PROBABILISTIC METHODS IN TELECOMMUNICATION

Benedikt Jahnel
WIAS Berlin

and

Wolfgang König
TU Berlin and WIAS Berlin

Lecture Notes

7 June, 2018

Contents

Preface

These are the lecture notes for a two-hours lecture, held by us in the summer semester 2018 at Technische Universität Berlin. We introduce mathematical concepts and theories for the rigorous probabilistic analysis of a certain type of a telecommunication system, an *multi-hop ad hoc system*. Such a system consists of users, base stations and message trajectories, jumping in several hops from user to user and/or base station. The term *ad hoc* expresses the fact that the messages do not have to jump to a base station directly and from there to their targets, but they use all the users present in the system for a sequence of hops to reach the target. Such a system has some advantages over systems that require direct hops to base stations: it does not need to have many (expensive!) base stations and can potentially absorb and handle more information. On the other hand, new questions arise, since the operator does not have any control on the locations of the users, which are decisive for the success of the message transmission: does every transmitter have always a connection, i.e., is another user available close by? Can this be iterated until the target has been reached? How long can the message trajectories be chosen, i.e., over how long distances can messages be transmitted in such a system? Further decisive questions are about problems coming from interference, i.e., the situation that too many messages are sent out in a part of the space at the same time, such that they hamper each other's successful delivery.

The main source of randomness sits in the locations of the users (in some models also the locations of the base stations), or in the randomness of the message trajectories. Let us begin with the user locations. It is generally acknowledged that the most fruitful mathematical model is the *Poisson point process (PPP)*, a random point cloud model in the d-dimensional euclidean space without clumping of the points and with a great degree of independence. This model is very elementary, but gives rise to high-level mathematics, and it can be extended in various directions by adding many types of features, in order to obtain more realistic models for certain situations. In Chapter 1, we introduce the PPP and prepare tools for a mathematical analysis. We also go a bit beyond the basics of the theory of PPPs by discussing a more refined setting, the *Cox process*, which is amenable to a more realistic modeling of the telecommunication area by adding for example a random street system.

The first big circle of questions concerns *connectivity*, i.e., the question how far messages can be transmitted via a multi-hop system from Poisson point to Poisson point, where every step has only a bounded length. In mathematical terms, this is the most fundamental question of *continuum percolation*, the question whether or not the PPP admits infinitely long multi-hop trajectories or not. The main mathematical model here is the *Boolean model*, which puts random closed areas around each of the Poisson points and asks about the size of the connected components of the union of all these areas, in particular for the existence of an unbounded component. We will introduce the concept and the most important results in Chapter 2. The

material of Chapters 1 and 2 summarizes the basics of a probabilistic theory that is sometimes called *stochastic geometry* and is the objective of many specialized lectures on probability because of its universal value for many models.

In Chapter 3 we now turn to a subject that is quite special for telecommunication, the *interference*. We introduce one of the main approaches (if not *the* main one) to express this phenomenon in a mathematically rigorous way: by means of the *signal-to-interference ratio*. We define the direct transmission of a hop as successful if the signal strength of that hop is larger than a given technical constant times the sum of all the signal strengths of all the hops that are present in the system. This induces a quite complex interaction in the PPP and destroys a great part of the independence, and it creates a new and more realistic connectivity structure. We introduce this concept in Chapter 3, demonstrate some model calculations and give some results about the percolation behavior of the resulting random graphs.

We are mostly interested in *large* systems and their overall behavior in summarizing terms. Hence, a great part of our mathematical analysis is devoted to asymptotic theories, like ergodic theorems (extensions of laws of large numbers) and *large-deviations analysis*. While the former is widely known and does not need to be introduced here, the latter may be less known in applications to communication systems. Roughly speaking, this theory provides the basis for analyzing the probabilities of events with an extremely small probability as a certain parameter diverges, and also the events themselves. This appears a very natural task for telecommunication systems, as many events (sometimes called *frustration events*) of a very low service quality need to be controlled and understood. In Chapter 4 we introduce the basics of the theory and apply it to an important setting relevant for telecommunication systems.

The theory developed so far is mainly static and thus can be interpreted as to represent a snap-shot view on an ad hoc system. In the final Chapter 5 we go beyond this setting by introducing a class of evolutionary processes on the network. More precisely, we give a short introduction to the theory of *interacting particle systems*, which are Markov jump processes in continuous time. Initially studied in the field of statistical physics, the framework of interacting particle systems has been subsequently used to model a great variety of situations such as opinion formation, spread of infections or traffic behavior. In our setting of an ad hoc telecommunications system, interacting particle systems can for example be used to analyze the *random spread of malware* and possible counter measures.

Certainly, there are many important questions that we do not touch in these notes, like coding questions, movement of the users, the introduction of time in the optimization of message routing, and much more.

In order to be able to present a useful wealth of material, we are not giving all the details of the proofs, but restrict at many places to explaining the main idea and strategy. The prerequisites that we will be relying on are the contents of two standard lectures on *Probability 1 and 2*, notably familiarity with calculations involving Poisson random variables, general measure theory, weak convergence of random variables, conditional expectations, stochastic processes in discrete and continuous time and the like.

We would like to thank our co-authors and co-workers Christian Hirsch, András Tóbiás and Alexander Wapenhans for their contributions to the text and illustrations. This work was supported by the WIAS Leibniz Group *Probabilistic Methods for Mobile Ad-Hoc Networks*.

Benedikt Jahnel and Wolfgang König

Berlin, in June 2018

Chapter 1

Device locations: Poisson point processes

In this chapter, we introduce the basic mathematical model for the random locations of many point-like objects in the Euclidean space, the *Poisson point process (PPP)*. This process will be used for modeling the places of users (i.e., their devices), additional boxes (supporting devices) and/or base stations in space. Apart from this interpretation in telecommunication, the PPP is universally applicable in many situations and is fundamental for the theory of stochastic geometry. The main assumption is a high degree of statistical independence of all the random points, which leads to many explicit and tractable formulas and to the validity of many properties that make a mathematical treatment simple. For these reasons, the PPP is the initial method of the choice practically in any spatial telecommunication modeling, and the most obvious starting point for a mathematical analysis. We will frequently refer to this application.

Poisson processes belong to the core subjects of probability theory, and there is a number of general mathematical introductory and deepening texts on this subject, e.g., [K95] or [LP17], as well as texts with emphasis towards applications in telecommunication and a chapter on Poisson processes, like [FM08, H12, P03]. Much more technical and comprehensive texts about general point processes are [DVJ03] and [Re87].

1.1 Point processes

In this section, we introduce random point clouds as random variables and discuss briefly some basics on topology and measurability. See [DVJ03, Appendix A2] and [Re87] for details and proofs.

To begin with, we fix a dimension $d \in \mathbb{N}$ and a measurable set $D \subset \mathbb{R}^d$, which in our interpretation is the *communication area* where measurability on $\mathbb{R}^d$ is considered with respect to the Borel-sigma algebra $\mathcal{B}(\mathbb{R}^d)$. In D, we assume that a random point cloud $\mathbb{X} = (X_i)_{i\in I}$, with a random index set I, is given. This is interpreted as the cloud of the *locations of the devices* (users, supporting devices, base stations etc.). We would like to have that, with probability one, these locations do not coincide or accumulate anywhere in D, i.e., that $X_i \neq X_j$ for any $i \neq j$, and that any compact subset of D receives only finitely many of the X_i. Hence, the index set I is at most countable. Actually, we do not want to distinguish the points, but indeed look only at the *set* $\mathbb{X} = \{X_i \colon i \in I\}$ or, equivalently, at the *point measure* $\sum_{i\in I} \delta_{X_i}$. In other words, we

would like to have a random variable $\mathbb{X}$ with values in the set

$$\mathbb{S}(D) = \{x \subset D \colon \#(x \cap A) < \infty \text{ for any bounded } A \subset D\}. \tag{1.1.1}$$

The elements of $\mathbb{S}(D)$ are called *locally finite sets*, and their point measures are *Radon measures*, i.e., measures that assign a finite value to any compact set. We call such an element a *point cloud in* D, and we will often make no difference between the set $x = \{x_i \colon i \in I\}$ and its point measure $\sum_{i \in I} \delta_{x_i}$.

Prospectively, we want to describe the distribution of a *random point cloud* $\mathbb{X} = (X_i)_{i \in I}$ in D, which we call a *point process*, i.e., a random variable with values in $\mathbb{S}(D)$. For this, we need a measurable structure on the state space $\mathbb{S}(D)$. We will now introduce two natural ones. Both have the advantage that they come from some topology, i.e., both ones are Borel-σ-algebras. Hence, it will be enough to introduce topologies. If we consider $\mathbb{S}(D)$ as a set of point measures, then a common way to describe a topology is to test elements of $\mathbb{S}(D)$ against a suitable class of functions. More precisely, we consider functionals of the form

$$S_f(x) = \Big\langle f, \sum_{i \in I} \delta_{x_i} \Big\rangle = \int f(y) \sum_{i \in I} \delta_{x_i}(\mathrm{d}y) = \sum_{i \in I} \int f(y)\, \delta_{x_i}(\mathrm{d}y) = \sum_{i \in I} f(x_i), \tag{1.1.2}$$

with $f \colon D \to \mathbb{R}$ taken from a suitable class of functions. Here we wrote $\langle f, \nu \rangle$ for the integral of f with respect to a measure ν. This approach is an adaptation of the well-known characterization of the weak topology on the set of (probability) measures to the current setting of point measures. Note that point measures are in general not normalized and in fact often have infinite total mass, which would render $S_f(x)$ equal to ∞ for many functions f. It makes more sense to test the point cloud only in local areas, and this is what we want to do now. The two sets of test functions are

$$\mathcal{C}_{\mathrm{c}}(D) = \text{the set of continuous functions } D \to \mathbb{R} \text{ with compact support}$$

and

$$\mathcal{M}(D) = \text{the set of measurable functions } D \to \mathbb{R} \text{ with compact support.}$$

Note that, in the definition of $\mathcal{C}_{\mathrm{c}}(D)$ and $\mathcal{M}(D)$, instead of a compact support, we could equivalently also talk of a bounded support.

Definition 1.1.1 (Vague and τ-topology on $\mathbb{S}(D)$)**.** *1. The* vague topology *is the smallest topology on $\mathbb{S}(D)$ such that, for any function $f \in \mathcal{C}_{\mathrm{c}}(D)$, the map $x \mapsto S_f(x)$is continuous.*

2. The τ-topology *is the smallest topology on $\mathbb{S}(D)$ such that, for any function $f \in \mathcal{M}(D)$, the map $x \mapsto S_f(x)$ is continuous.*

Remark 1.1.2. 1. Hence, the vaguely measurable structure on $\mathbb{S}(D)$ is given as the coarsest σ-algebra such that all the functionals S_f with $f \in \mathcal{C}_{\mathrm{c}}(D)$ are measurable, and the τ-measurability is given by the same with $f \in \mathcal{M}(D)$.

2. Obviously, every vaguely open set is also τ-open, i.e., the τ-topology is finer than the vague one. This implies that every $\mathbb{S}(D)$-valued random variable that is vaguely measurable is also τ-measurable.

3. We will in the following, if nothing else is stated, consider the τ-measurability (i.e., the Borel measurability induced by the τ-topology) and call this just *measurability*. Hence, vaguely measurable functions are in particular measurable. As soon as we will consider convergence in Section 1.7, we will mainly work with vague measurability, since convergence of integrals against continuous functions are much easier to handle.

4. One can easily extend the above topologies to locally compact topological spaces D, and we will make use of that later in Section 1.4.

5. Let us note that the theory of point processes can be developed in much greater generality (see, e.g., [LP17]) where instead of $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ a general measurable space $(W, \mathcal{W})$ is considered without any reference to topologies. Then, $\mathbb{S}(D)$ is usually replaced by the space $\mathbf{N}$ of all measures that can be written as a countable sum of measures ν with the property that $\nu(B) \in \mathbb{N}_0$ for all $B \in \mathcal{W}$. A σ-algebra on $\mathbf{N}$ can then for example be defined via generating sets of the form $\{\nu \in \mathbf{N} | \colon \nu(B) = k\}$ with $B \in \mathcal{W}, k \in \mathbb{N}_0$. ◇

Like in the well-known *Portmanteau theorem*[1], there are a number of useful characterizations of these topologies. We pick just one. For a given cloud $x = (x_i)_{i \in I} \in \mathbb{S}(D)$, we denote the number of its points in a given measurable set $A \subset D$ by

$$N_x(A) = \#\{i \in I \colon x_i \in A\} = S_{\mathbb{1}_A}(x) \in \mathbb{N}_0 \cup \{\infty\}, \tag{1.1.3}$$

where we wrote $\mathbb{1}_A(z) = 1$ if $z \in A$ and $\mathbb{1}_A(z) = 0$ otherwise for the indicator function on A.

Lemma 1.1.3 (Characterization of distributions)**.** *The distribution of an $\mathbb{S}(D)$-valued random variable $\mathbb{X}$ is uniquely determined by the distributions of all the vectors $(N_{\mathbb{X}}(A_1), \dots, N_{\mathbb{X}}(A_k))$ with $k \in \mathbb{N}$ and measurable bounded sets $A_1, \dots, A_k \subset D$.*

In analogy with stochastic processes with parameter set $\mathbb{N}_0$ instead of D, one can see these vectors as defining the finite-dimensional distributions of the point process $\mathbb{X}$.

Another important characterization of the distribution of an $\mathbb{S}(D)$-valued random variable $\mathbb{X}$ is in terms of its *Laplace transform* defined by

$$\mathcal{L}_{\mathbb{X}}(f) = \mathbb{E}[\mathrm{e}^{-\sum_{i \in I} f(X_i)}] \in [0, 1], \qquad f \colon D \to [0, \infty) \text{ measurable.} \tag{1.1.4}$$

Lemma 1.1.4 (Laplace transform fixes distributions)**.** *The distribution of an $\mathbb{S}(D)$-valued random variable $\mathbb{X}$ is uniquely determined by its Laplace transform for all measurable nonnegative functions f with compact support.*

Remark 1.1.5. There are obvious analogues to Lemma 1.1.3 and 1.1.4 for vague measurability, where the boundaries of the sets $A_1, \dots, A_k$ are required to be nullsets with respect to the Lebesgue measure on D, respectively where the Laplace transform is taken only for continuous nonnegative functions f with compact support. ◇

For describing the distribution of a random point cloud, it appears natural to do this in terms of a measure μ on D, which gives a first rough idea how many points of $\mathbb{X}$ are located in a given set.

[1] The Portmanteau theorem states that weak convergence of probability measures (defined by convergence of all the integrals against continuous bounded functions) is equivalent to convergence of their masses of any measurable set whose boundary is a nullset with respect to the limiting measure.

Definition 1.1.6 (Intensity measure)**.** *The* intensity measure μ *of a random point cloud* $\mathbb{X}$ *in* D *is defined by*

$$\mu(A) = \mathbb{E}[N_{\mathbb{X}}(A)], \qquad A \subset D \text{ measurable.} \tag{1.1.5}$$

However, the intensity measure is by far not enough to fully characterize the distribution of the random point process $\mathbb{X}$.

An important large class of random point clouds are the stationary ones. By $A + z = \{a + z \colon a \in A\}$ we denote the spatial shift by $z \in \mathbb{R}^d$ of a set $A \subset \mathbb{R}^d$.

Definition 1.1.7 (Stationary random point clouds)**.** *A random point cloud* $\mathbb{X}$ *in* $\mathbb{R}^d$ *(more precisely, its distribution) is called* stationary *if its distribution is identical to the one of* $\mathbb{X} + z$ *for any* $z \in \mathbb{R}^d$*.*

Observe that the term makes sense only for $D = \mathbb{R}^d$. Sometimes, the term *homogeneous* is used instead of stationarity. The intensity measure of a stationary random point cloud is invariant under shifts and therefore equal to a multiple of the Lebesgue measure on $\mathbb{R}^d$, which we will denote by Leb in the sequel.

Lemma 1.1.8 (Intensity of stationary point processes)**.** *Let* $\mathbb{X}$ *be a point process on* $\mathbb{R}^d$ *with intensity measure* μ *such that* $\mu([0,1]^d) < \infty$*. If* $\mathbb{X}$ *is stationary, then* $\mu = \lambda\text{Leb}$ *with* $\lambda = \mu([0,1]^d)$*.*

Proof. By stationarity $\mu(B + x) = \mu(B)$ for all measurable $B \subset \mathbb{R}^d$ and $x \in \mathbb{R}^d$, but λLeb is the only measure with these two properties. □

1.2 Definition and first properties of the Poisson point process

In this section, we introduce a very particular random point process, which is characterized by a very high degree of independence.

Definition 1.2.1 (Poisson point process)**.** *Let* μ *be a measure on* D *that gives finite values for any bounded subset, i.e., a Radon measure. We call the random point process* $\mathbb{X}$ *a* Poisson point process (PPP) *with* intensity measure μ *if, for any* $k \in \mathbb{N}$ *and any pairwise disjoint bounded measurable sets* $A_1, \dots, A_k \subset D$*, the counting variables* $N_{\mathbb{X}}(A_1), \dots, N_{\mathbb{X}}(A_k)$ *are independent Poisson-distributed random variables with parameters* $\mu(A_1), \dots, \mu(A_k)$*, i.e., if*

$$\mathbb{P}\big(N_{\mathbb{X}}(A_1) = n_1, \dots, N_{\mathbb{X}}(A_k) = n_k\big) = \prod_{i=1}^{k} \Big[\mathrm{e}^{-\mu(A_i)} \frac{\mu(A_i)^{n_i}}{n_i!}\Big], \qquad n_1, \dots, n_k \in \mathbb{N}_0. \tag{1.2.1}$$

In Figure 1.2.1 we present a realization of a PPP. Let us further make a number of comments:

Remark 1.2.2. 1. We can certainly also drop the set D, i.e., put it equal to $\mathbb{R}^d$, since the dependence on D can be absorbed in μ. Indeed, a measure μ on D can be trivially extended to a measure on $\mathbb{R}^d$ with the value zero, and the PPPs that are induced by μ on D and the one that is induced by its extension on $\mathbb{R}^d$ are equal to each other in distribution, after restricting to D.

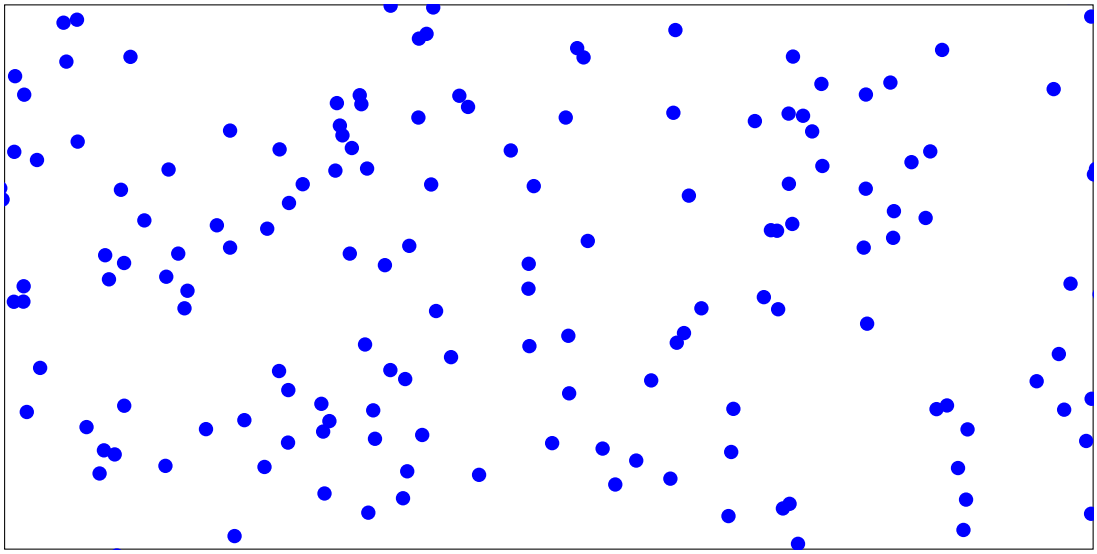

Figure 1.2.1: Realization of a stationary PPP.

2. From Lemma 1.1.3 it follows that each Radon measure on $\mathbb{R}^d$ is the intensity measure of a unique PPP, up to distribution.

3. It is clear that the space D (or $\mathbb{R}^d$) can be widely be generalized for Definition 1.2.1 to make sense; in fact we need only a measure μ on an arbitrary measurable space D that can be written as a countable sum of finite measures on D, see, e.g., [LP17]. For our setting of point processes, based on (1.1.1), we need to assume that D is a locally compact topological space.

4. What we called the intensity measure of a PPP in Definition 1.2.1 is consistent with Definition 1.1.6, as is seen easily. ◇

May be here is a good point to recall and collect some important properties of the Poisson distribution.

Remark 1.2.3 (The Poisson distribution)**.** The Poisson distribution with parameter $\alpha \in (0, \infty)$, or a Poisson-distributed random variable N with parameter α, is given by $\mathrm{Po}_\alpha(k) = \mathbb{P}(N = k) = \mathrm{e}^{-\alpha}\frac{\alpha^k}{k!}$ for $k \in \mathbb{N}_0$. It has expectation and variance equal to α. Here are more interesting properties:

1. Its generating function is $\sum_{k\in\mathbb{N}_0} \mathrm{Po}_\alpha(k)s^k = \mathbb{E}[s^N] = \mathrm{e}^{-(1-s)\alpha}$ for $s \in [0, \infty)$.

2. Its Laplace transform is $\mathcal{L}_\alpha(r) = \sum_{k\in\mathbb{N}_0} \mathrm{Po}_\alpha(k)\mathrm{e}^{-rk} = \mathbb{E}[\mathrm{e}^{-rN}] = \exp\{\alpha(\mathrm{e}^{-r} - 1)\}$ for $r \in [0, \infty)$.

3. The sum of an arbitrary finite number of independent Poisson-distributed random variables is also Poisson-distributed, and the parameter is the sum of the parameters of the summands.

4. The Poisson limit theorem states that the distribution of a sum of n independent Binomial-distributed random variables with parameter α/n converges towards Po_α as $n \to \infty$.

5. Given a Poisson-distributed random variable N with parameter α and independent and identically distributed (i.i.d.) Bernoulli random variables $Y_1, \dots, Y_N$ with parameter p, then $\sum_{i=1}^{N} Y_i$ is Poisson-distributed with parameter $p\alpha$.

6. If N is Poisson-distributed with parameter αt with $\alpha, t > 0$, then $\mathbb{P}(N \geq k) = \mathbb{P}(E_1 + \dots + E_k \leq t)$ for any $k \in \mathbb{N}$, where $(E_i)_{i\in\mathbb{N}}$ is a sequence of i.i.d. variables having the exponential distribution with parameter α.

The superposition principle 5. is the main reason why the Poisson distribution is the 'right' distribution for the PPP, which can be seen by considering the distribution of the PPP non-disjoint sets. ◇

Definition 1.2.1 is only in terms of the counting variables $N_{\mathbb{X}}(A)$, but we would like to have the object $\mathbb{X}$ also as an explicit $\mathbb{S}(D)$-valued random variable. This is provided by the following construction.

