wireless Communications* **14**(2), 1060–1070.

Nagamatsu, H., Miyoshi, N., & Shirai, T. (2014), Padé approximation for coverage probability in cellular networks, *in* 'WiOpt 2014 : The 12th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks', pp. 693–700.

Nakata, I. & Miyoshi, N. (2014), 'Spatial stochastic models for analysis of heterogeneous cellular networks with repulsively deployed base stations', *Performance Evaluation* **78**, 7–17.

National Institute of Standards and Technology. Digital Library of Mathematical Functions (2012). Release 1.0.5 of 2012-10-01. Available at http://dlmf.nist.gov/.

Nigam, G., Minero, P., & Haenggi, M. (2014), 'Coordinated multipoint joint transmission in heterogeneous networks', *IEEE Transactions on Communications* **62**(11), 4134–4146.

Nigam, G., Minero, P., & Haenggi, M. (2015), 'Spatiotemporal cooperation in heterogeneous cellular networks', *IEEE Journal on Selected Areas in Communications* **33**(6), 1253–1265.

Novlan, T., Ng, B. L., Zhang, J. C., Chen, H., & Liu, L. (2015), 'Application of stochastic geometry in modeling future LTE-A and 5G wireless networks'. Talk at the Simons Conference on Networks and Stochastic Geometry, University of Texas, Austin, TX. Available at www.ma.utexas.edu/conferences/simons2015/Modeling3GPP_Samsung_Final.pdf.

Panchenko, D. (2013), *The Sherrington-Kirkpatrick Model*. New York: Springer.

Papoulis, A. (1991), *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York: McGraw-Hill.

Pitman, J. & Yor, M. (1997), 'The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator', *The Annals of Probability* **25**(2), 855–900.

Rázc, S., Tari, A., & Telek, M. (2006), 'A moments based distribution bounding method', *Mathematical and Computer Modelling* **43**(11), 1367–1382.

Rényi, A. (1967), 'Remarks on the Poisson Process', *Studia Scientiarum Mathematicarum Hungarica* **2**, 119–123.

Renzo, M. D. & Guan, P. (2014), 'A mathematical framework to the computation of the error probability of downlink MIMO cellular networks by using stochastic geometry', *IEEE Transactions on Communications* **62**(8), 2860–2879.

Riihijärvi, J. & Mähönen, P. (2012), 'A spatial statistics approach to characterizing and modeling the structure of cognitive wireless networks', *Ad Hoc Networks* **10**(5), 858–869.

Ripley, B. D. (1976), 'The second-order analysis of stationary point processes', *Journal of Applied Probability* **13**, 255–266.

Ross, N. & Schuhmacher, D. (2017), 'Wireless network signals with moderately correlated shadowing still appear Poisson', *IEEE Transactions on Information Theory* **63**(2), 1177–1198.

Salehi, M., Mohammadi, A., & Haenggi, M. (2017), 'Analysis of D2D underlaid cellular networks: SIR meta distribution and mean local delay', *IEEE Transactions on Communications* **65**(7), 2904–2916.

Schilcher, U., Toumpis, S., Haenggi, M., Crismani, A., Brandner, G., & Bettstetter, C. (2016), 'Interference functionals in Poisson networks', *IEEE Transactions on Information Theory* **62**(1), 370–383.

Sesia, S., Toufik, I., & Baker, M., eds (2011), *LTE, The UMTS Long Term Evolution: From Theory to Practice*, 2nd ed. Hoboken, NJ: Wiley.

Shirai, T. & Takahashi, Y. (2003), 'Random point fields associated with certain Fredholm determinants I: Fermion, Poisson and Boson point processes', *Journal of Functional Analysis* **205**(2), 414–463.

Suryaprakash, V., Møller, J., & Fettweis, G. (2015), 'On the modeling and analysis of heterogeneous radio access networks using a Poisson cluster process', *IEEE Transactions on Wireless Communications* **14**(2), 1035–1047.

Tanbourgi, R., Dhillon, H. S., & Jondral, F. K. (2015), 'Analysis of joint transmit–receive diversity in downlink MIMO heterogeneous cellular networks', *IEEE Transactions on Wireless Communications* **14**(12), 6695–6709.

Temme, N. M. (2003), 'Large parameter cases of the Gauss hypergeometric function', *Computational and Applied Mathematics* **153**, 441–462.

Torrisi, G. L. & Leonardi, E. (2014), 'Large deviations of the interference in the Ginibre network model', *Stochastic Systems* **4**(1), 173–205.

Vu, T.-T., Decreusefond, L., & Martins, P. (2014), 'An analytical model for evaluating outage and handover probability of cellular wireless networks', *Wireless Personal Communications* **74**(4), 1117–1127.

Wang, Y., Cui, Q., Haenggi, M., & Tan, Z. (2017), 'On the SIR meta distribution for Poisson networks with interference cancellation', *IEEE Wireless Communications Letters*. to appear.

Wang, Y., Haenggi, M., & Tan, Z. (2017), The meta distribution of the SIR for cellular networks with power control. ArXiv, http://arxiv.org/abs/1702.01864v1.

Wei, H., Deng, N., Zhou, W., & Haenggi, M. (2016), 'Approximate SIR analysis in general heterogeneous cellular networks', *IEEE Transactions on Communications* **64**(3), 1259–1273.

Yazdanshenasan, Z., Dhillon, H. S., Afshang, M., & Chong, P. H. J. (2016), 'Poisson hole process: Theory and applications to wireless networks', *IEEE Transactions on Wireless Communications* **15**(11), 7531–7546

Yu, B., Yang, L., Ishii, H., & Mukherjee, S. (2015), 'Dynamic TDD support in macrocell-assisted small cell architecture', *IEEE Journal on Selected Areas in Communications* **33**(6), 1201–1213.

Zhang, X. & Andrews, J. G. (2015), 'Downlink cellular network analysis with multi-slope path loss models', *IEEE Transactions on Communications* **63**(5), 1881–1894.

Zhang, X. & Haenggi, M. (2013), On decoding the *k*th strongest user in Poisson networks with arbitrary fading distribution, *in* '47th Asilomar Conference of Signals, Systems and Computers (Asilomar'13)', Pacific Grove, CA.

Zhang, X. & Haenggi, M. (2014), 'A stochastic geometry analysis of inter-cell interference coordination and intra-cell diversity', *IEEE Transactions on Wireless Communications* **13**(12), 6655–6669.

Zhang, X., Baccelli, F., & Heath, R. W. (2015), An indoor correlated shadowing model, *in* '2015 IEEE Global Communications Conference (GLOBECOM)', pp. 1–7.

Zorzi, M. & Pupolin, S. (1995), 'Optimum transmission ranges in multihop packet radio networks in the presence of fading', *IEEE Transactions on Communications* **43**(7), 2201–2205.

Index