Lemma 1.2.4 (Construction of a PPP)**.** *Assume that μ is a measure on D with $\mu(D) \in (0, \infty)$. Let $N(D)$ be a Poisson random variable with parameter $\mu(D)$. Put $I = \{1, \dots, N(D)\}$. Given $N(D)$, let $\mathbb{X} = (X_i)_{i\in I}$ be a collection of independent random points in D with distribution $\mu(\cdot)/\mu(D)$. Then the counting variables defined in (1.1.3) form a PPP with intensity measure μ in the sense of Definition 1.2.1.*

This construction works *a priori* only for finite intensity measures, but if μ is infinite, but σ-finite, then one can decompose D into countably many measurable sets with finite μ-measure, construct the point process on the partial sets according to Lemma 1.2.4 independently and put all these point processes together in order to obtain a PPP with intensity measure μ on D. It is an exercise to show that this construction works and that the resulting point process is independent of the chosen decomposition of D; this is basically the proof of Lemma 1.2.8 below.

Remark 1.2.5 (Absolutely continuous intensity measures)**.** If μ has a Lebesgue density, more generally if it has no atoms (sites $x \in D$ with $\mu(\{x\}) > 0$), then it is easy to see that the points of the corresponding PPP are almost surely located at mutually distinct sites, i.e., $X_i \neq X_j$ for any $i \neq j$. Since we want to model the locations of human beings, we will assume this in most of the following. ◇

Example 1.2.6.

1. The *standard PPP* on $D = \mathbb{R}^d$ is obtained for the intensity measure $\lambda\mathrm{Leb}$, where $\lambda \in (0, \infty)$ is the *intensity*. Since Leb is shift-invariant, the corresponding point process is stationary (see Definition 1.1.7) and often referred to as a *homogeneous PPP*. Actually, since λLeb is the only shift-invariant measure on $\mathbb{R}^d$, every shift-invariant PPP has this as its intensity measure for some $\lambda \in (0, \infty)$, see Lemma 1.1.8. Furthermore, its distribution is also isotropic, i.e., rotationally invariant.

2. For $D = \mathbb{Z}^d$ and μ the counting measure on D, the corresponding PPP is a discrete variant of the standard PPP. It is obtained by realizing independent and identically distributed Poisson variables for each $z \in \mathbb{Z}^d$ and putting that number of points into z. Alternatively, one can, for any finite set $\Lambda \subset \mathbb{Z}^d$, generate a Poisson random variable $N(\Lambda)$ with parameter $\#\Lambda$, and distribute $N(\Lambda)$ points independently and uniformly over Λ, decompose $\mathbb{Z}^d$ into such sets and add all these points independently in all of $\mathbb{Z}^d$; the resulting superposition is the desired PPP.

3. If we want to model the locations of users in a given city D, then μ should reflect areas with low density like lakes, forest and fields, where the Lebesgue density of μ should be low, and high densities like highly frequented areas and places, where it should be high.

4. The one-dimensional standard PPP has enormous importance in the modeling of random times, in particular in the theory of time-homogeneous Markov chains in continuous time. For $d = 1$, another interesting characterization of the PPP is possible: the distances of neighboring pairs of points of the PPP are i.i.d. exponentially distributed random variables with the same parameter as the PPP has; see also the last property that we mention in Remark 1.2.3. This is directly connected with the famous property of *memorylessness* of the process of times at which the points appear, see also Chapter 5. However, we will not elaborate on these nice properties here, since we are mainly interested in $d \geq 2$. ◇

Example 1.2.7 (Contact distance)**.** The *contact distance* of a space point $u \in D$ to a set $x = \{x_i : i \in I\} \in \mathbb{S}(D)$ is defined by

$$\operatorname{dist}(u, x) = \inf\{\|u - x_i\| : i \in I\}. \tag{1.2.2}$$

This is the radius of the largest ball around u that contains no point of x. If $\mathbb{X}$ is a PPP with intensity measure μ, then the distribution function of $\operatorname{dist}(u, \mathbb{X})$ is easy to find:

$$\mathbb{P}(\operatorname{dist}(u, \mathbb{X}) < r) = \mathbb{P}(B_r(u) \cap \mathbb{X} \neq \emptyset) = 1 - \mathbb{P}(N_{\mathbb{X}}(B_r(u)) = 0) = 1 - \mathrm{e}^{-\mu(B_r(u))}, \tag{1.2.3}$$

where $B_r(u)$ is the open ball around u with radius r. ◇

Let us derive some important and simple properties of Poisson point processes. First, we identify the distribution of the superposition of independent such processes.

Lemma 1.2.8 (Superposition of PPPs)**.** *Let μ_1 and μ_2 be two measures on D and let $\mathbb{X}^{(1)} = (X_i^{(1)})_{i \in I_1}$ and $\mathbb{X}^{(2)} = (X_i^{(2)})_{i \in I_2}$ be independent PPPs with intensity measures μ_1 and μ_2, respectively. Then, $\{X_i^{(1)} : i \in I_1\} \cup \{X_i^{(2)} : i \in I_2\}$ is a PPP with intensity measure $\mu_1 + \mu_2$.*

The proof uses the well-known property of the sum of independent Poisson random variables to be again Poisson, see Remark 1.2.3. An extension to superpositions of countably many independent PPPs is straightforward.

Another operation that goes well with PPPs is random thinning, i.e., the random removal of some of the points.

Lemma 1.2.9 (Random thinning of PPPs)**.** *Let $\mathbb{X} = (X_i)_{i \in I}$ be a PPP in D with intensity measure μ. With a probability $p \in [0, 1]$, given $\mathbb{X}$, we keep independently any of the particles X_i. Then, the remains are a PPP with intensity measure $p\mu$.*

Also proving this is an elementary exercise, which is based on Property (5) in Remark 1.2.3.

Now we can easily realize many PPPs on one probability space with many different intensity measures:

Corollary 1.2.10 (Realization of superpositions)**.** *Let μ be a measure on D with $\mu(D) \in (0, \infty)$, then we can, for any $p \in [0, 1]$, construct the PPPs with intensity measure $p\mu$ on one probability space as follows: Given a Poisson-distributed random variable N with parameter $\mu(D)$, we pick N i.i.d. random sites $X_1, \ldots, X_N$ in D with distribution $\mu/\mu(D)$ and N i.i.d. random variables $U_1, \ldots, U_N$ that are uniformly distributed on $[0, 1]$, independently of $X_1, \ldots, X_N$. Then $\{X_i : U_i \leq p\}$ is a PPP with intensity measure $p\mu$.*

Here is another, quite general, way to construct a PPP out of another one: by a measurable mapping.

Theorem 1.2.11 (Mapping theorem)**.** *Let $\mathbb{X} = (X_i)_{i\in I}$ be a PPP with intensity measure μ on $D \subset \mathbb{R}^d$, and let $f\colon D \to \mathbb{R}^s$ be a measurable map. Then, $f(\mathbb{X}) = (f(X_i))_{i\in I}$ is a PPP in $\mathbb{R}^s$ with intensity measure $\mu \circ f^{-1}$.*

Let us note that, in the context of telecommunications, in order to avoid that the image PPP has an intensity measure which is not atom-less, we must assume that $\mu(f^{-1}(\{y\})) = 0$ for any $y \in \mathbb{R}^s$.

In some situations, the following formulas may be useful. Their proofs are exercises.

Lemma 1.2.12. *Let $\mathbb{X}$ be a PPP on D with intensity measure μ such that $\mu(D) \in (0,\infty)$. Then, for any measurable function $f\colon \mathbb{S}(D) \to [0,\infty)$,*

$$\mathbb{E}[f(\mathbb{X})] = \mathrm{e}^{-\mu(D)} f(\emptyset) + \mathrm{e}^{-\mu(D)} \sum_{n\in\mathbb{N}} \frac{1}{n!} \int_{D^n} f\Big(\{x_1,\dots,x_n\}\Big)\, \mu^{\otimes n}\big(\mathrm{d}(x_1,\dots,x_n)\big).$$

Lemma 1.2.13. *Let X be a PPP on D with intensity measure μ, and let A_1 and A_2 be two measurable subsets of D with $\mu(A_1), \mu(A_2) < \infty$, then the covariance of $N_{\mathbb{X}}(A_1)$ and $N_{\mathbb{X}}(A_2)$ is equal to $\mu(A_1 \cap A_2)$.*

1.3 The Campbell moment formulas

As always, we let $D \subset \mathbb{R}^d$ be a measurable (bounded or unbounded) subset of $\mathbb{R}^d$, the communication area. In Section 1.1, we defined two topologies on $\mathbb{S}(D)$ by means of the maps $x \mapsto S_f(x) = \sum_{i\in I} f(x_i)$ (see (1.1.2)) for certain functions $f\colon D \to \mathbb{R}$. Hence, the expectation $\mathbb{E}[S_f(\mathbb{X})]$ will be an important tool for characterizing the distribution of a random point process $\mathbb{X}$. We will give a formula for this.

Furthermore, the Laplace transform $\mathcal{L}_{\mathbb{X}}(f)$ (see (1.1.4)) turned out in Lemma 1.1.4 to uniquely determine the distribution of a random point process $\mathbb{X}$, hence it will also be useful to have explicit formulas for this. This functional has the great advantage that it always yields a finite value for nonnegative functions f and has very nice properties with respect to convergence of the point process, as an application of the bounded convergence theorem is always possible, see Section 1.7. We will also give a handy formula for this in the case of a PPP.

Theorem 1.3.1 (Campbell's theorem)**.** *Let $\mathbb{X}$ be a point process on D with intensity measure μ, and let $f\colon D \to \mathbb{R}$ be integrable with respect to μ, then*

$$\mathbb{E}[S_f(\mathbb{X})] = \int_D f(x)\, \mu(\mathrm{d}x). \tag{1.3.1}$$

If $\mathbb{X}$ is even a PPP, then for nonnegative f,

$$\mathcal{L}_{\mathbb{X}}(f) = \mathbb{E}[\mathrm{e}^{-S_f(\mathbb{X})}] = \exp\Big(\int_D (\mathrm{e}^{-f(x)} - 1)\, \mu(\mathrm{d}x)\Big). \tag{1.3.2}$$

The proofs are easily done with the help of a measure-theoretic induction and the fact that $\mathbb{E}[\mathrm{e}^{-\gamma Y}] = \exp\big(\lambda(\mathrm{e}^{-\gamma} - 1)\big)$ for any Poisson-distributed random variable Y with parameter $\lambda > 0$ and any $\gamma \in \mathbb{R}$.

1.4 Marked Poisson point processes

To each of the points in a point process, we may add some individual information, which may also be random. In a model for the locations for the users, this may be the strength of the transmitted signal sent out from the location of this user, or the fact whether or not the user is sending or receiving, and much more. We call this additional information the *mark* of the user. Hence, a *marked point process* in D is nothing but a random element of $\mathbb{S}(D \times \mathcal{M})$, where $\mathcal{M}$ is the set of marks, and we extended the definition of $\mathbb{S}(D)$ in (1.1.1) in an obvious way. We will write an element always of $\mathbb{S}(D \times \mathcal{M})$ as $\{(x_i, m_i) \colon i \in I\}$ or as the point measure $\sum_{i \in I} \delta_{(x_i, m_i)}$ and call m_i the mark of x_i.

In order to be able to use the notion of (1.1.1) in this way, we need to give $\mathcal{M}$ a topological structure and equip it with the corresponding Borel-σ algebra $\mathcal{B}(\mathcal{M})$. In order that we can use the topologies introduced in Section 1.1, we also assume that $\mathcal{M}$ is locally compact (see Remark 1.1.2(4).) We call $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ the *mark space*. We could equip it with a probability measure and can introduce a marked point process in which all the marks are independent of the users that they are attached to, but we would like to admit a spatial dependence. For doing this in a mathematically correct way, we need a probability kernel $K \colon D \times \mathcal{M} \to [0,1]$, i.e., a map such that $K(x, \cdot)$ is a probability measure on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ for any $x \in D$ and $K(\cdot, G)$ is measurable for any $G \in \mathcal{B}(\mathcal{M})$. Then $K(x_i, \cdot)$ is the distribution of the mark that is attached to the point x_i, and it may therefore depend on this point, but not on the index i. We keep the assumption that all the marks are independent and obtain the following notion.

Definition 1.4.1 (Marked PPP)**.** *Let $\mathbb{X} = (X_i)_{i \in I}$ be a PPP in D with finite intensity measure μ, and let $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ be a measurable space, the* mark space*. Furthermore, let K be a probability kernel from D to $\mathcal{M}$. Given $\mathbb{X}$, let $(m_i)_{i \in I}$ be an independent collection of $\mathcal{M}$-valued random variables with distribution $\otimes_{i \in I} K(X_i, \cdot)$ (where the i-th factor acts on m_i). Then, the point process $\mathbb{X}_K = ((X_i, m_i))_{i \in I}$ in $D \times \mathcal{M}$ (respectively, the point process $\sum_{i \in I} \delta_{(X_i, m_i)}$) is called a* K-marked Poisson point process (K-MPPP) *or a* K-marking *of the PPP $\mathbb{X}$.*

If K does not depend on the first argument, then $(m_i)_{i \in I}$ is even an i.i.d. collection, given $\mathbb{X}$, and then one calls $(\mathcal{M}, \mathcal{B}(\mathcal{M}), K)$ the mark space. We correspond to this marking as an *independent marking*. Let us calculate the Laplace transform of an MPPP.

Lemma 1.4.2 (Laplace transform of an MPPP)**.** *The Laplace transform of the K-marking $\mathbb{X}_K$ in Definition 1.4.1 is given by*

$$\mathcal{L}_{\mathbb{X}_K}(g) = \mathcal{L}_{\mathbb{X}}(g^*), \qquad g \colon D \times \mathcal{M} \to [0, \infty) \text{ measurable and compactly supported,} \tag{1.4.1}$$

where

$$g^*(x) = -\log\Big(\int_{\mathcal{M}} \mathrm{e}^{-g(x,y)} \, K(x, \mathrm{d}y) \Big), \qquad x \in D. \tag{1.4.2}$$

Proof. We use the so-called tower property of conditional expectations and the independence over i to see that

$$\begin{aligned} \mathcal{L}_{\mathbb{X}_K}(g) &= \mathbb{E}\Big[\mathrm{e}^{-\sum_{i \in I} g(X_i, m_i)}\Big] = \mathbb{E}\Big[\prod_{i \in I} \mathrm{e}^{-g(X_i, m_i)}\Big] = \mathbb{E}\Big[\mathbb{E}\Big[\prod_{i \in I} \mathrm{e}^{-g(X_i, m_i)} \Big| \mathbb{X}\Big]\Big] \\ &= \mathbb{E}\Big[\prod_{i \in I} \mathbb{E}\big[\mathrm{e}^{-g(X_i, m_i)} \mid X_i\big]\Big]. \end{aligned}$$

Observe that

$$\mathbb{E}\big[\mathrm{e}^{-g(X_i,m_i)}\big|X_i\big] = \int_{\mathcal{M}} K(X_i,\mathrm{d}y)\,\mathrm{e}^{-g(X_i,y)} = \mathrm{e}^{-g^*(X_i)}.$$

Substituting this yields the assertion. □

Now let us come back to the assumption that the mark space $\mathcal{M}$ is a locally compact topological space, see the remarks at the beginning of this section. Then Lemma 1.4.2 easily implies that the K-MPPP $\mathbb{X}_K$ is nothing but a usual PPP on $D\times\mathcal{M}$ with intensity measure $\mu\otimes K$ [2], where we slightly extended Definition 1.2.1 in the spirit of Remark 1.2.2(3).

Theorem 1.4.3 (Marking theorem)**.** *Let the situation of Definition 1.4.1 be given, and assume that $(\mathcal{M},\mathcal{B}(\mathcal{M}))$ is locally compact and is equipped with the Borel-σ-algebra. Then $\mathbb{X}_K$ is in distribution equal to the PPP on $D\times\mathcal{M}$ with intensity measure $\mu\otimes K$.*

Proof. From the last assertion in Remark 1.1.1, we know that the distribution of a point process in $D\times\mathcal{M}$ is uniquely determined by its Laplace transform. Hence, we only have to show that the Laplace transform of a PPP with intensity measure $\mu\otimes K$ is identical to the one of a K-MPPP. Apply (1.3.2) to Lemma 1.4.2 to see that

$$\begin{aligned}\mathcal{L}_{\mathbb{X}}(g^*) &= \exp\Big(\int_D (\mathrm{e}^{-g^*(x)}-1)\,\mu(\mathrm{d}x)\Big) = \exp\Big(\int_D\Big(\int_{\mathcal{M}} K(x,\mathrm{d}y)\,\mathrm{e}^{-g(x,y)}-1\Big)\mu(\mathrm{d}x)\Big)\\ &= \exp\Big(\int_D\int_{\mathcal{M}}\big(\mathrm{e}^{-g(x,y)}-1\big)\,\mu(\mathrm{d}x)K(x,\mathrm{d}y)\Big)\\ &= \exp\Big(\int_{D\times\mathcal{M}}(\mathrm{e}^{-g}-1)\mathrm{d}(\mu\otimes K)\Big).\end{aligned}$$

Now consult (1.3.2) once more (for $D\times\mathcal{M}$ instead of D) to see that this is the Laplace transform of a PPP with intensity measure $\mu\otimes K$. □

Having seen this, it is also clear that, it is not necessary to normalize K, since one can construct a realization of such a PPP also by first taking $\mathbb{X}$ as a PPP with intensity measure $K(\mathcal{M})\mu$ and then pick the marks with distribution $K/K(\mathcal{M})$. The reason is that, for any $c\in(0,\infty)$, the measures $c\mu\otimes K/c$ and $\mu\otimes K$ coincide.

It is also clear that, for any K-marked PPP $\{(X_i,m_i)\colon i\in I\}$ with intensity measure μ and mark measure K, the projected process $\{X_i\colon i\in I\}$ is a PPP with intensity measure μ.

Example 1.4.4. Since the d-dimensional Lebesgue measure is the d-fold product measure of the one-dimensional Lebesgue measure, one could think that the standard PPP in $\mathbb{R}^d$ can be seen as a marked PPP in $\mathbb{R}^{d-1}$ with marks in $\mathbb{R}$. However, since the Lebesgue measure on $\mathbb{R}$ is not finite, this is not covered by Definition 1.4.1. If the last factor $\mathbb{R}$ is replaced by some bounded measurable set and the Lebesgue measure by the restriction, then this interpretation is correct. ◇

[2] Note that the measure $\mu\otimes K$ is defined by $\mu\otimes K(B)=\int_{B^{(1)}}\mu(\mathrm{d}x)\,K(x,B_x^{(2)})$, where $B^{(1)}=\{x\in D\colon\exists y\in\mathcal{M}\colon(x,y)\in B\}$ and $B_x^{(2)}=\{y\in\mathcal{M}\colon(x,y)\in B\}$.

1.5 Conditioning: the Palm version

Let $\mathbb{X}$ be a stationary point process on $\mathbb{R}^d$ that is, the distribution of $\mathbb{X}$ equals the one of $\mathbb{X}+z$ for any $z \in \mathbb{R}^d$. We would like to imagine that we are standing in one of the points $X_i \in \mathbb{X}$ of the process and watch the other points from there. In other words, we are interested in the process $\mathbb{X} - X_i$, seen from the perspective of X_i. To be sure, this point X_i should be a *typical* one, i.e., not a point that is sampled according to any specific criterion. Since we are in a stationary setting, the randomly chosen point X_i can be assumed to be located at the origin. Hence, as we will explain heuristically below, we would like to look at the conditional version of $\mathbb{X}$ given $0 \in \mathbb{X}$. The definition of this object needs some care, since the event $\{0 \in \mathbb{X}\}$ has probability zero for stationary processes. The mathematically sound setup for this is *Palm theory*. Let us start by giving the associated existence and uniqueness result including the proof.

Theorem 1.5.1 (Refined Campbell theorem)**.** *Suppose that $\mathbb{X}$ is a stationary point process on $\mathbb{S}(\mathbb{R}^d)$ with finite positive intensity λ. Then, there exists a unique probability measure P^o on $\mathbb{S}(\mathbb{R}^d)$ such that*

$$\lambda^{-1}\mathbb{E}\Big[\sum_{i\in I} f(X_i, \mathbb{X}-X_i)\Big] = \int E^o[f(x,\cdot)]\,\mathrm{d}x, \qquad f\colon \mathbb{R}^d \times \mathbb{S}(\mathbb{R}^d) \to [0,\infty) \text{ measurable.} \quad (1.5.1)$$

The measure P^o is called the Palm distribution *of $\mathbb{X}$.*

Proof. We prove the statement for $f = \mathbb{1}_{B\times A}$ for bounded measurable $B \subset \mathbb{R}^d, A \subset \mathbb{S}(\mathbb{R}^d)$. The full statement then follows by the usual monotone class arguments. We define

$$\nu_A(B) = \mathbb{E}\Big[\sum_{i\in I} \mathbb{1}\{X_i \in B\}\mathbb{1}\{\mathbb{X}-X_i \in A\}\Big].$$

Then, by stationarity,

$$\begin{aligned}\nu_A(B+z) &= \mathbb{E}\Big[\sum_{i\in I} \mathbb{1}\{X_i - z \in B\}\mathbb{1}\{\mathbb{X}-X_i \in A\}\Big]\\ &= \mathbb{E}\Big[\sum_{i\in I} \mathbb{1}\{X_i \in B\}\mathbb{1}\{\mathbb{X}-X_i \in A\}\Big] = \nu_A(B),\end{aligned}$$

and thus ν_A is also stationary. Further, since $\nu_A(B) \le \mathbb{E}(N_{\mathbb{X}}(B)) = \lambda|B|$, ν_A is also locally finite and thus ν_A must be equivalent to $\lambda_A\mathrm{Leb}$ with $\lambda_A = \nu_A([0,1]^d)$. Then, defining the probability measure $P^0(A) = \lambda_A/\lambda$, we have

$$\mathbb{E}\Big[\sum_{i\in I} \mathbb{1}\{X_i \in B\}\mathbb{1}\{\mathbb{X}-X_i \in A\}\Big] = \lambda P^0(A)|B|.$$

Conversely, for $B \subset \mathbb{R}^d$ with $0 < |B| < \infty$, the equality (1.5.1) yields

$$P^0(A) = (\lambda|B|)^{-1}\mathbb{E}\Big[\sum_{i\in I} \mathbb{1}\{X_i \in B\}\mathbb{1}\{\mathbb{X}-X_i \in A\}\Big],$$

which shows uniqueness. □

Remark 1.5.2. 1. It is convenient to introduce an $\mathbb{S}(\mathbb{R}^d)$-valued random variable $\mathbb{X}^*$ with distribution P^o on the original probability space. For example, if f does not depend on x, we write

$$\lambda^{-1}\mathbb{E}\Big[\sum_{i\in I} f(\mathbb{X}-X_i)\Big] = \mathbb{E}[f(\mathbb{X}^*)].$$

2. Let us provide some illustration and interpretation for the Palm distribution. The l.h.s. of equation (1.5.1) can be interpreted as the probability of an event for $\mathbb{X}$, seen from a 'typical' user $X_i \in \mathbb{X}$, i.e., from a user that is picked uniformly at random. But what means 'uniformly' for an infinite point cloud? And what about the r.h.s. of (1.5.1)? To give some substance to the idea of picking a 'typical' point, we pick X_i uniformly at random from the stationary point process $\mathbb{X}$ with intensity measure λLeb for some $\lambda \in (0,\infty)$, in some compact set A with positive Lebesgue measure, say a centered box. That is, we consider the distribution of $\sum_{i\in I}\delta_{X_i}(A)\delta_{\{\mathbb{X}-X_i\}}$, properly normalized. Note that the normalization is $1/\mathbb{E}[\sum_{i\in I}\delta_{X_i}(A)] = 1/\mathbb{E}[N_{\mathbb{X}}(A)] = 1/\lambda\mathrm{Leb}(A)$. The probability of an event Γ in $\mathbb{S}(\mathbb{R}^d)$ is then equal to

$$\frac{1}{\lambda\mathrm{Leb}(A)}\mathbb{E}\Big[\sum_{i\in I}\delta_{X_i}(A)\mathbb{1}\{\mathbb{X}-X_i\in\Gamma\}\Big].$$

Actually, it turns out that this does not depend on A. Indeed, considering a partition $(D_k)_{1\le k\le n}$ of A consisting of connected Lebesgue-positive sets, we can rewrite this as

$$\frac{1}{\lambda\mathrm{Leb}(A)}\sum_{k=1}^{n}\mathbb{E}\Big[\sum_{i\in I}\delta_{X_i}(D_k)\mathbb{1}\{\mathbb{X}-X_i\in\Gamma\}\,\Big|\,N_{\mathbb{X}}(D_k)>0\Big]\,\mathbb{P}(N_{\mathbb{X}}(D_k)>0).$$

Now, considering the limit of finer and finer partitions, we observe that $\mathbb{P}(N_{\mathbb{X}}(D_k)>0) = \lambda|D_k| + o(\lambda|D_k|)$ and thus the above sum, in the spirit of a Riemann sum, should converge to a limiting expression of the form

$$\frac{1}{\lambda\mathrm{Leb}(A)}\lambda\int_A \mathbb{E}\big[\mathbb{1}\{\mathbb{X}-x\in\Gamma\}|x\in\mathbb{X}\big]\,\mathrm{d}x. \tag{1.5.2}$$

Now, by translation invariance, $\mathbb{E}\big[\mathbb{1}\{\mathbb{X}-x\in\Gamma\}|x\in\mathbb{X}\big] = \mathbb{P}(\mathbb{X}\in\Gamma|o\in\mathbb{X})$, where we write o for the origin. We arrive at the heuristic equality

$$\frac{1}{\mathbb{E}[N_{\mathbb{X}}(A)]}\mathbb{E}\Big[\sum_{i\in I}\delta_{X_i}(A)\mathbb{1}\{\mathbb{X}-X_i\in\Gamma\}\Big] = \mathbb{P}(\mathbb{X}\in\Gamma|o\in\mathbb{X}).$$

Note that the right-hand side is independent of A. The equality explains heuristically the relationship between the idea of a PPP seen from a typical point and the process conditioned on having a point at the origin. The distribution $\mathbb{P}(\mathbb{X}\in\cdot|o\in\mathbb{X}) = P^0(\cdot) = \mathbb{P}(\mathbb{X}^*\in\cdot)$ is then the Palm version of the distribution of $\mathbb{X}$.

3. Following the same line of ideas, a closely related result can be formulated, called the *reduced Campbell-Little-Mecke formula*, which does not use stationarity. It states the existence of the *reduced Palm distribution* $P_x^!$ for $x\in\mathbb{R}^d$ that is characterized by the equation

$$\mathbb{E}\Big[\sum_{i\in I} f(X_i,\mathbb{X}-\delta_{X_i})\Big] = \int E_x^![f(x,\cdot)]\,\mu(\mathrm{d}x),$$

where μ is the intensity measure of $\mathbb{X}$. Note that here the process is not shifted but rather a random point is removed, compare also to the expression (1.5.2). ◇

Theorem 1.5.1 is formulated for general stationary point processes. In the special case of a homogeneous PPP, the Palm distribution has a particularly simple form, as can be understood from the following argument. Since the event $\{o \in \mathbb{X}\}$ has measure zero under the PPP $\mathbb{X}$, we instead condition, for some $\varepsilon > 0$, on the event $\{N_{\mathbb{X}}(B_\varepsilon(0)) = 1\}$ (which has positive probability) and perform the limit $\varepsilon \downarrow 0$. Let us see what this gives for a counting variable $N_{\mathbb{X}}(A)$ for some bounded open $0 \in A \subset \mathbb{R}^d$. Then,

$$\begin{aligned}\mathbb{P}(N_{\mathbb{X}}(A) = n \mid N_{\mathbb{X}}(B_\varepsilon(0)) = 1) &= \frac{\mathbb{P}(N_{\mathbb{X}}(A \setminus B_\varepsilon(0)) = n-1, N_{\mathbb{X}}(B_\varepsilon(0)) = 1)}{\mathbb{P}(N_{\mathbb{X}}(B_\varepsilon(0)) = 1)} \\ &= \mathbb{P}(N_{\mathbb{X}}(A \setminus B_\varepsilon(0)) = n-1) \\ &\to \mathbb{P}(N_{\mathbb{X}}(A \setminus \{0\}) = n-1) = \mathbb{P}(N_{\mathbb{X}\cup\{0\}}(A) = n)\end{aligned}$$

as $\varepsilon \downarrow 0$. This suggests that the limiting conditioned process should be nothing but the process $\mathbb{X} \cup \{0\}$. This is made precise in the following result. It states that Poisson processes are even characterized by this property.

Theorem 1.5.3 (Stationary Mecke-Slivnyak theorem)**.** *Let $\mathbb{X}$ be a stationary point process with intensity $\lambda > 0$. Then, $\mathbb{X}$ is a PPP if and only if*

$$\mathbb{E}[f(\mathbb{X}^*)] = \mathbb{E}[f(\mathbb{X} \cup \{0\})], \qquad f\colon \mathbb{S}(\mathbb{R}^d) \to [0,\infty) \text{ measurable.}$$

Remark 1.5.4.

1. The stationary Mecke-Slivnyak theorem is a special case of the more general *Mecke-Slivnyak theorem*, which does not use stationarity but some mild assumptions on the intensity measure μ. It states that a PPP $\mathbb{X}$ is characterized by the equation

$$\mathbb{E}\Big[\sum_{i\in I} f(X_i, \mathbb{X})\Big] = \int \mathbb{E}\big[f(x, \mathbb{X} \cup \{x\})\big]\, \mu(\mathrm{d}x), \qquad f\colon \mathbb{R}^d \times \mathbb{S}(\mathbb{R}^d) \to [0,\infty) \tag{1.5.3}$$

   In other words, for PPP, the reduced Palm distribution is equal to the original distribution.

2. The Mecke-Slivnyak theorem can also be seen as a generalization of Campbell's theorem 1.3.1, which considers functions $f(X_i, \mathbb{X}) = f(X_i)$ not depending on $\mathbb{X}$. ◇

Example 1.5.5 (Contact distance distribution for homogeneous PPPs)**.** Recall from Remark 1.2.7 the contact distance $\mathrm{dist}(u, x)$ of a space point $u \in D$ to a point set $x = \{x_i\colon i \in I\} \in \mathbb{S}(D)$. If $\mathbb{X}$ is a homogeneous PPP with intensity λ, then $\mathbb{P}(\mathrm{dist}(u, \mathbb{X}) \le r) = \mathbb{P}(\mathrm{dist}(o, \mathbb{X}^*) \le r)$ for all $u \in D$. In words, the distance of a typical point from a homogeneous PPP to its nearest neighbor in the PPP is distributed exactly as the distance from any fixed point. ◇

1.6 Random intensity measures: Cox point processes

Modeling a system of telecommunication devices in space via a homogeneous PPP represents a situation where no information about the environment or any preferred behavior of the devices is available. To some degree this can be compensated by the use of a non-homogeneous PPP with intensity measure μ, where now areas can be equipped with higher or lower user density. Thereby we leave the mathematically nicer setting of spatial stationarity, but at least we keep the spatial independence. Nevertheless, also the independence of devices is an assumption that is often violated in the real world and user behavior is usually correlated. One way to incorporate dependencies into the distribution of devices in space is to use *Cox point processes*.

In simple words, a *Cox point processes* is a PPP with random intensity measure Λ. The *directing random measure* Λ can be interpreted as a *random environment* and the resulting processes is thus constructed via a two-step stochastic procedure. More specifically, let $D \subset \mathbb{R}^d$ be a measurable set, and we assume Λ to be a random element of the space of all σ-finite measures $\mathcal{M}(D)$ on D equipped with the smallest sigma algebra such that all evaluation mappings $\mu \mapsto \mu(B)$ from $\mathcal{M}(D)$ to $[0,\infty]$ are measurable for all measurable $B \subset D$. We call such Λ a *random measure* on D.

Definition 1.6.1 (Cox point processes)**.** *Let Λ be a random measure on D, then the PPP $\mathbb{X}$ with random intensity measure Λ is called a* Cox point process directed by Λ.

For a realization of Cox point process see Figure 1.6.1. Let us make some comments.

Remark 1.6.2 (Properties of Cox point processes)**.** 1. The expected number of points in a measurable volume $A \subset D$ is given by the expected intensity of A, i.e.,

$$\mathbb{E}[N_{\mathbb{X}}(A)] = \mathbb{E}[\mathbb{E}[N_{\mathbb{X}}(A)|\Lambda]] = \mathbb{E}[\Lambda(A)].$$

2. The Laplace transform of a Cox point process is given by

$$\mathcal{L}_{\mathbb{X}}(f) = \mathbb{E}[\mathrm{e}^{-S_f(\mathbb{X})}] = \mathbb{E}\Big[\exp\Big(\int_D (\mathrm{e}^{-f(x)} - 1)\,\Lambda(\mathrm{d}x)\Big)\Big], \tag{1.6.1}$$

for all measurable $f\colon D \to [0,\infty)$. ◇

The theory of Cox point processes provides a broad setting for modeling interesting spatial telecommunication systems.

Example 1.6.3 (Absolutely continuous random fields)**.** A large class of random environments Λ is given by measures having a non-negative random field $\ell = \{\ell_x\}_{x\in\mathbb{R}^d}$ as a Lebesgue density, i.e., $\Lambda(\mathrm{d}x) = \ell_x\,\mathrm{d}x$. On $\mathbb{R}^d$, one often assumes ℓ to be stationary. For example, this includes random *measures modulated by a random closed set* Ξ, see [CSKM13, Section 5.2.2]. Here, $\ell_x = \lambda_1 \mathbb{1}\{x \in \Xi\} + \lambda_2 \mathbb{1}\{x \notin \Xi\}$ with parameters $\lambda_1, \lambda_2 \geq 0$. For instance, Ξ could be given by the Boolean model $\bigcup_{j\in J} B_r(Y_j)$ of a PPP $(Y_j)_{j\in J}$, see Chapter 2, interpreted as a random configuration of hot spots. Another important example is a random measure induced by a *shot-noise field*[3], see [CSKM13, Section 5.6]. Here, $\ell_x = \sum_{j\in J} k(x - Y_j)$ for some non-negative integrable kernel $k\colon \mathbb{R}^d \to [0,\infty)$ with compact support and a PPP $(Y_j)_{j\in J}$. ◇

Example 1.6.4 (Random street systems)**.** Very interesting for a realistic modeling of an urban area is a random environment Λ that is defined as the restriction of the Lebesgue measure to a random segment process S in $\mathbb{R}^d$, we think of $d = 2$. That is, S is a point process in the space of line segments [CSKM13, Chapter 8], which we want to assume as stationary for simplicity. Observe that S is a union of one-dimensional subsets, in particular a nullset with respect to the two-dimensional Lebesgue measure. However, there is a natural measure ν_1 on S that attaches a finite and positive value to each bounded non-trivial line segment (indeed, its length). This measure is the one-dimensional Hausdorff measure ν_1 on S. Then we put $\Lambda(\mathrm{d}x) = \nu_1(S \cap \mathrm{d}x)$ and obtain a random measure on $\mathbb{R}^2$ that is concentrated on S. Indeed, this random environment Λ is singular with respect to the two-dimensional Lebesgue measure.

[3]The name *shot noise* = *Schroteffekt* comes from the choice of k as a Gaussian-shaped function, which approaches the outcome of a lead shot.

There are a number of interesting choices of S, some of which have high relevance as models for street systems of cities. In Remark 1.6.5 we introduce several tessellations, the Poisson-Voronoi, Poisson-Delaunay and the Poisson line tessellations. See also Figure 1.6.1 for an illustration. ◇

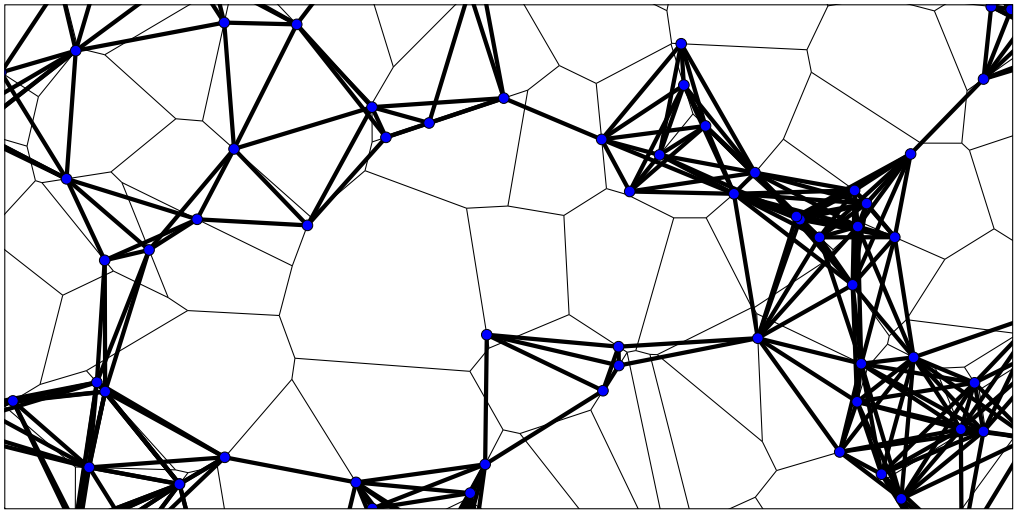

Figure 1.6.1: Realization of the Gilbert graph, see Chapter 2, of users confined to a street system given by a Poisson-Voronoi tessellation.

Remark 1.6.5 (Tessellations)**.** Prominent examples of singular random environments as in Example 1.6.4 are given by directing random measures Λ derived from some tessellation process S. These random tessellations are also of independent interest and can be used in a number of seemingly unrelated branches of mathematics, such as numerical methods for partial differential equations.

The most common one is the *Poisson-Voronoi tessellation (PVT)*, which we introduce now. Consider a PPP $\mathbb{X} = (X_i)_{i\in I}$ in D. We assign to each $X_i \in \mathbb{X}$ the *cell*

$$z(X_i) = z(X_i, \mathbb{X}) = \Big\{x \in D\colon \|x - X_i\| \le \inf_{j\in I} \|x - X_j\|\Big\}. \tag{1.6.2}$$

In words, the interior of $z(X_i)$ contains all points in D that are closer to X_i than to any other point in $\mathbb{X}$. Now, D is partitioned into cell interiors and cell boundaries, which motivates the term tessellation. It can be proved that the cell boundaries are polygon lines. See Figure 1.6.2 on the left for an illustration. If the underlying PPP $\mathbb{X}$ is homogeneous, then the distribution of the PVT is translation invariant and isotropic, i.e., invariant with respect to rotations around the origin. A number of important characteristics of the PVT, such as the expected cell volume etc., can be calculated from the intensity of $\mathbb{X}$, see [OBSC00, Table 5.1.1].

In telecommunication applications, one can see X_i as the location of a base station and $z(X_i)$ as its serving zone, but this is not the interpretation that we are after here. Instead, we interpret S as a random street system, and indeed there is some statistical evidence indicating that Poisson-Voronoi tessellations give decent fits to street systems in central European cities.

Let us mention two more examples of tessellation processes. First, the *Poisson-Delaunay tessellation (PDT)* is the dual tessellation corresponding to the PVT. Here line segments are

drawn such that any two cell centers X_i are connected by a line if and only if this line crosses exactly one cell boundary (or face in higher dimensions) in the PVT S. The PDT naturally has very similar locality properties as PVT, but completely different behavior for examples of its vertex degree. More precisely, the typical Poisson-Voronoi cell has 6 line segments leading to a degree of 6 for the typical Poisson-Delaunay vertex. But, in the PVT, with probability one, only 3 line segments meet in a vertex. See Figure 1.6.2 on the right for an illustration.

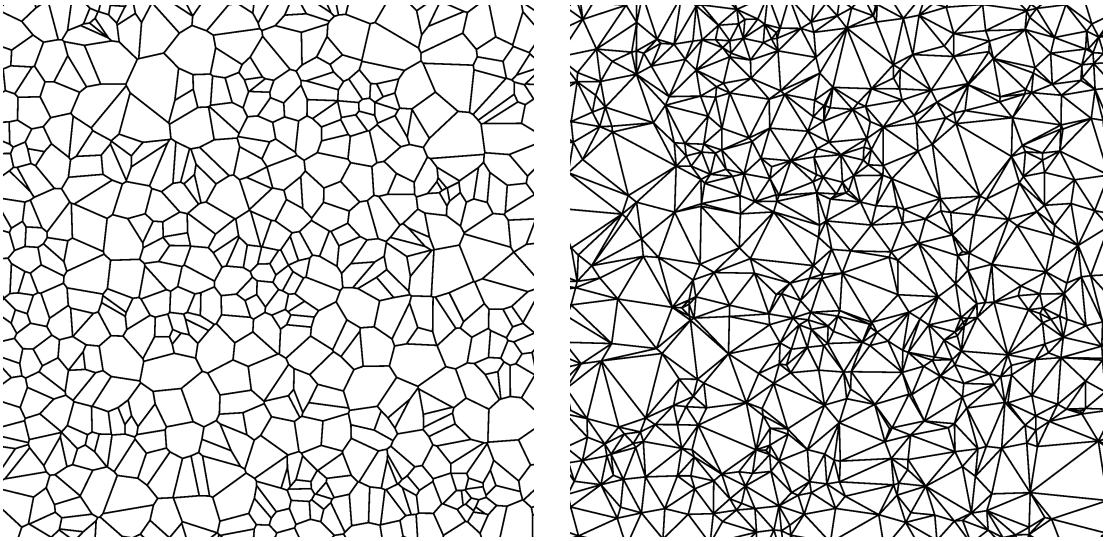

Figure 1.6.2: Realizations of the PVT (left) and PDT (right).

Second, as another example of a tessellation process with relevance to telecommunications, let us mention *Manhattan grids (MG)*, with the particular example of the *rectangular Poisson line process (RPLT)*. The RPLT consists of perpendicular lines through the points of independent Poisson point processes representing landmarks along each axis in $\mathbb{R}^2$. Despite its popularity in stochastic geometry, this model has the serious drawback of only being able to represent street systems where the distance between successive streets is exponentially distributed. This constraint can be removed by replacing the Poisson renewal process by a *stationary renewal process*, that is, by a renewal process that is statistically invariant under shifts along the axis. See Figure 1.6.3 on the left for an illustration. The MG can be further refined by putting additional rectangular lines inside the boxes given by the MG. This construction gives rise *nested Manhattan grids (NMG)*, see Figure 1.6.3 on the right for an illustration. ◇

Often, a key to the mathematical analysis of Cox point processes is their mixing properties: how strong and how far reaching are spatial stochastic dependencies? In the remainder of this section we introduce the concept of *stability* as a tool to measures these dependencies. We denote by Λ_B the restriction of a measure Λ to a set $B \subset \mathbb{R}^d$. Further, let $Q_r(x) = x + [-r/2, r/2]^d$ denote the cube with side length $r > 0$ centered at $x \in \mathbb{R}^d$ and put $Q_r = Q_r(o)$. We define $\mathrm{dist}(\varphi, \psi) = \inf\{|x - y| \colon x \in \varphi, y \in \psi\}$ for the distance between sets $\varphi, \psi \subset \mathbb{R}^d$.

Definition 1.6.6 (Stabilizing random measures)**.** *A random measure Λ on $\mathbb{R}^d$ is called* stabilizing*, if there exists a random field of stabilization radii $\mathcal{R} = \{R_x\}_{x \in \mathbb{R}^d}$, defined on the same probability space as Λ, such that*

1. *$(\Lambda, \mathcal{R})$ is jointly stationary,*

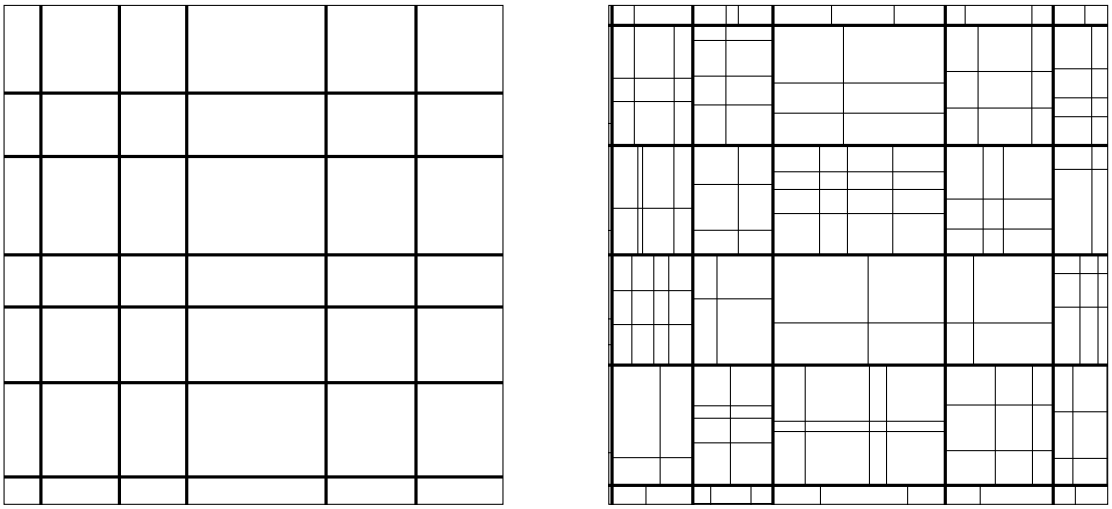

Figure 1.6.3: Realizations of the MG (left) and the NMG (right).

2. $\lim_{r\uparrow\infty} \mathbb{P}(\sup_{y\in Q_r\cap\mathbb{Q}^d} R_y < r) = 1$, *and*

3. *for all* $r \geq 1$, *the random variables*

$$\Big\{ f(\Lambda_{Q_r(x)}) \mathbb{1}\{ \sup_{y\in Q_r(x)\cap\mathbb{Q}^d} R_y < r\} \Big\}_{x\in\varphi}$$

are independent for all bounded measurable functions $f\colon \mathcal{M}(D) \to [0,\infty)$ *and all finite* $\varphi \subset \mathbb{R}^d$ *with* $\operatorname{dist}(x, \varphi\setminus\{x\}) > 3r$ *for all* $x \in \varphi$.

A strong form of stabilization is given if Λ is *b-dependent* in the sense that Λ_A and Λ_B are independent whenever $\operatorname{dist}(A,B) > b$. The two models of Example 1.6.3 are b-dependent for some b, and the random measure Λ concentrated on the Poisson-Voronoi tessellation S of Example 1.6.4 is stabilizing.

Lemma 1.6.7 (The PVT is stabilizing)**.** *The stationary PVT on* $\mathbb{R}^d$ *is stabilizing.*

Proof. The proof rests on the definition of the radius of stabilization as $R_x = \inf\{\|X_i - x\| : X_i \in \mathbb{X}\}$, for details see [CHJ17]. □

1.7 Convergence of point processes

Later, we want to discuss and analyze approximations of point processes, in order to arrive at manageable formulas for complex situations. Hence, we need to discuss also convergence issues for point measures, which we do here. See also [DVJ03, Appendix A2] and [Re87]. The basis was laid in Section 1.1, where two notions of distributions of point processes are discussed. Here we proceed by one step and provide tools for characterizing convergence. One example that we find important is the high-density limit, which we discuss at the end of this section. Here we encounter the situation that the point process converges towards some deterministic measure, i.e., not towards a point process. For this sake, we have to extend the set of random measures for which we consider convergence in distribution.

As before, we fix a measurable set $D \subset \mathbb{R}^d$. The appropriate general setting of the following is the setting where D is just some locally compact topological space, but, as in the preceding sections, we just keep this in mind and proceed with $D \subset \mathbb{R}^d$. Instead of locally finite point configurations $(x_i)_{i\in I} \in \mathbb{S}(D)$ or corresponding point measures $\sum_{i\in I} \delta_{x_i}$, we will more generally look at *Radon measures* on D, i.e., measures that assign to each compact subset of D a finite value. We want to characterize convergence of sequences of such measures in a natural way. We will do this for the two topologies that we introduced in Section 1.1, the vague and the τ-topology.

First, observe that the definition of vague and of τ-convergence of Radon measures on D directly derives from a slight extension of Definition 1.1.1, i.e., it is defined by convergence of all the test integrals against all the continuous, respectively the measurable, functions $D \to \mathbb{R}$ with compact support. A variant of the Portmanteau theorem shows that this is the same as convergence of the measures of compact subsets (whose boundary is a nullset with respect to the limiting measure, for the vague case). This makes it easy to show, e.g., that on $D = \mathbb{R}$, the measure $\frac{1}{n}\sum_{i\in\mathbb{Z}} \delta_{i/n}$ converges towards the Lebesgue measure.

Example 1.7.1. Vague and τ-convergence are local notions and say nothing about the total mass, as one sees in the examples that δ_n on $D = \mathbb{R}$ converges to the zero measure and that the measure on $E = \mathbb{R}$ with Lebesgue density $\mathbb{1}_{[-n,n]}$ converges towards the Lebesgue measure as $n \to \infty$. ◇

Remark 1.7.2 (Metrizability and measurability)**.** There is a metric on the set of Radon measures on D that induces the vague topology. Hence, the topological space of such measures is indeed a metric space. Its measurable structure is then given by the Borel σ-field. In particular, the maps $\mu \mapsto \mu(A)$ are measurable for any relative compact set $A \subset D$ with boundary a nullset; actually these maps form a basis of this Borel σ-field; i.e., it is the smallest σ-field that makes these maps measurable. ◇

Remark 1.7.3 (Relative compactness of sets of measures)**.** It can be deduced from Prohorov's theorem that a sequence $(\mu_n)_{n\in\mathbb{N}}$ of Radon measures on D is relatively compact in the vague topology (i.e., that each subsequence contains a further subsequence that vaguely converges) if and only if, for any relatively compact set $A \subset D$, the sequence $(\mu_n(A))_{n\in\mathbb{N}}$ is bounded. ◇

Now we turn to sequences of *random* point measures, i.e., sequences of random variables taking values in the set of point measures, and want to characterize possible limits. In principle, this has been settled by the preceding, since such measures are embedded in the set of Radon measures, and we established the topology of vague convergence on that set. Hence, it is, as a topological space, also a measurable space, and convergence of random variables taking values in that space is to be understood in terms of *weak convergence*, sometimes also called *convergence in distribution*. To summarize this, we denote the set of Radon measures on D by $\mathfrak{R}(D)$, then we can recast the notion of the convergence of point processes, or more generally Radon measures, as follows.

Definition 1.7.4. *A sequence $(\pi_n)_{n\in\mathbb{N}}$ of random Radon measures on D converges weakly (or converges in distribution) towards a random Radon measure π on D if, for any continuous bounded function $\Phi\colon \mathfrak{R}(D) \to \mathbb{R}$, we have $\lim_{n\to\infty} \mathbb{E}(\Phi(\pi_n)) = \mathbb{E}(\Phi(\pi))$.*

Note that the continuity of Φ refers to any of two respective topologies that we consider on the set of Radon measures, the vague and the τ-topology. In general, this characterization of

weak convergence is not very helpful, as it is *a priori* difficult to characterize all the continuous bounded functions on $\mathfrak{R}(D)$. However, if all the measures $\pi, \pi_1, \pi_2, \dots$ are point measures, then the situation is simpler:

Lemma 1.7.5 (Weak convergence of point measures in the vague- respectively τ-topology)**.** *A sequence $(\pi_n)_{n\in\mathbb{N}}$ of random point measures on D converges weakly (or in distribution) towards a random point measure π if and only if the following holds. For any $k \in \mathbb{N}$ and for all relative compact sets $A_1, \dots, A_k \subset E$ (additionally satisfying $\pi(\partial A_i) = 0$ for the vague topology) almost surely for all $i \in \{1, \dots, k\}$, the vector $(\pi_n(A_1), \dots, \pi_n(A_k))$ converges weakly towards the vector $(\pi(A_1), \dots, \pi(A_k))$ as $n \to \infty$, i.e., if and only if, for any $n_1, \dots, n_k \in \mathbb{N}_0$,*

$$\lim_{n\to\infty} \mathbb{P}\big(\pi_n(A_1) = n_1, \dots, \pi(A_k) = n_k\big) = \mathbb{P}\big(\pi(A_1) = n_1, \dots, \pi(A_k) = n_k\big).$$

We will also need convergence of Radon measures against (possibly deterministic) measures that are not point measures. For this, the possibly most handy criterion is the following. We denote the *Laplace transform* of a random measure π on D by

$$\mathcal{L}_\pi(f) = \mathbb{E}\Big[\mathrm{e}^{-\int_D f(x)\,\pi(\mathrm{d}x)}\Big].$$

This notation is slightly misleading, since $\mathcal{L}_\pi$ does not depend on π, but only on its distribution.

Lemma 1.7.6 (Convergence and Laplace transforms)**.** *A sequence $(\pi_n)_{n\in\mathbb{N}}$ of Radon measures on D converges weakly (or in distribution) towards some measure π on D if and only if the Laplace transforms converge, i.e., for any test function $f\colon D \to [0,\infty)$ with compact support (continuous for the vague topology, just measurable for the τ-topology),*

$$\lim_{n\to\infty} \mathcal{L}_{\pi_n}(f) = \mathcal{L}_\pi(f).$$

Later we will be interested in the high-density limit of a Poisson process in a compact subset of $\mathbb{R}^d$, in particular in the deviations away from the limit. Here we establish the limit itself.

Lemma 1.7.7 (Convergence of empirical measures)**.** *Let $D \subset \mathbb{R}^d$ be compact and μ a positive and finite measure with Lebesgue density on D. For $\lambda \in (0,\infty)$, let $\mathbb{X}^{(\lambda)} = (X_i^{(\lambda)})_{i\in I_\lambda}$ be a Poisson point process in D with intensity measure $\lambda\mu$. Then, as $\lambda \to \infty$, the random point measure*

$$L_\lambda = \frac{1}{\lambda} \sum_{i\in I_\lambda} \delta_{X_i^{(\lambda)}} \tag{1.7.1}$$

converges weakly (in both the vague and the τ-topology) towards μ as $\lambda \to \infty$.

Proof. According to Lemma 1.7.6, it is sufficient to check the convergence of the Laplace transform. Let $f\colon D \to [0,\infty)$ be measurable, then, according to Campbell's theorem,

$$\mathcal{L}_{L_\lambda}(f) = \mathcal{L}_{\mathbb{X}^{(\lambda)}}(f/\lambda) = \exp\Big(\int_D (\mathrm{e}^{-f(x)/\lambda} - 1)\,\lambda\mu(\mathrm{d}x)\Big). \tag{1.7.2}$$

For any x, we see that the integrand with respect to $\mu(\mathrm{d}x)$ converges towards $-f(x)$. If f is integrable with respect to μ, then we can apply the dominated convergence theorem, since $1 - \mathrm{e}^{-y} \le y$ for any $y \in \mathbb{R}$, and therefore the integrand is bounded in absolute value by $f(x)$. Hence, we see that the Laplace transform converges towards $\exp(-\int_D f(x)\,\mu(\mathrm{d}x)) = \mathcal{L}_\mu(f)$

in this case. If f is not integrable with respect to μ, then we estimate $\mathcal{L}_{L_\lambda}(f)$ first against $\mathcal{L}_{L_\lambda}(f \wedge K)$ for some cutting parameter K, derive convergence towards $\exp\{-\int f \wedge K \, \mathrm{d}\mu\}$ and use then the monotone monotone-convergence theorem for letting $K \to \infty$ to see that $\mathcal{L}_{L_\lambda}(f)$ converges towards $0 = \mathcal{L}_\mu(f)$. In both cases, we have verified the convergence of the Laplace transform for all nonnegative measurable test functions f. □

Recall the interpretation that we rely on. The points X_i are the locations of the users (or other devices) of the telecommunication system in the communication area D. The interpretation of Lemma 1.7.7 is that the dense cloud of users in D approaches the density of the intensity measure μ, i.e., a multitude of microscopic information (every single user location) is approximated by some much simpler macroscopic object, a density, for which there are good perspectives for further analysis. One might argue that such a limiting setting is useless for describing human beings, since they cannot be squeezed infinitely strongly, but we are heading for approximate formulas, and many of the situations are in reality quite well described by approximations via asymptotic formulas.

Chapter 2

Coverage and connectivity: the Boolean model

In this chapter, we discuss mathematical approaches to the two most fundamental questions about spatial telecommunication models:

- *Coverage:* How much of the area can be reached by the signals emitted from the users, respectably the base stations?
- *Connectivity:* How far can a message travel through the system in a multihop-functionality?

To do this, we introduce and study the most basic model for message transmission within a spatial system formed by a PPP $\mathbb{X} = \{X_i : i \in I\}$ of users or base stations, the *Boolean model*. In this model, which we introduce in Section 2.1, to each location X_i a random closed set Ξ_i is attached, the local communication zone that can be reached by a signal emitted from X_i. Then $\bigcup_{i\in I}(X_i + \Xi_i)$ is the *communication area*, the set of locations that can be reached by any message transmission. In Section 2.2, we study questions about the coverage of a given compact set $C \subset \mathbb{R}^d$, i.e., about the probability that C can be reached by some signal. These are local questions. In contrast, in Section 2.3, we consider global questions about whether or not the communication area possesses an unbounded connected component. This we will do only for homogeneous PPPs and only for balls Ξ_i of a given fixed radius. In this simple but fundamental setting, we will distinguish two drastically different scenarios, the occurrence versus the absence of *percolation*. The distinction is one important example of a *phase transition* and lies at the heart of a beautiful theory called *continuum percolation*. Both phases are non-trivial, as we formulate in Section 2.3. In order to carry out the proof for that in Section 2.5, we first need to rely on the discrete counterpart of the the theory, which we will prepare for in Section 2.4. Furthermore, we establish in Section 2.6 additional relevant connectivity questions related to percolation, and in Section 2.7 we discuss some peculiarities on percolation that arise for Cox point processes.

See [BB09a] and [BB09b] (which we follow in Sections 2.1 and 2.2) for application of the Boolean model to telecommunication, in particular coverage and percolation properties, and [BR06], [MR96] and [FM08] for mathematical proofs of continuum percolation properties. Standard references on the discrete part of the theory are [G89] and [BR06]. The Section 2.6 about further interpretations of the percolation probability in telecommunications is largely self

contained except for the shape theorem, see [YCG11]. The results about continuum percolation for Cox point processes presented in Section 2.7 are less standard and taken from [CHJ17].

2.1 The Boolean model

Let $\mathbb{X} = \{X_i : i \in I\}$ be a point process in $\mathbb{R}^d$ with intensity measure μ. Again, we interpret the X_i as the locations of the users or base stations of a spatial telecommunication system. Now, we extend the model by adding a random closed set $\Xi_i \subset \mathbb{R}^d$ around each user X_i and interpret $X_i + \Xi_i$ as the area that can be reached by a signal that is emitted from X_i. We call Ξ_i the local communication zone around X_i. The idea is that the strength of the signal decays quickly in the distance, and a certain least strength is necessary for a successful transmission. Typical choices for Ξ_i are centered balls with random or deterministic radius, but more complex choices are thinkable and have their right, e.g., when environmental conditions have to be taken care of. For example, if X_i is located on a street, then its local communication area Ξ_i will be shaped by the houses left and right of the location X_i and will be approached by some rectangle, depending on the location X_i. Apart from that, we will take the random sets Ξ_i, $i \in I$, as independent. Hence, we would like to see the Ξ_i as marks attached to the users X_i. Note that, under this assumption, the case of Ξ_i being given by a Poisson-Voronoi cell, see (1.6.2), is not covered.

Definition 2.1.1 (Boolean model)**.** *Let $\mathbb{X} = \{X_i : i \in I\}$ be a PPP in $\mathbb{R}^d$, and let K be a probability kernel from $\mathbb{R}^d$ to the set of closed subsets of $\mathbb{R}^d$. Consider the K-marking $\mathbb{X}_K = \sum_{i\in I} \delta_{(X_i,\Xi_i)}$ according to Definition 1.4.1). Then, the random set $\Xi_{\mathrm{BM}} = \bigcup_{i\in I}(X_i + \Xi_i)$ is called a* Boolean model*.*

We have formally taken the set of all closed subsets of $\mathbb{R}^d$ as the mark space, and there is also a natural σ-algebra on this set to turn this into a measurable space. However, this space is not a locally compact topological space, and therefore the above definition, strictly speaking, does not fall into Definition 1.4.1. However, there is no problem to concentrate the kernel on a much smaller set of closed sets, e.g., indexed by $\mathbb{R}^l$ for some set of parameters l in a natural way that turns it into a locally compact topological space. One important example is the set of centered balls (or squares, or rectangles, ...) with a random radius. We will only think of such examples in these notes.

For simplicity, we will from now consider only independent K-markings, i.e., we will assume that the random sets Ξ_i are i.i.d., not depending on the location X_i that they are attached to. That is, the kernel K is just one probability measure, which we will drop from the notation. The Boolean model Ξ_{BM} is interpreted as the total communication area, i.e., the (random) set that is covered by the signals emitted from the set $\mathbb{X}$.

2.2 Coverage properties

Let Ξ_{BM} be a Boolean model in the sense of Section 2.1, i.e., a PPP with an independent marking in a locally compact topological mark space. In this section we provide notions and methods to determine probabilities of *coverage*, i.e., events that a given set or point lies in Ξ_{BM}. That is, we look only at one single transmission step from some X_i. Mathematically, this amounts to the study of the local structure of the random set Ξ_{BM}, i.e., a local question.

By Ξ we denote a generic random closed set that we use in our marking, i.e., a random variable having the distribution K. From now on, we will not use the kernel K anymore, but will use $\mathbb{P}$ and $\mathbb{E}$ for probability and expectation with respect to Ξ. We will assume that its distribution satisfies

$$\mathbb{E}[\mu(C-\Xi)]<\infty, \qquad \text{for any compact } C\subset\mathbb{R}^d, \tag{2.2.1}$$

where $C-\Xi=\{x-y\colon x\in C, y\in\Xi\}$. For example if $C=\{1\}$ and $\Xi=[-R,R]$ with some $R\in(0,\infty)$, then $C-\Xi=[-R+1,1+R]$ is the set of user locations x such that $x+\Xi$ intersects C. Condition 2.2.1 ensures that the expected number of grains Ξ communicating with any compact C is finite. In particular, under this condition, also Ξ_{BM} is a random closed set itself.

The *capacity functional* of Ξ is defined as the function

$$T_\Xi(C)=\mathbb{P}(\Xi\cap C\neq\emptyset), \qquad C\subset\mathbb{R}^d \text{ compact}. \tag{2.2.2}$$

This function can be seen as an equivalent of the distribution function of a real random variable; actually it determines the distribution of Ξ, according to *Choquet's theorem*, see [M75].

Lemma 2.2.1. *For any compact set $C\subset\mathbb{R}^d$, the number*

$$N_{\mathbb{X}_{\mathrm{BM}}}(C)=\#\{i\in I\colon (X_i+\Xi_i)\cap C\neq\emptyset\}$$

is a Poisson random variable with parameter $\mathbb{E}[\mu(C-\Xi)]$.

Proof. Observe that the point process

$$\sum_{i\in I}\delta_{X_i}\mathbb{1}\{(X_i+\Xi_i)\cap C\neq\emptyset\}$$

is an independent thinning of $\mathbb{X}$ (recall Lemma 1.2.9) with (space-dependent) thinning probability

$$p_C(x)=\mathbb{P}((x+\Xi)\cap C\neq\emptyset)=\mathbb{P}(x\in C-\Xi).$$

In the same way as in the proof of Lemma 1.2.9, one sees that this process is a PPP with intensity measure $p_C(x)\,\mu(\mathrm{d}x)$. Furthermore, $N_{\mathbb{X}_{\mathrm{BM}}}(C)$, the total number of its points, is a Poisson random variable with parameter equal to $\int_{\mathbb{R}^d}p_C(x)\,\mu(\mathrm{d}x)$. With the help of Fubini's theorem, we identify this parameter as follows

$$\int p_C(x)\,\mu(\mathrm{d}x)=\int\mathbb{P}(x\in C-\Xi)\,\mu(\mathrm{d}x)=\mathbb{E}\Big[\int\mathbb{1}\{x\in C-\Xi\}\,\mu(\mathrm{d}x)\Big]=\mathbb{E}[\mu(C-\Xi)],$$

which ends the proof. □

Lemma 2.2.2. *The capacity functional is identified as*

$$T_\Xi(C)=1-\mathrm{e}^{-\mathbb{E}[\mu(C-\Xi)]}, \qquad C\subset\mathbb{R}^d \text{ compact}.$$

Proof. Observe that $T_\Xi(C)=\mathbb{P}(N_{\mathbb{X}_{\mathrm{BM}}}(C)>0)$ and use Lemma 2.2.1. □

From now on we restrict to the stationary (or homogeneous) Boolean model, by which we mean that the intensity measure μ of the underlying PPP is equal to λLeb for some $\lambda \in (0,\infty)$, and we call λ the intensity of the Boolean model. It is clear that then also the capacity functional of the Boolean model,

$$T_{\Xi_{\mathrm{BM}}}(C) = \mathbb{P}\Big(C \cap \bigcup_{i\in I}(X_i + \Xi_i) \neq \emptyset\Big)$$

(where the probability extends over the PPP $\mathbb{X}$ and the family of the Ξ_i's) is shift-invariant, i.e., $T_{\Xi_{\mathrm{BM}}}(z+C) = T_{\Xi_{\mathrm{BM}}}(C)$ for any $z \in \mathbb{R}^d$ and any compact set $C \subset \mathbb{R}^d$. Also the *volume fraction*

$$p = \frac{\mathbb{E}[\mathrm{Leb}(\Xi_{\mathrm{BM}} \cap B)]}{\mathrm{Leb}(B)} \tag{2.2.3}$$

does not depend on the compact set $B \subset \mathbb{R}^d$, as long as it has positive Lebesgue measure. The volume fraction has the nice interpretation as the probability that the origin is covered by the Boolean model, as

$$p = \frac{1}{\mathrm{Leb}(B)}\int_B \mathbb{E}[\mathbb{1}\{x \in \Xi_{\mathrm{BM}}\}]\,\mathrm{d}x = \mathbb{E}[\mathbb{1}\{0 \in \Xi_{\mathrm{BM}}\}] = \mathbb{P}(0 \in \Xi_{\mathrm{BM}}) = T_{\Xi_{\mathrm{BM}}}(\{0\}).$$

In particular, Lemma 2.2.2 tells us that $p = 1 - \mathrm{e}^{-\mathbb{E}[\mathrm{Leb}(\Xi)]}$.

Remark 2.2.3 (Covariance of coverage variables)**.** The volume fraction p is the expectation of the *coverage variable* at the origin, $\mathbb{1}\{0 \in \Xi_{\mathrm{BM}}\}$. The expectation of the product of the two coverage variables $\mathbb{1}\{0 \in \Xi_{\mathrm{BM}}\}$ and $\mathbb{1}\{z \in \Xi_{\mathrm{BM}}\}$ can be calculated in an elementary way as

$$\widetilde{C}(z) = \mathbb{E}\big[\mathbb{1}\{0 \in \Xi_{\mathrm{BM}}\}\,\mathbb{1}\{z \in \Xi_{\mathrm{BM}}\}\big] = \mathbb{P}(0 \text{ and } z \text{ lie in } \Xi_{\mathrm{BM}}) = 2p - 1 + (1-p)^2\mathrm{e}^{-\lambda\mathbb{E}[\mathrm{Leb}(\Xi\cap(\Xi+z))]}.$$

The function $\widetilde{C}(z)$ is usually referred to as the *covariance function* of the Boolean model. It is the probability that two points separated by the vector z are covered. The covariance of the two coverage variables is given by

$$C(z) = \widetilde{C}(z) - p^2 = -(1-p)^2(1 - \mathrm{e}^{-\lambda\mathbb{E}[\mathrm{Leb}(\Xi\cap(\Xi+z))]}).$$

$\diamond$

The *coverage probability* of a given compact set $C \subset \mathbb{R}^d$ by a random closed set Ξ is defined as $\mathbb{P}(C \subset \Xi)$. In general, it is difficult to give explicit expressions for this quantity; however it is clear that it is not larger than $T_\Xi(C)$, and we have equality for singletons C. In the literature, there are asymptotic results for the coverage probability for the homogeneous Boolean model with Ξ equal to a centered ball of radius rR, where R is a positive random variable and r is a parameter. These results are precise in the limit $\lambda \to \infty$ of a highly dense PPP and $r \downarrow 0$ of very small communication radii. We present one such result, see [J86, Lemma 7.3].

Theorem 2.2.4 (Asymptotic coverage probability)**.** *Assume that $d = 2$ and let $C \subset \mathbb{R}^2$ be a compact set whose boundary is a Lebesgue null set. Consider the Boolean model*

$$\Xi_{\mathrm{BM}} = \bigcup_{i\in I}(X_i + B_{rR}(0)),$$

where the random radius R satisfies $\mathbb{E}[R^{2+\varepsilon}] < \infty$ for some $\varepsilon > 0$. Put

$$\phi(\lambda, r) = \lambda r^2\pi\mathbb{E}[R^2] - \log\frac{\mathrm{Leb}(C)}{\pi r^2\mathbb{E}[R^2]} - 2\log\log\frac{\mathrm{Leb}(C)}{\pi r^2\mathbb{E}[R^2]} - \log\frac{\mathbb{E}[R]^2}{\mathbb{E}[R^2]}.$$

Then

$$\mathbb{P}(C \subset \Xi_{\mathrm{BM}}) = \exp\big\{ - \mathrm{e}^{-\phi(\lambda,r)}\big\} \qquad \text{as } \lambda \to \infty, r \downarrow 0,$$

provided that $\phi(\lambda, r)$ *tends to some limit in* $[-\infty, \infty]$.

One can use this result for finding, for a given compact set C, the number of Poisson points, depending on the size of the local communication balls that are needed for covering C with the communication area with a certain given positive probability. Indeed, for a given $u \in \mathbb{R}$, couple λ and r such that $\phi(\lambda(r), r) \to u$ with

$$\lambda(r) = \frac{1}{\lambda r^2 \pi \mathbb{E}[R^2]}\Big(u + \log \frac{\mathrm{Leb}(C)}{\pi r^2 \mathbb{E}[R^2]} + 2 \log\log \frac{\mathrm{Leb}(C)}{\pi r^2 \mathbb{E}[R^2]} + \log \frac{\mathbb{E}[R]^2}{\mathbb{E}[R^2]}\Big).$$

Then, the coverage probability converges in the limit $r \downarrow 0$

$$\mathbb{P}(C \subset \Xi_{\mathrm{BM}}) \to \exp\big\{ - \mathrm{e}^{-u}\big\}.$$

2.3 Long-range connectivity in the homogeneous Boolean model

In this section, we consider the question of *connectivity over long distances* in the Boolean model, i.e, the question how far a message can travel through the system if it is allowed to make an unbounded number of hops. That is, we assume that a message can hop from user to user arbitrarily often, as long as it does not leave the communication area, and we ask how long the distance is that it can travel. In other words, we consider a multi-hop functionality and use the system of users as a wireless ad hoc system, which carries the message trajectories without usage of base stations. We will consider this question only for a very special, but fundamental, version of this model: the homogeneous Boolean model on the entire space $\mathbb{R}^d$ with deterministic local communication zones that are simply balls of a fixed radius. Hence, the Boolean model has only one effective parameter left, but it will turn out that it gives rise to a beautiful mathematical theory that is called *continuum percolation*. We will encounter an interesting phase transition in this parameter: for large values, there is a possibility that the message can travel unboundedly far, and for small values its trajectory will always be bounded. Furthermore, we will be able to attack a number of further important quantities in later sections.

We assume that $\mathbb{X} = \{X_i \colon i \in I\}$ is a homogeneous PPP with intensity $\lambda \in (0, \infty)$, and the random closed set Ξ_i that we put around each user location X_i is just a deterministic closed ball $B_{R/2}(X_i)$ with a fixed radius $R \in (0, \infty)$. A message can now hop from X_i to X_j if and only if $\|X_i - X_j\| \le R$, i.e., if and only if the straight line between them entirely lies in the communication area $\bigcup_{i \in I} \mathbb{B}_{R/2}(X_i)$. This is the case if and only if the closed balls $B_{R/2}(X_i)$ and $B_{R/2}(X_j)$ intersect. Therefore, we now have basically two mathematical models that express connectivity: either we conceive $\mathbb{X}$ as a random geometric graph (the so-called *Gilbert graph*) by drawing an edge between any two points X_i and X_j with distance $\le R$, or we consider the Boolean model

$$\Xi_{\mathrm{BM}} = \bigcup_{i \in I} B_{R/2}(X_i) \tag{2.3.1}$$

and consider connectivity in the usual topological sense for subsets of $\mathbb{R}^d$. We will proceed in the latter model. Recall that the process is homogeneously distributed over $\mathbb{R}^d$, and that we have just two parameters, the intensity λ of $\mathbb{X}$ and the diameter R of the balls. We will write $\mathbb{P}_\lambda$ and

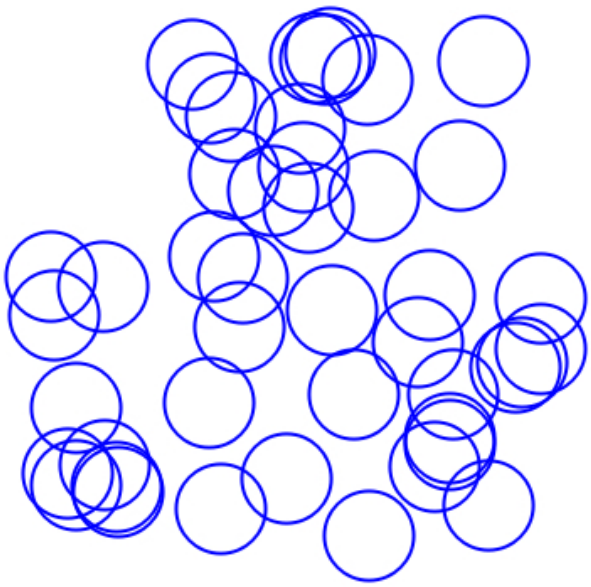

Figure 2.3.1: Realization of a homogeneous Boolean model.

$\mathbb{E}_\lambda$ for the probability and expectation in this model. In Figure 2.3.1 we present a realization of such a Boolean model.

Let us denote by $\mathcal{C}_{R/2}(x)$ the cluster (= connected component) of Ξ_{BM} that contains x. Then $\mathcal{C}_{R/2}(X_i)$ is the set of those space points in $\mathbb{R}^d$ that can be reached by a multi-hop trajectory starting at X_i through $\mathbb{X}$. One of the most decisive properties of Ξ_{BM} is whether or not it has unboundedly large components. We say that Ξ_{BM} *percolates* or that *percolation occurs* if the answer is yes. In this case, we also say that the points in the unbounded component are connected to ∞. The notion of percolation is the base of everything that follows. It gave the theory its name "continuum percolation", since it is about the continuous space $\mathbb{R}^d$. Interestingly, the first paper that introduced this model in 1961, see [G61], explicitly took wireless multi-hop communication as the prime example and motivation. There is a theory of discrete percolation (usually motivated by water leakage through porous stones), which we will encounter in Section 2.4 below.

Remark 2.3.1 (Percolation and message transmission)**.** In the event that percolation occurs, there is at least one component $\mathcal{C}$ that is unboundedly large and contains infinitely many users X_i. As a consequence, messages that are emitted from such a user can reach, at least theoretically, infinitely many other users and can travel infinitely far. Note that this does not say anything about the reach of messages that are emitted from a bounded component, which certainly also do exist. In this event, the quality of transmission service is drastically better than in the complement event, such that we are highly interested in those values of the parameters that make this possible. ◇

The most important quantity is the *percolation probability*

$$\theta(\lambda, R/2) = \mathbb{P}_\lambda(\mathrm{Leb}(\mathcal{C}_{R/2}(o)) = \infty). \tag{2.3.2}$$

Since the PPP is locally finite, this is equal to the probability that $\mathcal{C}_{R/2}(o)$ contains infinitely many points of $\mathbb{X}$.

Remark 2.3.2 (Percolation probability via Palm calculus)**.** Recall the Palm distribution P_λ^o of $\mathbb{X}$ from Section 1.5. Then, by the stationary Mecke-Slivnyak theorem 1.5.3, the percolation probability can be identified as

$$\theta(\lambda, R) = P_\lambda^o(\mathrm{Leb}(\mathcal{C}_{R/2}(o)) = \infty).$$

By the interpretation of the Palm distribution, this quantity can be also interpreted as the probability that the cluster $\mathcal{C}_R(X)$ has infinite Lebesgue measure, for a randomly chosen user $X \in \mathbb{X}$. The advantage of the Palm interpretation is that it can be easily extended to the setting of Cox point processes, see Section 2.7. ◇

Because of the homogeneity of the PPP in space (more precisely, since the intensity measure is invariant under scaling), it is clear that the influence of the parameter R can easily be extracted as follows:

$$\theta(\lambda, R) = \theta(\lambda R^d), \qquad \lambda, R \in (0, \infty), \tag{2.3.3}$$

where we write $\theta(\lambda) = \theta(\lambda, 1)$. Therefore, we will consider only $\theta(\lambda)$ in the following and will drop R from the notation. It is clear from an application of Lemma 1.2.10 that θ is an increasing function. Hence, we can define the *critical percolation threshold*

$$\lambda_{\mathrm{cr}} = \inf\{\lambda \in (0, \infty) \colon \theta(\lambda) > 0\} = \sup\{\lambda \in (0, \infty) \colon \theta(\lambda) = 0\} \in [0, \infty], \tag{2.3.4}$$

where we put $\sup \emptyset = 0$ and $\inf \emptyset = \infty$. (From (2.3.3), we obtain that $\lambda_{\mathrm{cr}}(R) = \lambda_{\mathrm{cr}} R^{-d}$ if $\lambda_{\mathrm{cr}}(R)$ denotes the critical percolation threshold for the Boolean model with interaction radius R.)

The start of the theory is the following.

Theorem 2.3.3 (The critical threshold is positive and finite)**.** *For any $d \in \mathbb{N} \setminus \{1\}$, $\lambda_{\mathrm{cr}} \in (0, \infty)$.*

The theorem says that, for some (sufficiently small) $\lambda \in (0, \infty)$, the percolation probability is zero, while for sufficiently large ones, it is positive, giving rise to non-trivial *sub-and supercritical regimes*. This is a clear manifestation of the occurrence of a *phase transition*. We give more comments in Section 2.5 below. Before we prove Theorem 2.3.3 there, we must have a closer look at the discrete version of percolation theory, as this will provide an inevitable mathematical base.

2.4 Intermezzo: phase transition in discrete percolation

As it is very often the case in the theory of point processes in continuous space, proofs rest on a comparison with the (much simpler) setting of a fixed geometry. Here, instead of a PPP on $\mathbb{R}^d$, we will consider a Bernoulli field on the edges of $\mathbb{Z}^d$. That is, we put on every edge e between neighboring sites in $\mathbb{Z}^d$ a Bernoulli random variable $\xi_e \in \{0, 1\}$ with parameter $p \in [0, 1]$, and assume that all these random variables are i.i.d.. Then $(\xi_e)_e$ is a Bernoulli field. The edge e is called *open* if $\xi_e = 1$ and *closed* otherwise. A realization of such a random field is said to exhibit percolation if there exists an infinite connected component of open edges, where connectedness is defined in the usual sense on $\mathbb{Z}^d$. In this way, we are now working on the basic model of the theory of *discrete bond percolation*, or just *percolation*. This likewise is a beautiful mathematical theory, which studies analogous questions, but is further developed than continuum percolation, due to the simpler setting of a fixed, discrete geometry. The standard reference for this theory

is [G89], but see also [BR06]. The main motivating application idea is a porous stone with hyper-cubic microstructure, one water source in the middle and walls between neighboring cells that are permeable to water with a certain probability, and then the question is whether or not the surface of the stone is wet somewhere or not.

Note that we have now just one parameter p, the probability for openness. We denote probability and expectation in this model by $\mathbb{P}_p$ and $\mathbb{E}_p$. In this model, we call two sites in $\mathbb{Z}^d$ connected if there is a sequence of open edges that form a nearest-neighbor path between the two sites. We write $C(x)$ for the cluster (= connected component) that contains $x \in \mathbb{Z}^d$. Similarly to 2.3.2 we introduce the percolation probability

$$\theta(p) = \mathbb{P}_p(\#C(o) = \infty) \tag{2.4.1}$$

and the critical percolation threshold

$$p_{\mathrm{cr}} = \inf\{p \in [0,1] \colon \theta(p) > 0\} \in [0,1]. \tag{2.4.2}$$

Here is the start of the theory of percolation, a result that will later be instrumental for proving the continuous version, Theorem 2.3.3:

Theorem 2.4.1 (Non-trivial critical threshold for Bernoulli bond percolation)**.** *For any $d \in \mathbb{N} \setminus \{1\}$, we have that $p_{\mathrm{cr}} \in (0,1)$.*

Proof. The proof naturally comes in two parts. First, we prove existence of a subcritical phase, i.e., that $p_{\mathrm{cr}} > 0$. For this, note that the event that the origin is connected to infinity is contained in the event that a self-avoiding path of open edges of a fixed length $n \in \mathbb{N}$ starts at the origin. Let Ψ_n denote the set of self-avoiding n-step paths starting at the origin, then, for any $n \in \mathbb{N}$, we see that

$$\begin{aligned}\theta(p) &\leq \mathbb{P}_p(\text{there exists } \eta \in \Psi_n \text{ such that } \xi_e = 1 \text{ for all } e \in \eta)\\ &\leq \sum_{\eta \in \Xi_n} \mathbb{P}_p(\xi_e = 1 \text{ for all } e \in \eta)\\ &\leq |\Psi_n| p^n \leq (2dp)^n.\end{aligned}$$

But, if $p < 1/(2d)$, then this quantity tends to zero as n tends to infinity and thus, $\theta(p) = 0$. Therefore, $p_{\mathrm{cr}} \geq 1/(2d)$. (If one would estimate $|\Psi_n|$ against $2d(2d-1)^{n-1}$, then even the bound $p_{\mathrm{cr}} \geq 1/(2d-1)$ would follow, which proves absence of a supercritical regime in one spatial dimension.)

The proof for the existence of a supercritical phase, i.e., that $p_{\mathrm{cr}} < 1$, is more complicated. It suffices to prove that $\theta(p) > 0$ for $p \in (0,1)$ close to 1. Note that if percolation occurs in dimension $d = 2$, then it also occurs for higher dimensions since there it is even easier to percolate. (This is an idea that is difficult to make precise for continuum percolation.) In other words, the critical threshold for percolation is a decreasing function of the dimension and it suffices to prove existence of a supercritical phase for dimension 2. This we will do now.

The strategy of the proof is an example of the famous *Peierls' argument*, which leverages the probabilistic costs of creating a blocking interface in the following sense. Consider the shifted lattice $\mathbb{Z}^2_* = \mathbb{Z}^2 + (1/2, 1/2)$ and call an edge e^* in $\mathbb{Z}^2_*$ closed if the unique edge e in $\mathbb{Z}^2$ that crosses e^* is open and *vice versa*. Now, if the origin is not connected to infinity, there must exist

a finite blocking interface of open edges in $\mathbb{Z}_*^2$ that surrounds the origin and contains one point $(n+1/2, 1/2)$ for some $n \in \mathbb{N}$. Thus, we can bound

$$\begin{aligned} 1-\theta(p) &\leq \sum_{n\in\mathbb{N}} \mathbb{P}_p\big(\text{there exists an open interface } \eta^* \subset \mathbb{Z}_*^2 \\ &\qquad\qquad \text{surrounding } (0,0) \text{ and passing } (n+1/2,1/2)\big) \\ &\leq \sum_{n\in\mathbb{N}} \mathbb{P}_p\big(\text{there exists an open path in } \mathbb{Z}_*^2 \text{ of length } 2n+4, \text{ passing } (n+1/2,1/2)\big) \\ &\leq \sum_{n\in\mathbb{N}} (4(1-p))^{2n+4}. \end{aligned}$$

The factor 4^{2n+4} is the number of paths of length $2n+4$. Now, for p close to one, this sum is strictly smaller than 1 and thus, $\theta(p) > 1$. □

Let us mention some of many interesting and important results on discrete percolation.

Remark 2.4.2 (Survey on results on discrete percolation)**.**

1. *Numerical value.* For Bernoulli bond percolation on $\mathbb{Z}^2$, the critical threshold is proven to be $p_{\mathrm{cr}} = 1/2$ based on the self-duality of $\mathbb{Z}^2$. Except for a few examples (such as the triangular lattice to mention another one), there is no formula for p_{cr}, and its numerical value can only be approached via simulations.

2. *Other criticality notions.* Critical behavior of Bernoulli percolation can also be based on related, but different, quantities other than the percolation function, e.g., the value from which on the expected size of the cluster containing the origin is infinite: $p'_{\mathrm{cr}} = \inf\{p \in [0,1] \colon \mathbb{E}_p(\#(C(o))) = \infty\}$. For Bernoulli bond percolation on $\mathbb{Z}^2$, it is known that $p'_{\mathrm{cr}} = p_{\mathrm{cr}}$.

3. *Number of infinite clusters.* The random field $\mathbb{P}_p$ is invariant (in distribution) under lattice translations and is ergodic. This implies that, in the super-critical regime $p > p_{\mathrm{cr}}$, a percolation cluster appears almost surely (i.e, $\#(C(x)) = \infty$ for some $x \in \mathbb{Z}^d$), since this event is measurable with respect to the tail-sigma-algebra. Further, it can be shown that this infinite cluster is unique almost surely.

4. *Sizes of finite clusters.* A lot of work has been dedicated to further understand the clustering behavior in the two regimes. For example, in the subcritical regime, the probability that the origin is connected to the complement of a centered box of side-length n is known to be small exponentially fast in n. This is one ingredient of the proof of $p'_{\mathrm{cr}} = p_{\mathrm{cr}}$. In a certain sense, the cardinality of any of the finite clusters (in both regimes) is known to be a random variable with exponential tails, which implies that the largest of the finite clusters in a box of radius n have about $\log n$ sites.

5. *Less independence.* Theorem 2.4.1 can be generalized with respect to the independence. E.g., if the probability p for a bond to be open is allowed to depend on neighboring bonds at distance $\leq k$ for some $k \in \mathbb{N}$, then it can be shown that there exist two critical thresholds $p_{\mathrm{cr}}^{(1)} \leq p_{\mathrm{cr}}^{(2)}$ such that below $p_{\mathrm{cr}}^{(1)}$ there is no infinite cluster almost surely and above $p_{\mathrm{cr}}^{(2)}$ there is an infinite cluster almost surely.

6. *Continuity of θ.* Another big topic in the field is to determine the continuity properties of $p \mapsto \theta(p)$. It can be shown that it is continuous in $[0,1] \setminus \{p_{\mathrm{cr}}\}$ and right-continuous at p_{cr}.

Left-continuity at p_{cr} so far could only be established for $d = 2$ and $d \geq 11$. In particular for $3 \leq d \leq 5$ it is one of the big open questions for Bernoulli bond percolation.

7. *Behavior of θ near criticality.* It is widely believed that, on general graphs in place of $\mathbb{Z}^d$, the behavior of θ close to p_{cr} is governed by a power law that depends only on some local features of the underlying geometry. More precisely, in two dimensions it is expected that $\theta(p) = (p - p_{\mathrm{cr}})^{\beta + o(1)}$ as $p \downarrow p_{\mathrm{cr}}$, with the *critical exponent* $\beta = 5/36$. This has been shown rigorously only for site percolation on the triangular lattice. ◇

In Section 2.5 it will be important for us to note that discrete Bernoulli percolation on $\mathbb{Z}^d$ or any other lattice can be also considered with openness attached to sites rather than bonds. Here it is the sites in the lattice that are independently declared open with probability $p \in [0, 1]$ (and closed otherwise), and the notion of clusters (= connected components) is even more immediate than in the bond setting. The resulting model is called *Bernoulli site percolation*. One can convince oneself that it is harder to have site-percolation than to have bond-percolation, hence the site-percolation threshold is not smaller than the bond-percolation threshold. Also for the site-version of the model, versions of Theorem 2.4.1 have been proved for various lattices, see [G89] and [BR06]. In particular, we will rely in Section 2.5 on the non-triviality of the percolation threshold for site percolation on the triangular lattice in two dimensions.

2.5 Proof of phase transition in continuum percolation

We are now going to apply the non-triviality of the percolation threshold for discrete percolation to the proof of the corresponding result in continuum, Theorem 2.3.3, which we are really after. However, it will not be Theorem 2.4.1 that we directly apply, but its version for site percolation on the triangular grid in two dimensions, see [BR06], and correspondingly we will prove Theorem 2.3.3 only for $d = 2$.

The advantage of using the triangular lattice is that, in the dual face-percolation of hexagons, neighboring hexagons always share an edge and never only a vertex, like on $\mathbb{Z}^2$. Additionally, this comparison gives better bounds for the critical threshold. Indeed, site percolation on the triangular lattice also has the critical threshold $p_{\mathrm{cr}} = 1/2$, see [BR06, Theorem 8, Chapter 5]. The comparison is done via some appropriate discretization.

Proof of Theorem 2.3.3 for $d = 2$. Again, the proof naturally comes in two parts. For both directions we will use a partition (up to boundaries of Lebesgue null sets) of $\mathbb{R}^d$ into open hexagons A_x^s of side-length $s > 0$ centered at some points $x \in \mathbb{R}^2$. By default, we assume that the origin is one of them. Note that the centers x form a triangular lattice $\mathcal{T}_s$, where the bonds are drawn across the edges of neighboring hexagons. We call the center x of one of these hexagons A_x^s *open* if it contains at least one Poisson point (i.e, at least one of the X_i's) and *closed* otherwise. Note that the probability for x to be open is given by

$$p_s = 1 - \exp(-\lambda \mathrm{Leb}(A_x^s)), \tag{2.5.1}$$

and note further that $\mathrm{Leb}(A_x^s) = 3\sqrt{3}s^2/2$. It is clear from the homogeneity and the Poissonian nature of the PPP that openness of the sites in $\mathcal{T}_s$ defines a Bernoulli field of i.i.d. random variables. Hence, $\mathcal{T}_s$ is a Bernoulli site-percolation model as we briefly discussed at the end of Section 2.4.

Now, in order to show that $\lambda_{\mathrm{cr}} < \infty$, note that any two points in neighboring hexagons have distance at most $\sqrt{13}s$. Recall that each ball around the Poisson points has diameter one. Hence, if we pick s so small that $\sqrt{13}s < 1$, then site-percolation on $\mathcal{T}_s$ implies percolation of the Boolean model $\mathbb{X}_{\mathrm{BM}}$. Hence, if λ is sufficiently large such that $p_s = 1 - \exp(-\lambda 3\sqrt{3}s^2/2)$ is larger than the critical threshold $1/2$ for site percolation on $\mathcal{T}_s$, then we have percolation of the Boolean model. Explicitly, we have

$$\lambda_{\mathrm{cr}} \leq (26 \log 2)/(3\sqrt{3}).$$

Now we show that $\lambda_{\mathrm{cr}} > 0$. Note that, if $s > 1$, percolation of the Boolean model implies site-percolation on $\mathcal{T}_s$. Hence, if λ is sufficiently small such that $p_s = 1 - \exp(-\lambda 3\sqrt{3}s^2/2)$ is smaller than the threshold $1/2$, then there is no percolation of the Boolean model. Calculating again, we arrive at

$$\lambda_{\mathrm{cr}} \geq (2 \log 2)/(3\sqrt{3}),$$

which completes the proof. □

As for discrete percolation, we give now a list of further important results, most of which are analogous, and the underlying proofs are often based on the discrete counterparts.

Remark 2.5.1 (More results on continuum percolation)**.**

1. *Numerical value.* The numerical value of the critical threshold is unknown in general. Rigorous bounds in $d = 2$ are $0.174 < \lambda_{\mathrm{cr}} < 0.843$ [MR96, 3.9], the numerical value is $\lambda_{\mathrm{cr}} \approx 0.6763475$, derived by computer simulations [QZ07].

2. *Uniqueness of infinite cluster.* As in the discrete case, the Boolean model is invariant under translations and ergodic in any dimension. Since percolation is a tail event, in the super-critical regime, a percolation cluster appears almost surely. Further, it can be shown that the infinite cluster is also unique almost surely.

3. *Size of bounded clusters.* Again, as in the discrete case, in the subcritical regime, the probability that the origin is connected to the complement of a centered box of side-length n, becomes small exponentially fast in n.

4. *Complement of communication zone.* It is equally interesting to study the vacant area $\mathbb{R}^d \setminus \Xi_{\mathrm{BM}}$. It can be shown that in the Boolean model there is at most one unbounded component in the vacant area.

5. *Random radii.* The above proof can be extended to also cover the case of a Boolean model with centered balls of random radii > 0. The results here are the following. Let R be a random variable having the distribution of the radius of one of the balls. For $d \geq 2$ a supercritical regime exists, i.e., $\lambda_{\mathrm{cr}} < \infty$. On the other hand, if $\mathbb{E}[R^{2d-1}] < \infty$, then $\lambda_{\mathrm{cr}} > 0$. For $d = 1$, if $\mathbb{E}[R] < \infty$, then $\lambda_{\mathrm{cr}} = \infty$, and if $\mathbb{E}[R] = \infty$, then $\lambda_{\mathrm{cr}} = 0$. ◇

2.6 More about the percolation probability

In this section we give some more information about the percolation probability θ. We return to the setting where the radius of the balls around the Poisson points X_i is some arbitrary, fixed $R \in (0, \infty)$, and the intensity of the PPP $\mathbb{X}$ is λ. We write $\mathbb{P}$ and $\mathbb{E}$ for probability and expectation.

Lemma 2.6.1 (Continuity of θ and exponential approach to 1 for large intensities)**.** *For any $R \in (0,\infty)$, the map $\lambda \mapsto \theta(\lambda, R)$ is continuous in $[0,\infty) \setminus \{\lambda_{\mathrm{cr}}\}$ with asymptotics*

$$\lim_{\lambda\to\infty} \frac{1}{\lambda} \log(1-\theta(\lambda,R)) = -\mathrm{Leb}(B_R(o)). \tag{2.6.1}$$

The lower bound in equation (2.6.1) is easy to understand, since $\theta(\lambda, R)$ is lower bounded by the probability that the origin is isolated (i.e., that its R-ball has no Poison points), i.e., $1-\theta(\lambda,R) \geq \exp(-\lambda \mathrm{Leb}(B_R(o)))$. For a full proof, see [P91, Corollary of Theorem 3]. Also consider [FM08, Theorem 2.6.3] for a nice introduction in two dimensions. By the scale invariance, there is also an obvious corresponding result for $R \to \infty$. In Figure 2.6.1 we present a sketch of the graph of the percolation probability.

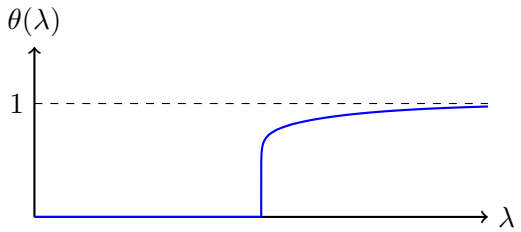

Figure 2.6.1: Approximative form of the percolation probability.

The fact that, in the super-critical regime, with probability one, there is only one infinite component (see Remark 2.5.1) allows us to represent the probability of existence of a connection between two far distant users via the percolation probability. More precisely, let

$$p_x = \mathbb{P}(o \leftrightsquigarrow x) = \mathbb{P}(\text{there exists a path in } \mathbb{X} \cup \{o,x\} \text{ connecting } o \text{ and } x) \tag{2.6.2}$$

denote the probability that the origin is connected to a point x, then we have the following result.

Theorem 2.6.2 (Two-point connectivity)**.** *For any $\lambda, R \in (0,\infty)$, we have that $\lim_{\|x\|\to\infty} p_x = \theta(\lambda,R)^2$.*

We write from now on $Q_s(x) = [x-\frac{s}{2}, x+\frac{s}{2}]^d$ for the square around $x \in \mathbb{R}^d$ with side-length s.

Proof. The main idea is that the only way that o and a distant site x can be connected is that each of them belongs to the infinite cluster, two events that depend only on (sufficiently large) neighborhoods of o respectively x. The probabilities of these two events are roughly given by the percolation probability each, and they become asymptotically independent if the distance is large. Let us give some details.

For any $y \in \mathbb{R}^d$, let E_y denote the event that there exists a path (i.e., polygon line connecting y and points of $\mathbb{X}$) in $\Xi_{\mathrm{BM}} \cup B_R(y)$, starting in $y \in \mathbb{R}^d$ and leaving the square $Q_{\|x\|/3}(y)$. Then

$$|p_x - \theta(\lambda,R)^2| \leq |p_x - \mathbb{P}(E_o \cap E_x)| + |\mathbb{P}(E_o \cap E_x) - \theta(\lambda,R)^2|. \tag{2.6.3}$$

Note that E_o and E_x are independent and have the same probability. Hence, the second summand can be bounded from above by

$$|\mathbb{P}(E_o \cap E_x) - \theta(\lambda,R)^2| \leq 2|\mathbb{P}(E_o) - \theta(\lambda,R)| = 2(\mathbb{P}(E_o) - \theta(\lambda,R)).$$

The right hand side is 2 times the probability of the origin being connected to $\mathbb{R}^d \setminus Q_{\|x\|/3}(o)$ but not to infinity. This is not larger than 2 times the probability of the existence of an sufficiently thick interface of vacant space surrounding the origin at distance at least $\|x\|/3$. In the super-critical regime, the probability for this event tends to zero (even exponentially fast) as $\|x\| \to \infty$, see [MR96, Theorem 1].

The first summand on the right-hand side of (2.6.3), since the event that $o \leftrightsquigarrow x$ is contained in the event $E_o \cap E_x$, is equal to the probability that $E_o \cap E_x$ occurs but not $\{o \leftrightsquigarrow x\}$. Note that this event is contained in the event that there exist two disjoint components in $(\Xi_{\mathrm{BM}} \cup B_R(o)) \cap Q_{\|x\|}(o)$ of diameter at least $\|x\|/6$. But, the probability for this event also tends to zero (even exponentially fast) as $\|x\| \to \infty$, see [MR96, Lemma 4.1].

In the sub-critical regime, $\theta(\lambda, R) = 0$ and $p_x \le \mathbb{P}(E_o)$, and $\mathbb{P}(E_o)$ tends to zero as $\|x\| \to \infty$ (even exponentially fast), see Remark 2.5.1 4.. □

A similar relation can be established between the percolation probability and the proportion of connected pairs of points in a growing volume. For this, let us denote by

$$\pi_s = (s^d\lambda)^{-2}\mathbb{E}\big[\#\{(X_i, X_j) \in (\mathbb{X} \cap Q_s(o))^2 \colon X_i \leftrightsquigarrow X_j\}\big] \tag{2.6.4}$$

the expected number of pairs of connected Poisson points in a square of side-length s, divided by the expected number all Poisson point pairs. Then we have the following result.

Theorem 2.6.3 (Expected number of connected pairs)**.** $\lim_{s\to\infty} \pi_s = \theta(\lambda, R)^2$.

Proof. By the Mecke-Slivnyak theorem (1.5.3),

$$\begin{aligned}
\pi_s &= (s^d\lambda)^{-2}\mathbb{E}\Big[\sum_{i\in I}\sum_{j\in I} \mathbb{1}\{X_i \leftrightsquigarrow X_j\}\mathbb{1}\{X_i, X_j \in Q_s(o)\}\Big] \\
&= s^{-2d}\lambda^{-1}\int_{Q_s(o)} \mathbb{E}\Big[\sum_{j\in I} \mathbb{1}\{x \leftrightsquigarrow X_j\}\mathbb{1}\{X_j \in Q_s(o)\}\Big]\,\mathrm{d}x + (s^d\lambda)^{-1} \\
&= s^{-2d}\int_{Q_s(o)}\int_{Q_s(o)} \mathbb{P}(x \leftrightsquigarrow y)\,\mathrm{d}x\mathrm{d}y + (s^d\lambda)^{-1} \\
&= s^{-2d}\int_{Q_s(o)}\int_{Q_s(o)} \mathbb{P}(o \leftrightsquigarrow (y-x))\,\mathrm{d}x\mathrm{d}y + (s^d\lambda)^{-1} \\
&= s^{-2d}\int_{Q_s(o)}\int_{Q_s(x)} \mathbb{P}(o \leftrightsquigarrow z)\,\mathrm{d}z\mathrm{d}x + (s^d\lambda)^{-1}.
\end{aligned}$$

Hence, $|\pi_s - \theta(\lambda, R)^2| \le s^{-2d}\int_{Q_s(o)}\int_{Q_s(x)} |\mathbb{P}(o \leftrightsquigarrow z) - \theta(\lambda, R)^2|\,\mathrm{d}x\mathrm{d}z + (s^d\lambda)^{-1}$ and by Theorem 2.6.2, there exists an $r > 0$ such that for $z \in \mathbb{R}^d \setminus Q_r(o) = (Q_r(o))^{\mathrm{c}}$ we have $|\mathbb{P}(o \leftrightsquigarrow z) - \theta(\lambda, R)^2| < \varepsilon$. But then we can estimate for sufficiently large s that

$$\begin{aligned}
|\pi_s - \theta(\lambda, R)^2| &\le s^{-2d}\int_{Q_s(o)} \mathrm{Leb}(Q_s(x) \cap Q_r(o))\mathrm{d}x + \varepsilon s^{-2d}\int_{Q_s(o)} \mathrm{Leb}(Q_s(x) \cap (Q_r(o))^c)\mathrm{d}x + \varepsilon \\
&\le s^{-d}\mathrm{Leb}(Q_r(o)) + 2\varepsilon.
\end{aligned}$$

Since ε was arbitrary, this finishes the proof. □

We have seen that connectivity characteristics of a network represented by the Boolean model can be expressed in terms of percolation probabilities. For many types of telecommunication services it is essential not only to know of the existence of a connection, but to be able to satisfy constraints on the number of hops. Let us finish this section by explaining how a hop constraint can be captured via the *stretch factor*, a fundamental characteristic in continuum percolation.

To begin with, we extend the definition of the connectivity of two points by imposing a constraint on the number of hops. For $k \in \mathbb{N}$, we write $o \overset{k}{\rightsquigarrow} x$ if o and x can be connected in $\Xi_{\mathrm{BM}} \cup B_R(o) \cup B_R(x)$ in at most k hops, and we write

$$p_{k,x} = \mathbb{P}(o \overset{k}{\rightsquigarrow} x).$$

To understand the asymptotic behavior of $p_{k,x}$ for large values of k and distant points x, we require a crucial auxiliary result known as the *shape theorem*. Let $\mathcal{C}$ denote the unique infinite connected component in the super-critical phase of percolation in the Boolean model.

Theorem 2.6.4 (Shape theorem)**.** *For $\lambda > \lambda_{\mathrm{cr}}$ there exists a deterministic and finite* stretch factor $\rho = \rho(\lambda, R)$ *such that*

$$\lim_{\substack{\|X_i - X_j\| \to \infty \\ X_i, X_j \in \mathcal{C}}} \frac{T(X_i, X_j)}{\|X_i - X_j\|} = \rho(\lambda, R),$$

where $T(x, y)$ denotes the smallest number of hops from x to y.

In words, the shape theorem states that the minimum number of hops to connect points in the unique infinite connected component grows linearly in the distance between the end points, and the proportionality factor depends only on the two parameters. Loosely speaking, the shape theorem shows that, asymptotically for large distances, the metric induced by shortest paths between points becomes simply a scalar multiple of the Euclidean metric.

The definition of $T(x, y)$ generalizes to arbitrary points in $\mathbb{R}^d$ via $T(x, y) = T(q(x), q(y))$ with the point $q(x) \in \mathbb{X}$ denoting the closest Poisson point to x that is contained in the infinite connected component. In particular, by sub-additivity, $T(x, y) \leq T(x, z) + T(z, y)$ for any $x, y, z \in \mathbb{R}^d$, so that $\mathbb{E}T(o, ne_1) \leq n\mathbb{E}T(o, e_1)$. Hence, in expectation,

$$\rho(\lambda, R) \leq \mathbb{E}T(o, e_1).$$

The proof of Theorem 2.6.4 rests on the famous *Kingman's sub-additive ergodic theorem*, see [YCG11]. With this, we can extend the representation formula in Theorem 2.6.2 to the setting of a bounded number of hops.

Theorem 2.6.5. *Let $\lambda > \lambda_{\mathrm{cr}}$ and $r > 0$ be arbitrary, then*

$$\lim_{\|x\| \to \infty} p_{r\|x\|, x} = \theta(\lambda, R)^2 \mathbb{1}\{\rho(\lambda, R) \leq r\}.$$

Sketch of proof. First, the factor θ^2 appears for the same reason as in Theorem 2.6.2: (1) distant points fare only connected if both are in the unique infinite connected component, and (2) the events of being in the infinite component are close to being independent if the considered points are far away from each other. By Theorem 2.6.4, the stretch factor converts constraints on the number of hops into a constraint on the Euclidean distance of the endpoints, thereby giving rise to the indicator on the right-hand side of the asserted limit. □

2.7 Percolation for Cox point processes

The proofs of the results about the percolation probability of the Boolean model in the preceding sections all benefit from the spatial homogeneity provided by the underlying PPP. However, some of them should hold also under weaker assumptions than independence. We have already introduced a way to quantify spatial correlations, at least for Cox point processes, via the stabilization property, see Definition 1.6.6. Let us close this chapter by presenting a version of Theorem 2.3.3 for stationary stabilizing Cox point processes.

Before we can do this, we have to properly set up the percolation probability for stationary Cox point processes with intensity $\lambda\Lambda$ for some random directing measure Λ and some intensity $\lambda \in (0,\infty)$. Recall from Section 1.5 the definition of the Palm version of a stationary Cox point process $\mathbb{X}$, i.e.,

$$E^o_\lambda[f(\cdot)] = \frac{1}{\lambda \mathbb{E}[\Lambda(Q_1(o))]} \mathbb{E}\Big[\sum_{i\in I} \mathbb{1}\{X_i \in Q_1(o)\} f(\mathbb{X} - X_i)\}\Big], \qquad f\colon \mathcal{S}(\mathbb{R}^d) \to [0,\infty) \text{ measurable.}$$

Now we work with the Boolean model for the Cox process with deterministic centered balls of radius $R \in (0,\infty)$ around each of the Poisson points. The notion of connection and clusters is the same as before, see Figure 2.7.1 for an illustration. Then the percolation probability is

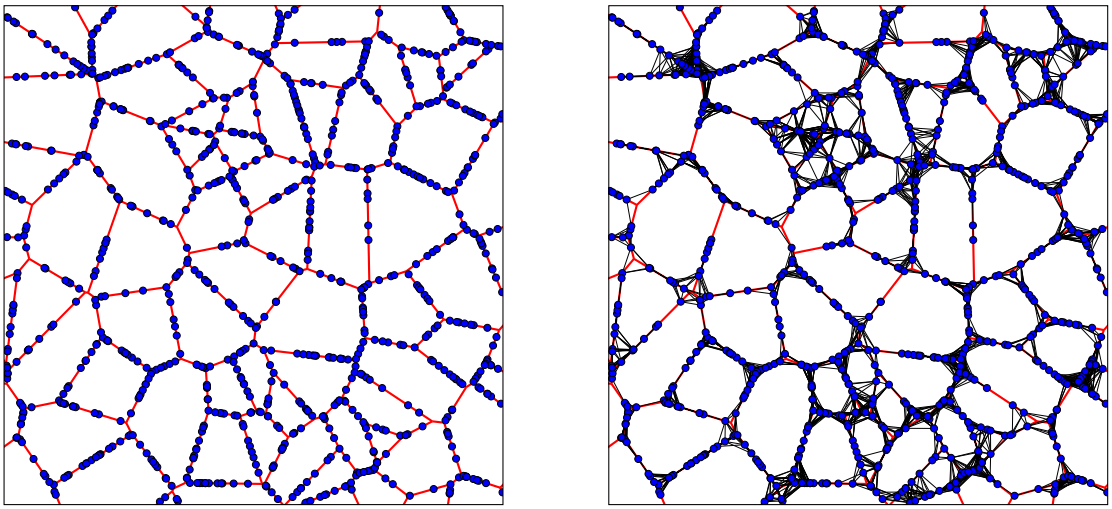

Figure 2.7.1: Illustration of the Boolean model for a Cox point process with random intensity measure based on a PVT.

defined as

$$\theta(\lambda, R) = P^o_\lambda(\mathrm{Leb}(\mathcal{C}_R(o)) = \infty)),$$

and the associated critical intensity is given by

$$\lambda_{\mathrm{cr}}(R) = \inf\{\lambda > 0\colon \theta(\lambda, R) > 0\}.$$

Note that we cannot impose any scaling invariance that would allow us to eliminate the dependence on the reach R. We have the following result on the existence of a sub-critical phase.

Theorem 2.7.1. *If the random intensity measure Λ is stabilizing, then $\lambda_{\mathrm{cr}}(R) > 0$.*

The proof again rests on a comparison with Bernoulli percolation; we do not give details. This time, the comparison is first for a dependent discrete percolation system in the sense of Remark 2.4.2 6.. In an additional step, the dependency can be stochastically bounded by an Bernoulli site percolation with sufficiently low openness probability.

Note that, in order to establish existence of a supercritical regime, it is not sufficient to assume stabilization for the environment. Trivially, the environment may not put positive weight on a connected region in space and consequently, even if λ is very large, the emergence of a percolation cluster is impossible. This scenario can be avoided by putting another condition on the environment which ensures enough connectivity. Recall the notation $Q_s = [-s/2, s/2]^d$.

Definition 2.7.2 (Connectedness of random measures)**.** *A stabilizing random measure Λ with stabilization radii R_x, $x \in \mathbb{R}^d$, is called* asymptotically essentially connected *if for all $n \geq 1$, whenever $\sup_{y \in Q_{2n(o)}} R_y < n/2$, we have that*

1. *$supp(\Lambda_{Q_n(o)}) \neq \emptyset$ and*
2. *$supp(\Lambda_{Q_n(o)})$ is contained in a connected component of $supp(\Lambda_{Q_{2n}(o)})$.*

The Poisson-Voronoi tessellation S of Example 1.6.4 does have this property.

Lemma 2.7.3 (The PVT is asymptotically essentially connected)**.** *The stationary PVT on $\mathbb{R}^d$ is asymptotically essentially connected.*

Proof. The proof rests again on the definition of the radius of stabilization as $R_x = \inf\{\|X_i - x\| : X_i \in \mathbb{X}\}$, for details see [CHJ17]. □

Now we can state our result about the existence of a super-critical phase for Cox point processes.

Theorem 2.7.4. *If the random intensity measure Λ is asymptotically essentially connected, then $\lambda_{\mathrm{cr}}(R) < \infty$.*

Again the proof works via a discretization and comparison with a dependent Bernoulli percolation process, which can then be dominated by a super-critical independent Bernoulli percolation process.

Chapter 3

Interference: the signal-to-interference ratio

Let us proceed with our first steps in the application to a spatial random telecommunication network. Our basic model is a PPP $\mathbb{X} = \{X_i : i \in I\}$ in $\mathbb{R}^d$, which is the set of locations of the users (or devices, or base stations, ...) of the system. We imagine that, at the time that we consider, from each of the locations a signal is emitted. Hence, there are possibly a number of signals floating around the area. As a consequence, it may very well be that a given signal cannot be successfully transmitted, since it is suppressed by too much noise. This effect is called *interference*. In this section, we discuss one of the most widely used mathematical models for describing interference, and show how to handle some of the most relevant properties. In Section 3.1, we introduce the basic tools for describing the signal strengths mathematically, and in Section 3.2, we introduce and discuss a criterion for successful transmission in terms of the *signal-to-interference ratio*. In Section 3.3 we present some results about the percolation properties of the variant of the Boolean model where edges are drawn based on the signal-to-interference ratio.

3.1 Describing interference

The signals are propagated instantly and isotropically (i.e., radially symmetrically) into all space directions from the point of origin. However, there are fading effects by acoustic obstacles like the medium or by more concrete ones like trees, fences, houses and so on, which make the signal strength smaller and smaller as the distance to the transmission site increases. This fading is typically expressed in two ways. First, in terms of a (random) mark at the transmission site, accounting for the smaller non distant-based obstruction of the signal. Secondly, the fading of the signals because of distance, is expressed in terms of a function $\ell : (0, \infty) \to [0, \infty)$, the *path-loss function*, i.e., $\ell(r)$ is the strength of the signal in a distance r from the transmitter. This function should be decreasing with $\lim_{r\to\infty} \ell(r) = 0$ and $\lim_{r\downarrow 0} \ell(r) = 1$, since the received strength cannot be larger than the emitted strength. Let us assume, that the transmission power of each signal is identical, say one.

Very important is the kind of decay of $\ell(r)$ for large r. The general ansatz is that it should decay like a power of r, say $\ell(r) \approx r^{-\alpha}$ as $r \to \infty$, for some $\alpha > 0$. The parameter α models the average path-loss in the medium that we consider: the more acoustic obstacles are there (e.g.,

in areas with a high density of houses or walls), the larger α should be picked. Often, one just determines that $\alpha > d$ to ensure integrability, see below. Therefore, typical choices of ℓ are

$$\ell(r) = r^{-\alpha} \qquad \text{or} \qquad \ell(r) = (1+r)^{-\alpha} \qquad \text{or} \qquad \ell(r) = \min\{1, r^{-\alpha}\}. \tag{3.1.1}$$

Note that the first choice, $\ell(r) = r^{-\alpha}$, is for our application not senseful, since it explodes for small distances, but on the other side it has nice mathematical properties, for example a perfect-scaling property, such that many quantities of interest can explicitly be calculated.

Let us assume that at some site $y \in \mathbb{R}^d$ one receiver is located. The total amount of signals that (s)he receives is equal to

$$I_{\mathbb{X}}(y) = \sum_{i \in I} \ell(\|X_i - y\|), \qquad y \in \mathbb{R}^d, \tag{3.1.2}$$

which is called the *interference* at y. Note that the interference at the origin, $I_{\mathbb{X}} = I_{\mathbb{X}}(o)$, is equal to $S_{\ell \circ \|\cdot\|}(\mathbb{X}) = \sum_{i \in I} \ell(\|X_i\|)$, as defined in (1.1.2). Let us have a look at its expectation.

Example 3.1.1 (Mean interference at the origin)**.** How large is the expected sum of the signals that one receiver at the origin experiences? Campbell's formula from Theorem 1.3.1 gives

$$\mathbb{E}[I_{\mathbb{X}}] = \mathbb{E}[S_{\ell \circ \|\cdot\|}(\mathbb{X})] = \int_{\mathbb{R}^d} \ell(\|y\|)\, \mu(\mathrm{d}y),$$

where μ is the intensity measure of the point process $\mathbb{X}$. Let us study the important special case of a standard PPP $\mathbb{X}$ with intensity λ, and let us pick one of the two bounded choices of ℓ in (3.1.1). (Indeed, one can easily see that the perfect-scaling choice $\ell(r) = r^{-\alpha}$ leads to an integral that diverges for any value of α – either at ∞ or at zero). Hence, the finiteness of the integral exclusively hinges at the integrability of the integral of $\ell(\|\cdot\|)$ at infinity. The asymptotics for $y \to \infty$ are $\sim \|y\|^{-\alpha}$, hence we see that

$$\mathbb{E}[I_{\mathbb{X}}] < \infty \qquad \Longleftrightarrow \qquad \alpha > d.$$

This belongs to the reasons that often we will pick $\alpha > d$ or make the assumption that $y \mapsto \ell(\|y\|)$ is integrable with respect to μ: the interference is an integrable random variable. For both above choices of ℓ it is possible without difficulties to calculate the expectation of the interference. Campbell's formula helps us also to find its variance explicitly, but for its finiteness we even need to assume that $\alpha > 2d$. ◇

Example 3.1.2 (The distribution of the interference)**.** Let us describe a standard procedure for describing the distribution of the interference at zero, $I_{\mathbb{X}}$, for a PPP $\mathbb{X}$ with general intensity measure μ and general path-loss function ℓ. The two basic ingredients of this procedure are (1) an approximation with the interference coming from a large ball, and (2) decomposition according to the number of users in that ball. However, we will see that an explicit identification of the distribution is possible only in very particular cases.

Let $I_{\mathbb{X}}[B] = \sum_{i \in I:\, X_i \in B} \ell(\|X_i\|)$ denote the interference at zero coming from all users in a set B, and recall that $N_{\mathbb{X}}(B)$, the number of users in B, is Poisson-distributed with parameter $\mu(B)$. Furthermore, recall from Lemma 1.2.4 that, given the event $\{N_{\mathbb{X}}(B) = k\}$, the k users in B are independent and identically distributed sites with distribution $\mu(\mathrm{d}x)/\mu(B)$, restricted to

B. We denote by B_a the centered ball around the origin with radius a. Let us try to calculate the distribution function of $I_{\mathbb{X}}$. For any $t \in [0,\infty)$, we have

$$\begin{aligned}\mathbb{P}(I_{\mathbb{X}} \le t) &= \lim_{a\to\infty} \mathbb{P}(I_{\mathbb{X}}[B_a] \le t) = \lim_{a\to\infty} \sum_{k\in\mathbb{N}_0} \mathrm{e}^{-\mu(B_a)} \frac{\mu(B_a)^k}{k!} \mathbb{P}(I_{\mathbb{X}}[B_a] \le t \mid N_{\mathbb{X}}(B_a) = k) \\ &= \lim_{a\to\infty} \mathrm{e}^{-\mu(B_a)} \sum_{k\in\mathbb{N}_0} \frac{\mu(B_a)^k}{k!} \mathbb{P}_a\Big(\sum_{j=1}^{k} L_j \le t\Big),\end{aligned} \tag{3.1.3}$$

where $L_1, \dots, L_k$ are k i.i.d. $(0,\infty)$-valued random variables with the distribution of $\ell(\|Y\|)$. If $Y \in B_a$ is distributed according to $\mu(\mathrm{d}x)/\mu(B_a)$ that is,

$$\mathbb{P}_a(L_1 \le t) = \frac{1}{\mu(B_a)} \int_{B_a} \mathbb{1}\{\ell(\|y\|) \le t\}\, \mu(\mathrm{d}y), \qquad t \in (0,\infty).$$

It is not straight-forward to evaluate the right-hand side of (3.1.3), but such problems are ubiquitous. The last term is the distribution function of a sum of k i.i.d. random variables and hence it looks promising to proceed instead with either the Fourier transform (characteristic function) or with the Laplace transform, since in both cases the probability term is decomposed in a product of k identical things, i.e., in a k-th power. In both cases, the k-sum can therefore be evaluated with the help of the exponential series. Let us demonstrate this for the Laplace transform: for any $s > 0$, we have

$$\begin{aligned}\mathbb{E}[\mathrm{e}^{-sI_{\mathbb{X}}}] &= \lim_{a\to\infty} \mathrm{e}^{-\mu(B_a)} \sum_{k\in\mathbb{N}_0} \frac{\mu(B_a)^k}{k!} \mathbb{E}_a[\mathrm{e}^{-s\sum_{j=1}^{k} L_j}] \\ &= \lim_{a\to\infty} \mathrm{e}^{-\mu(B_a)} \sum_{k\in\mathbb{N}_0} \frac{\mu(B_a)^k}{k!} \mathbb{E}_a[\mathrm{e}^{-sL_1}]^k \\ &= \lim_{a\to\infty} \mathrm{e}^{-\mu(B_a)} \exp\big\{\mu(B_a)\mathbb{E}_a[\mathrm{e}^{-sL_1}]\big\} \\ &= \lim_{a\to\infty} \exp\Big\{\int_{B_a} \mathrm{e}^{-s\ell(\|y\|)}\, \mu(\mathrm{d}y) - \mu(B_a)\Big\} \\ &= \exp\Big\{\int_{\mathbb{R}^d} \big[\mathrm{e}^{-s\ell(\|y\|)} - 1\big]\, \mu(\mathrm{d}y)\Big\}.\end{aligned}$$

(This calculation is an alternate proof of Campbell's theorem 1.3.1 for Laplace transforms.)

A similar calculation applies to Fourier transforms. Indeed, for any $\omega \in \mathbb{R}$, we see that

$$\mathcal{F}_{I_{\mathbb{X}}}(\omega) = \mathbb{E}[\mathrm{e}^{\mathrm{i}\omega I_{\mathbb{X}}}] = \exp\Big\{\int_{\mathbb{R}^d} \big[\mathrm{e}^{\mathrm{i}\omega\ell(\|y\|)} - 1\big]\, \mu(\mathrm{d}y)\Big\}.$$

Now, in order to identify the distribution of $I_{\mathbb{X}}$, one has to find formulas for the inversion of the respective transforms.

An explicit evaluation of this integral seems possible only for very particular choices, for example μ the Lebesgue measure and $\ell(r) = r^{-\alpha}$. It is not clear how much one can learn form such model calculations, as the interference drastically depends on the choice of the path-loss function and on the users that are located close to the origin, and they may produce effects that are typically unwanted and unrealistic. ◇

Example 3.1.3 (Interference with individual transmission powers)**.** Let us consider the case where each user X_i emits a signal with an individual transmission strength $P_i \in (0,\infty)$. We assume that, given I, the PPP $\mathbb{X} = \{X_i \colon i \in I\}$ and the collection $P = (P_i)_{i\in I}$ of strengths are independent, and that $(P_i)_{i\in I}$ is an i.i.d. collection. Then $\{(X_i, P_i) \colon i \in I\}$ is a marked PPP. The interference at zero is now the random variable $I_{(\mathbb{X},P)} = \sum_{i\in I} P_i \ell(\|X_i\|)$. Its Laplace transform is calculated as

$$\mathbb{E}[\mathrm{e}^{-sI_{(\mathbb{X},P)}}] = \exp\Big\{ \int_{\mathbb{R}^d} \int_{(0,\infty)} \big[\mathrm{e}^{-sp\ell(\|y\|)} - 1\big]\, K(\mathrm{d}p)\mu(\mathrm{d}y)\Big\},$$

where K is the distribution of a transmission strength. In the special case $\mu = \lambda\mathrm{Leb}$ and $\ell(r) = r^{-\alpha}$, one can calculate in an elementary way that

$$\mathbb{E}[\mathrm{e}^{-sI_{(\mathbb{X},P)}}] = \exp\Big\{\lambda c_d \int_{(0,\infty)} (ps)^{\alpha/d}\, K(\mathrm{d}p)\Gamma(1-\tfrac{\alpha}{d})\Big\}$$

where Γ denotes the distribution function of the Gamma distribution. ◇

3.2 The signal-to-interference ratio

The interference $I_{\mathbb{X}}(y)$ at a given site y from transmissions from all the users $X_i \in \mathbb{X}$ is a measure of the total amount of noise that is received from all the messages that are present. Actually, the receiver at y is interested to successfully understand a signal that comes from a particular user X_i. The interference makes it hard that (s)he is successful. A precise mathematical criterion for this success is given in terms of the *signal-to-noise-and-interference ratio (SINR)*,

$$\mathrm{SINR}(X_i, y, \mathbb{X}) = \frac{P_i\ell(\|X_i - y\|)}{N + \sum_{j\in I\setminus\{i\}} P_j\ell(\|X_j - y\|)}, \qquad y \in \mathbb{R}. \tag{3.2.1}$$

where $N \in (0,\infty)$ is the general *noise* in the system, $P_i \in (0,\infty)$ is the signal power of the i-th user, and ℓ is a path-loss function as in Section 3.1. In words, $\mathrm{SINR}(X_i, y, \mathbb{X})$ is the quotient of the (wanted) signal strength that is received at y from the user X_i and the (unwanted) total sum of the basic noise and all the other signal strengths transmitted from all the other users. If $y = X_k$ for some $X_k \in \mathbb{X}$, then in the interference term also the signal from X_j to X_k is neglected, i.e., we consider $\sum_{j\in I\setminus\{i,k\}} P_j\ell(\|X_j - X_k\|)$.

If we want to neglect the general noise, then we put $N = 0$ and call the quantity in (3.2.1) the *signal-to-interference ratio (SIR)* and write it as $\mathrm{SIR}(X_i, y, \mathbb{X})$. Certainly, we can further simplify by putting $P_i = 1$ for any i. Both the noise N and the signal strengths P_i can be taken random.

The criterion is then formulated as:

$$\text{The person at } y \text{ can detect the signal from } X_i \quad \Longleftrightarrow \quad \mathrm{SINR}(X_i, y, \mathbb{X}) \geq \tau, \tag{3.2.2}$$

where $\tau \in (0,\infty)$ is a technical constant, which measures the fineness of the ability to filter the wanted signal from the unwanted ones. The area of sites that can detect the signal emitted from a given Poisson site $X_i \in \mathbb{X}$,

$$\mathcal{C}_{X_i} = \{y \in \mathbb{R}^d \colon \mathrm{SINR}(X_i, y, \mathbb{X}) \geq \tau\} \tag{3.2.3}$$

is called its *SINR cell*. The union of these cells,

$$\Xi_{\mathrm{SINR}} = \bigcup_{i\in I} \mathcal{C}_{X_i}, \tag{3.2.4}$$

is called the *SINR model* similar to the Boolean model in (2.3.1). We define the *SINR coverage probability*

$$p_o(y) = P^o(y \in \mathcal{C}_o) = T_{\Xi_{\mathrm{SINR}}}(\{y\}), \qquad y \in \mathbb{R}^d, \tag{3.2.5}$$

as the probability (under the Palm measure) that a given site can be reached by a signal emitted from the origin; recall the capacity functional introduced in (2.2.2).

Remark 3.2.1 (Calculating the coverage probability)**.** Note that this probability can be evaluated using that

$$p_o(y) = \mathbb{P}\Big(P_o \ell(\|y\|) \geq \tau\Big(N + \sum_{j\in I} P_j \ell(\|X_j - y\|)\Big)\Big).$$

The probability involves the PPP $\mathbb{X}$, the random noise N and the i.i.d. signal strength variables P_i; we assume that all these quantities are independent. It is difficult to obtain more explicit formulas in general; [BB09a] derives formulas involving Fourier transforms of the interference and of the signal strength variables. Here is an approach for the case that the signal strengths P_i are exponentially distributed with parameter c. This puts us in the lucky situation that we can employ the Laplace transforms of N and of the interference $I_{(\mathbb{X},P)}(y) = \sum_{j\in I} P_j \ell(\|X_j - y\|)$ as follows.

$$\begin{aligned}
p_o(y) &= \int_0^\infty \mathbb{P}(P \geq \tau s/\ell(\|y\|))\, \mathbb{P}(N + I_{(\mathbb{X},P)}(y) \in \mathrm{d}s) \\
&= \int_0^\infty \exp\{-c\tau s/\ell(\|y\|)\}\, \mathbb{P}(N + I_{(\mathbb{X},P)}(y) \in \mathrm{d}s) \\
&= \mathcal{L}_{N+I_{(\mathbb{X},P)}(y)}(c\tau/\ell(\|y\|)) \\
&= \mathcal{L}_N(c\tau/\ell(\|y\|))\, \mathcal{L}_{I_{(\mathbb{X},P)}(y)}(c\tau/\ell(\|y\|)),
\end{aligned}$$

where we used that the Laplace transform of a sum of two independent random variables is equal to the product of the two Laplace transforms. For calculating the transform of $I_{(\mathbb{X},P)}(y)$, one might use Example 3.1.3. In the simple case where $N = 0$ and $\ell(r) = r^{-\alpha}$ for some $\alpha \in (0, \infty)$, one can derive explicit results. ◇

Remark 3.2.2 (Avoiding interference)**.** In our interpretation, the sum over i in the sum in the denominator of the SIR in 3.2.1 actually extends over the set of those users who transmit precisely at the time that we consider. In a more complex model that takes a time development into consideration, certainly this set can vary from time to time. Even more, if many users desire to transmit messages, then it will be absolutely necessary to distribute the times of transmissions in such a way that, at each of the time instances, only a carefully chosen part of the messages is sent out, such that the sum in the denominator is not too large. A strategy that guarantees this will possibly allow only a certain density of user to transmit at a given time instance. Further, since we are looking at a multi-hop system, each message will need to be assigned several such hopping times, until it finally arrives at the intended receiver.

Finding or describing algorithms to achieve this belong to the theory of discrete optimization, but finding the abstract possibility to do this and proving bounds for the number of message transmissions per time unit is an exciting application of percolation theory. ◇

3.3 SINR percolation

Given the point cloud $\mathbb{X}$, the SINR gives us a more refined and more realistic rule to build a random graph with vertices given by $\mathbb{X}$. The classical Boolean model, where an edge is drawn between two points based on their mutual distance, can now be replaced by the *SINR model*, where a directed edge is drawn from X_i to X_j in $\mathbb{X}$ if $\mathrm{SINR}(X_i, X_j, \mathbb{X}) \geq \tau$. In many systems, one considers a message transmission $X_i \to X_j$ successful only if also a confirmation message $X_j \to X_i$ is successfully transmitted. Hence, it makes good sense to introduce symmetry in the bonds and consider the undirected graph having an edge between X_i and X_j in $\mathbb{X}$ if both $\mathrm{SINR}(X_i, X_j, \mathbb{X}) \geq \tau$ and $\mathrm{SINR}(X_j, X_i, \mathbb{X}) \geq \tau$.

Let us introduce an additional parameter $\gamma > 0$, which allows us to tune the interference and write

$$\mathrm{SINR}_\gamma(X_i, y, \mathbb{X}) = \frac{P_i \ell(\|X_i - y\|)}{N + \gamma \sum_{j \in I \setminus \{i\}} P_j \ell(\|X_j - y\|)}, \qquad y \in \mathbb{R}. \tag{3.3.1}$$

We call the corresponding graph the *SINR$_\gamma$ model*.

Remark 3.3.1 (Relative versus total interference)**.** The exclusion of the transmitter X_i in the interference term in the denominator in (3.3.1) is a standard model assumption. However, under a suitable change of the connectivity parameter τ, the reduced interference $\sum_{j \in I \setminus \{i\}} P_j \ell(\|X_j - y\|)$ can be replaced by the total interference $I_{(\mathbb{X},P)}(y)$. Indeed, the condition $\mathrm{SINR}_\gamma(X_i, y, \mathbb{X}) \geq \tau$ is equivalent to

$$P_i \ell(\|X_i - y\|) \geq \tau \Big(N + \gamma \sum_{j \in I \setminus \{i\}} P_j \ell(\|X_j - y\|) \Big) \quad = \tau \Big(N + \gamma I_{(\mathbb{X},P)}(y) - \gamma P_i \ell(\|X_i - y\|) \Big),$$

which is equivalent to $P_i \ell(\|X_i - y\|) \geq \frac{\tau}{1+\tau\gamma}(N + \gamma I_{(\mathbb{X},P)}(y))$. ◇

In contrast with the Boolean model, the SINR_γ model has far-reaching correlations, and one might think that its mathematical treatment should be much more difficult than the one of the Boolean model. This is in general true, but with respect to the degrees of its nodes, the SINR_γ model has surprisingly a rather simple property. Indeed, while the vertices in the Boolean model in an infinite space have unbounded degrees (in the sense that, with probability one, there exists an infinite sequence of users $(X_{i_n})_{n \in \mathbb{N}}$ such that each X_{i_n} has at least n neighbors), in the SINR_γ model this cannot happen.

Lemma 3.3.2 (The SINR_γ model has bounded degree)**.** *Let $\gamma > 0$, then in the SINR$_\gamma$ model, with probability one, each node has at most $1 + 1/(\gamma\tau)$ neighbors.*

Proof. Let X_0 be any user in $\mathbb{X}$ and denote by N_0 the number of its neighbors. If $N_0 \leq 1$, there is nothing to show. If $N_0 > 1$, then with probability one, there exists a neighbor of X_0 with smallest signal power, i.e., there exists $X_1 \in \mathbb{X}$ be such that for all $i = 2, \ldots, N_0$

$$P_1 \ell(\|X_1 - X_0\|) \leq P_1 \ell(\|X_i - X_0\|).$$

Then,

$$\begin{aligned} P_1 \ell(\|X_1 - X_0\|) &\geq \tau N + \tau\gamma \sum_{j \in I \setminus \{0,1\}} P_j \ell(\|X_j - X_0\|) \\ &\geq \tau\gamma (N_0 - 1) P_1 \ell(\|X_1 - X_0\|) \end{aligned}$$

and thus $N_0 \leq 1 + 1/(\gamma\tau)$. □

In particular, if $\gamma > 1/\tau$, then each node has at most one neighbor and the SINR_γ model has components of size at most two.

Remark 3.3.3 (SINR_γ model versus Boolean model)**.** 1. For $\gamma = 0$, the interference is neglected and the SINR_γ model becomes a Boolean model, where two users are connected independently from all the other users. Assuming that all users transmit at some fixed maximum power $P \in (0, \infty)$, the (random) radius of interaction is then given by

$$R = \sup\{r \geq 0 \colon \ell(r) \geq \tau N/P\}.$$

More specifically, for $\ell(r) = r^{-\alpha}$ we have that $R = (P/(\tau N))^{1/\alpha}$. Recalling the statement 5. in Remark 2.5.1, under some assumptions on the distribution of N, there exists a non-trivial $\lambda^*_{\mathrm{cr}} \in (0, \infty)$ marking the phase-transition point which separates a sub- and supercritical regime of percolation.

2. Note that for any $\gamma > 0$, realization-wise, the SINR_γ model is contained in the SINR_0 model since interference can only decrease the connectivity. Hence, for $\lambda < \lambda^*_{\mathrm{cr}}$, the SINR_γ model is subcritical.

3. For $\lambda > \lambda^*_{\mathrm{cr}}$ we know that

   (a) the SINR_0 model is supercritical and

   (b) for $\gamma > 1/\tau$, by Lemma 3.3.2, the SINR_γ model is subcritical.

   Hence, there must exist a critical $0 \leq \gamma_{\mathrm{cr}}(\lambda) \leq 1/\tau$ at which the SINR_γ model has a phase transition of percolation. ◇

The following theorem makes statements about $\gamma_{\mathrm{cr}}(\lambda)$ under rather strong model assumptions. A proof can be found in [DBT05].

Theorem 3.3.4. *Let $d = 2$ and assume $P_i = P$ and N to be fixed positive and finite numbers. Further assume that there exist $0 < \delta < \beta$ and $M > \tau N/P$ such that the pathloss function satisfies $\ell(r) = 0$ for all $r > \beta$ and $\tau N/P < \ell(r) < M$ for all $r < \delta$. Then, there exists $\lambda' < \infty$ and a function $\gamma_{\mathrm{cr}} \colon [\lambda', \infty) \to (0, \infty)$ such that, for $\lambda > \lambda'$ and $\gamma < \gamma_{\mathrm{cr}}(\lambda)$ there exists almost surely an infinite component in the SINR_γ model.*

A more complete statement for the existence of a super-critical phase in the presence of interference is given in [DFMMT06]. Let us stress that, with interference, the percolation probability is not a monotone function of the intensity any more. In fact, for a sufficiently large intensity, the connectivity of the model begins to decrease due to the effect of interference. The general picture is summarized in Figure 3.3.1.

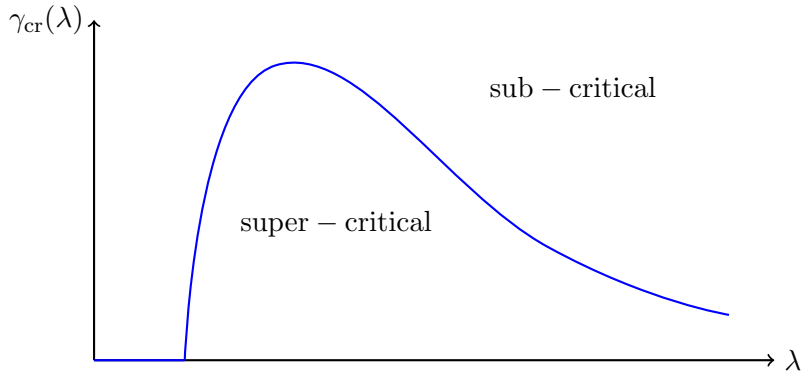

Figure 3.3.1: Sketch of the phase diagram for percolation of the SINR_γ model.

Chapter 4

Events of bad quality of service: large deviations

For measuring the quality of a random telecommunication system, it is certainly important to quantify its expected performance, i.e., the quality of service in a normal situation. However, it appears equally important to know also something about very bad situations, i.e., about random occurrences of events of a bad service, like an event of the form that only a small percentage of users actually are connected. For such unwanted, but hopefully rare, events, it is desirable to know (1) good upper bounds for the probability for this to happen, and (2) the characteristics of the situation that typically leads to this unlikely situation.

Mathematical tools for deriving such information are provided by the *theory of (the probabilities of) large deviations*, at least in situations where a certain parameter diverges that drives the unlikeliness of the event. If the theory of large deviations is applicable, it turns out that the probability under consideration decays even exponentially fast in the parameter, and the exponential rate is characterized in terms of a variational formula, which is amenable for a deeper investigation. In the connection with telecommunication, we mainly think here of the *high-density limit* $\lambda \to \infty$, where the presence of too many user in a bounded area lets the interference explode; see the end of Section 1.7, in particular Lemma 1.7.7 and Section 3.2.

In this chapter, we explain what large-deviations theory is and what it can achieve; see [DZ10] for a general account and [RS15] for an introduction to the theory with statistical physics flavor. Roughly speaking, it provides tools for expressing, for random events $A_1, A_2, A_3, \dots$ whose probabilities converge to zero, the exponential rate of this decay $\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(A_n)$, if it exists. In Section 4.1 we explain a typical example, in Section 4.2 we give a crashcourse on the theory, in Section 4.3 we proceed with the main example that we have in mind and give the large-deviation principle in the high-density situation, and in Section 4.4 we show an application to frustration events coming from a lack of connectivity due to too much interference.

4.1 Introductory example

The prototypical example is in terms of a random walk $S_n = X_1 + \cdots + X_n$, the partial sum of independent and identically distributed (i.i.d.) random variables X_i. For definiteness, assume that the X_i have expectation equal to zero and additionally very strong integrability properties, more precisely, have finite exponential moments of all orders. Then, according to the law of large

numbers, S_n/n converges to zero in probability, i.e., the probability of the event $\{|S_n/n| \geq x\}$ converges towards zero as $n \to \infty$ for any $x > 0$. Such an event is the prototypic example of an *event of a large deviation*. We want to find very good upper estimates for its probability, first for the upwards deviation $\{S_n \geq xn\}$.

With the help of the Markov inequality, one derives the following upper bound. For any $y > 0$, we have

$$\begin{aligned} \mathbb{P}(S_n \geq xn) &= \mathbb{P}(\mathrm{e}^{yS_n} \geq \mathrm{e}^{yxn}) \leq \mathrm{e}^{-yxn}\mathbb{E}[\mathrm{e}^{yS_n}] = \mathrm{e}^{-yxn}\mathbb{E}\Big[\prod_{i=1}^{n} \mathrm{e}^{yX_i}\Big] = \mathrm{e}^{-yxn}\mathbb{E}[\mathrm{e}^{yX_1}]^n \\ &= \Big(\mathrm{e}^{-yx}\mathbb{E}[\mathrm{e}^{yX_1}]\Big)^n. \end{aligned} \tag{4.1.1}$$

Note that we used that all the exponential (here: the positive) moments of X_1 are finite. The technique that we demonstrated in (4.1.1) is sometimes called the *exponential Chebyshev inequality*. Optimizing over y, we derive an exponentially decaying upper bound, which we may summarize as

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq xn) \leq -I(x), \qquad x \in (0, \infty), \tag{4.1.2}$$

where the *rate function* I is given as $I(x) = \sup_{y>0}[yx - \log \mathbb{E}(\mathrm{e}^{yX_1})]$. When turning to lower bounds, it turns out that the upper bound $I(x)$ is already extremely good and cannot be improved on the exponential scale. Indeed, somewhat deeper techniques (see Remark 4.1.1) show that in (4.1.2) also the opposite inequality holds and that a version of I for negative x holds as well. This may be loosely summarized by saying that

$$\mathbb{P}(S_n \approx xn) \approx \mathrm{e}^{-nI(x)}, \qquad n \to \infty, \tag{4.1.3}$$

for any $x \in \mathbb{R}$, with rate function

$$I(x) = \sup_{y\in\mathbb{R}}[yx - \log \mathbb{E}(\mathrm{e}^{yX_1})], \qquad x \in \mathbb{R}, \tag{4.1.4}$$

which is the Legendre transform of the map $y \mapsto \log \mathbb{E}(\mathrm{e}^{yX_1})$. One says that $(S_n/n)_{n\in\mathbb{N}}$ *satisfies a large-deviations principle (LDP)* with rate function I.

It is important to note that the formula (4.1.4) for the exponential rate is explicit and amenable to further analysis; it contains useful and characteristic information about the way how the large-deviations are typically realized. The theories of convex functions and variational calculus are helpful here.

Remark 4.1.1 (Lower bound)**.** The standard method to prove the lower bound in (4.1.3) is via the *Cramér transform*, which is an exponentially transformed probability distribution of the step variables X_i that drives the random walk to adopt the deviation event as its typical behavior. This transform $\widehat{\mathbb{P}}_a$ has a parameter $a \in \mathbb{R}$ and is defined by using the Radon-Nikodym density $\frac{1}{Z_a}\mathrm{e}^{aX_1}$ with respect to the distribution of X_1, where $Z_a = \mathbb{E}[\mathrm{e}^{aX_1}]$ is the normalizing constant. Hence, for any measurable set $A \subset \mathbb{R}$,

$$\widehat{\mathbb{P}}_a(X_1 \in A) = \frac{1}{Z_a}\mathbb{E}\big[\mathrm{e}^{aX_1}\mathbb{1}\{X_1 \in A\}\big]. \tag{4.1.5}$$

Now we fix $x \in \mathbb{R}$ and derive the lower bound in (4.1.3). It is elementary to calculate the expectation of X_1 under $\widehat{\mathbb{P}}_a$ explicitly in terms of the moment generating function $\varphi_a(x) =$

$\widehat{\mathbb{E}}_a[\mathrm{e}^{xX_1}]$ and to pick a in such a way that $\widehat{\mathbb{E}}_a(X_1) = x$. Indeed, it turns out that a is characterized as the unique number that realizes the maximum over y in (4.1.4). Now we rewrite the probability on the left in terms of this transformed measure:

$$\mathbb{P}(S_n \approx xn) = Z_a^n \widehat{\mathbb{E}}_a\big[\mathrm{e}^{-aS_n} \mathbb{1}\{S_n \approx xn\}\big] \approx Z_a^n \mathrm{e}^{-axn} \widehat{\mathbb{P}}_a(S_n \approx xn).$$

For the last term, if the event $\{S_n \approx xn\}$ is defined in a suitable way, e.g., as $\{|S_n/n - x| \le x_n\}$ for some $x_n \to 0$ (not too quickly), one can apply well-known asymptotics in the vein of the law or large numbers, to see that its exponential rate is zero. Hence, the exponential rate of $\mathbb{P}(S_n \approx xn)$ is shown to be equal to $-[ax - \log Z_a]$. Using the above mentioned characterization of a, we see that this is equal to $-I(x)$, and the proof of (4.1.3) is finished. ◇

4.2 Principles of large deviations

A proper formulation of a large-deviations principle is in terms of the weak convergence of the set function $\frac{1}{n}\log \mathbb{P}(S_n/n \in \cdot)$ towards the set function $-\inf\{I(x)\colon x \in \cdot\}$, i.e., in terms of upper bounds for closed sets and a lower bound for open sets.

Definition 4.2.1 (Large-deviations principle)**.** *Let $\mathcal{X}$ be a topological space and $(X_n)_{n\in\mathbb{N}}$ be a sequence of $\mathcal{X}$-valued random variables. Furthermore, let $I\colon \mathcal{X} \to [0,\infty]$ be a lower semicontinuous function. We say that $(X_n)_{n\in\mathbb{N}}$ (or equivalently, its distributions)* satisfy a large-deviations principle (LDP) *with speed n and* rate function I*, if, for any open set $G \subset \mathcal{X}$ and for any closed set $F \subset \mathcal{X}$,*

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(X_n \in G) \ge -\inf_{x\in G} I(x) \quad \text{and} \quad \limsup_{n\to\infty} \frac{1}{n}\log \mathbb{P}(X_n \in F) \le -\inf_{x\in F} I(x). \tag{4.2.1}$$

Hence, topology plays an important role in an LDP.

Example 4.2.2 (Cramér's theorem)**.** The introductory example of Section 4.1 is called *Cramér's theorem*. It states that, if a sequence of i.i.d. centered real-valued random variables $(X_i)_{i\in\mathbb{N}}$ is given, having all exponential moments finite, then the sequence of the empirical means $\frac{1}{n}(X_1 + \cdots + X_n) = \frac{1}{n}S_n$ satisfies an LDP with speed n and rate function I given by (4.1.4), the Legendre transform of the map $y \mapsto \log \mathbb{E}(\mathrm{e}^{yX_1})$. The function I is convex and non-negative and possesses the expected value of X_1, zero, as its only zero. The proof comes in several steps:

1. The proof of the upper bound for sets of the form $F = [x,\infty)$ with $x > 0$ is in (4.1.1).
2. Sets of the form $(-\infty, -x]$ are handled in the same way.
3. The proof of the corresponding lower bound is outlined in Remark 4.1.1.
4. General open, respectively closed, sets are handled by using that I is strictly increasing in $[0,\infty)$ respectively strictly decreasing in $(-\infty, 0]$. ◇

One useful tool says that continuous maps turn LDPs into new LDPs.

Theorem 4.2.3 (Contraction principle)**.** *If $(X_n)_{n\in\mathbb{N}}$ satisfies an LDP with rate function I in the topological state space $\mathcal{X}$, and if $F\colon \mathcal{X} \to \mathcal{Y}$ is a continuous map into another topological space, then also $(F(X_n))_{n\in\mathbb{N}}$ satisfies an LDP, and the rate function $J\colon \mathcal{Y} \to [0,\infty]$ is given by*

$$J(y) = \inf\{I(x)\colon x \in \mathcal{X}, F(x) = y\}, \qquad y \in \mathcal{Y}.$$

One of the cornerstones of the theory tells how to evaluate the exponential rate of expectations of an exponential function of variables that satisfy an LDP. We call a rate function $I\colon \mathcal{X} \to [0,\infty]$ *good* if its level sets $\{x \in \mathcal{X}\colon I(x) \le \alpha\}$ are compact.

Theorem 4.2.4 (Varadhan's lemma)**.** *If $(X_n)_{n\in\mathbb{N}}$ satisfies an LDP with good rate function I in the topological state space $\mathcal{X}$, and if $f\colon \mathcal{X} \to \mathbb{R}$ is continuous and bounded, then*

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{E}[\mathrm{e}^{nf(X_n)}] = \sup_{x\in\mathcal{X}} (f(x) - I(x)).$$

This is a substantial extension of the well-known *Laplace principle* that says that $\int_0^1 \mathrm{e}^{nf(x)}\,\mathrm{d}x$ behaves to first order like $\mathrm{e}^{n \max_{[0,1]} f}$ if $f\colon [0,1] \to \mathbb{R}$ is continuous.

4.3 LDP in the high-density setting

Now let us consider the limiting situation of importance for modeling large networks that we briefly mentioned at the end of Section 1.7, the high-density limit. Let a bounded communication area D and a Poisson point process $\mathbb{X}^\lambda = (X_i)_{i\in I_\lambda}$ in D with absolutely continuous intensity measure $\lambda\mu$ be given. Here λ is a positive parameter, which will be sent to infinity. According to Lemma 1.7.7, the *normalized empirical measure* $L_\lambda = \lambda^{-1} \sum_{i\in I_\lambda} \delta_{X_i}$ converges towards the intensity measure. That is, $L_\lambda \Longrightarrow \mu$ as $\lambda \to \infty$ in the τ-topology, i.e., when testing against measurable and bounded functions.

Now let us consider the large deviations of the empirical measure L_λ.

Lemma 4.3.1 (LDP for L_λ)**.** *Let $\mathbb{X}^\lambda = (X_i)_{i\in I_\lambda}$ be a PPP in a compact set $D \subset \mathbb{R}^d$ with intensity measure $\lambda\mu$, where $\lambda \in (0,\infty)$ and μ is an absolutely continuous measure on D. Then, the normalized empirical measure $L_\lambda = \lambda^{-1} \sum_{i\in I_\lambda} \delta_{X_i}$ satisfies, as $\lambda \to \infty$, an LDP on the set of measures on D with rate function given by*

$$I(m) = H(m|\mu) = \int f(x) \log f(x) \mu(\mathrm{d}x) - m(D) + \mu(D), \tag{4.3.1}$$

if the density $\mathrm{d}m/\mathrm{d}\mu = f$ exists, and $H(m|\mu) = \infty$ otherwise.

The term $H(m|\mu)$ is called the *Kullback-Leibler divergence* or *relative entropy* of m with respect to the reference measure μ.

Proof. Let us give an argument why this LDP should be true, at least in the weak topology. Pick some measure m on D with density with respect to μ. We want to heuristically evaluate the probability of the event $\{L_\lambda \approx m\}$. We approximate this by picking a decomposition of D into many small measurable sets $A_1, \dots, A_n$ with positive Lebesgue measure. Then we find, using the independence of $L_\lambda(A_1), \dots, L_\lambda(A_n)$ and the Poisson nature of $\mathbb{X}^\lambda$,

$$\begin{aligned}
\mathbb{P}(L_\lambda \approx m) &\approx \mathbb{P}\big(L_\lambda(A_k) \approx m(A_k)\, \forall k = 1, \dots, n\big) \\
&= \prod_{k=1}^n \mathbb{P}\Big(\sum_{i\in I_\lambda} \delta_{X_i}(A_k) \approx \lambda m(A_k)\Big) \\
&= \prod_{k=1}^n \Big[\frac{(\lambda\mu(A_k))^{\lambda m(A_k)}}{(\lambda m(A_k))!}\, \mathrm{e}^{-\lambda\mu(A_k)}\Big].
\end{aligned}$$

Now use the rough form of Stirling's formula, $n! = (\frac{n}{\mathrm{e}})^n \mathrm{e}^{o(n)}$, to see that

$$\begin{aligned}\mathbb{P}(L_\lambda \approx m) &\approx \prod_{k=1}^{n} \Big[\Big(\frac{\lambda\mu(A_k)\mathrm{e}}{\lambda m(A_k)} \Big)^{\lambda m(A_k)} \mathrm{e}^{-\lambda\mu(A_k)} \Big] \\ &\approx \exp\Big(-\lambda \sum_{k=1}^{n} \Big[m(A_k) \log \frac{m(A_k)}{\mu(A_k)} + \mu(A_k) - m(A_k) \Big] \Big) \\ &= \exp\Big(-\lambda H(m^{(n)} | \mu^{(n)}) \Big),\end{aligned}$$

where $m^{(n)}$ and $\mu^{(n)}$ are the coarsening projections of m and μ on the decomposition $(A_1, \ldots, A_n)$ of D. It is an exercise to see that their entropy converges towards the entropy of m and μ in the limit $n \to \infty$ of zero fineness of the decomposition. □

4.4 An application to user connectivity

Now let us demonstrate an example of what assertions can be deduced from the LDP of Lemma 4.3.1. This example is taken from [HJKP18]. We will focus on the frustration event of too many users being unable to send their messages to a single base station placed at the center of the compact communication area D. The network is assumed to carry a relaying functionality, that is, messages do not have to be delivered to the base station directly but can also use one intermediate relaying step. We call a user frustrated if any possible message route from him is blocked due to too low SIR along the message trajectory. We are working with the SIR introduced in Section 3.2 (with zero noise, $N = 0$), but we add a factor of $\gamma = 1/\lambda$ in front of the interference, in order to cope with the exploding amount of interference coming from an exploding number of users. (This is of course only a mathematical trick in order not to have to introduce time and to divide the message emissions on many time instances, as proposed in Remark 3.2.2). Hence, we can rewrite the SIR as a functional of the empirical measure of the PPP of users as $\mathrm{SIR}(X_i, x, L_\lambda)$, where for any measure ν on D, we write

$$\mathrm{SIR}(X_i, x, \nu) = \frac{\ell(|X_i - x|)}{\nu[\ell(|\cdot - x|)]},$$

and we write $\nu[f]$ for the integral of a function f with respect to ν. Note also that, as explained in Remark 3.3.1, there is no loss of generality in using the total interference.

Now we want to allow each message from some X_i to the base station at the origin o to make one direct step or at most one relaying step into some relay X_j. We require that each of the two steps $X_i \to X_j$ and $X_j \to o$ has to satisfy the interference condition that the SIR is not smaller than a given threshold $\tau \in (0, \infty)$. Let us write this in terms of the minimum SIR in a trajectory $x \to y \to o$

$$\mathrm{D}(x, y, o, \nu) = \min\{\mathrm{SIR}(x, y, \nu), \mathrm{SIR}(y, o, \nu)\}.$$

The maximum over the two trajectories $x \to o$ and $x \to y \to o$ for some y is then given by

$$\mathrm{R}(x, o, \nu) = \max\Big\{ \mathrm{SIR}(x, o, \nu), \max_{y \in X^\lambda} \mathrm{D}(x, y, o, \nu) \Big\}.$$

We will render the transmission of a message from X_i to o successful if $\mathrm{R}(X_i, o, L_\lambda) \geq \tau$ for a given threshold $\tau \in (0, \infty)$, i.e., if either the SIR of the direct or both SIRs of the two hops

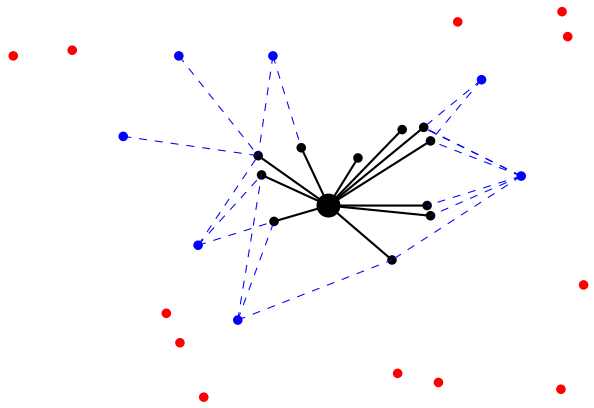

Figure 4.4.1: Realization of user configuration with black users directly linked to the origin, blue users indirectly connected and red users disconnected.

of some indirect two-hop link from X_i to o are larger than τ. In Figure 4.4.1 we present a realization of the network indicating direct and indirect uplinks.

We are interested in the large deviation behavior of the empirical measure of frustrated users, that is, of those ones whose message do not reach o. Their empirical measure may be written in terms of the function

$$\varphi_{\nu,\tau}(x) = \mathbb{1}\{\mathrm{R}(x, o, \nu) < \tau\}, \qquad x \in D,$$

as the measure with density $\varphi_{L_\lambda,\tau}$ with respect to L_λ, that is,

$$\mathfrak{M}_{L_\lambda}(\mathrm{d}x) = \varphi_{L_\lambda,\tau}(x)\, L_\lambda(\mathrm{d}x) = \frac{1}{\lambda} \sum_{i \in I_\lambda} \delta_{X_i}(\mathrm{d}x) \mathbb{1}\{\mathrm{R}(X_i, o, L_\lambda) < \tau\}.$$

Note that L_λ appears here at two places: as the ground measure and as inducing interference; a general definition of the measure $\mathfrak{M}_\nu$ for arbitrary measures ν is obvious. Examples of events that we could handle now are of the form

$$\{\mathfrak{M}_{L_\lambda}(D) - \mathfrak{M}_\mu(D) > \epsilon\},$$

In words, the event that the proportion of disconnected users is by $\epsilon > 0$ higher than expected. Now using the LDP of Lemma 4.3.1, one obtains that

$$\lim_{\lambda \uparrow \infty} \frac{1}{\lambda} \log \mathbb{P}(\mathfrak{M}_{L_\lambda}(D) > b) = -\inf\{H(\nu|\mu) \colon \mathfrak{M}_\nu(D) > b\}, \qquad b \in (0, \infty). \tag{4.4.1}$$

(Actually, it is not possible to apply Lemma 4.3.1 directly, since the event $\{\mathfrak{M}_{L_\lambda}(D) > b\}$ is not closed, and the map $\nu \mapsto \mathfrak{M}_\nu(D)$ is not continuous, but these are only technical constraints.)

The minimizers ν of $\inf\{H(\nu|\mu) \colon \mathfrak{M}_\nu(D) > b\}$ describe the typical behavior of the system conditioned on the atypical event $\{\mathfrak{M}_{L_\lambda}(D) > b\}$. It can be shown that for this event, in a domain D given by a centered disk, any minimizer must be rotationally invariant if $\mu = \mathrm{Leb}$ if only a direct uplink is allowed. A plot of the resulting radial density is given in Figure 4.4.2 in case of a disk with radius 5. It can be clearly seen that an event of too many disconnected users is typically achieved by putting slightly more users at the cell boundary. Even more prominently, it is entropically beneficial to also put more users next to the base station at the center. Those

users create a higher then expected interference, leading to more users being disconnected at the cell boundary.

In general however, finding more explicit properties of the minimizers is a hard analytic task and requires additional research.

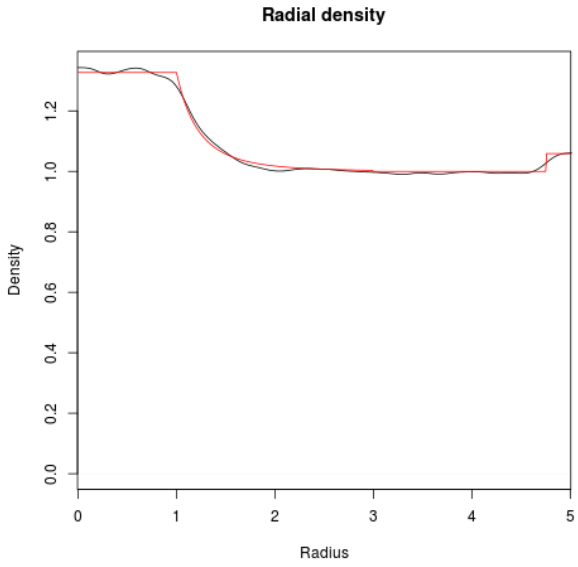

Figure 4.4.2: Plot of the spatial intensity as a function of the radius in black based on simulations. The corresponding analytic approximation result is shown in red.

Chapter 5

Random malware propagation: the contact process

In contrast to the preceding chapters, where we considered the propagation of *wanted* messages over the telecommunication system, we now turn to the propagation of *unwanted* ones, e.g., of malware. We introduce, on the set $\mathbb{X}$ of all the device locations, a Markovian time-dependent model of the set of *infected* devices at a given time. The main random mechanism that drives this infection has two elements: (1) each user that has an infected neighbor is also infected after an exponentially distributed random time, and (2) any infected device undergoes a spontaneous healing after another exponential random time.

This is a well-known mechanism of a Markovian random process on a state space of the form $E^{\mathbb{X}}$, the *contact process*; every device X_i has two possible states, infected or susceptible. The contact process is a prominent object in the theory of *interacting particle systems (IPS)*. In a nutshell, it is the theory of continuous-time Markov jump processes on discrete configuration spaces of the form $E^{\mathbb{X}}$, where $\mathbb{X}$ carries some neighboring structure, see [L85] for the standard text book on the subject. The contact process has not yet been investigated, to the best of our knowledge, on a PPP $\mathbb{X}$, but if the locations of the devices are replaced by the fixed sites in the standard grid $\mathbb{Z}^d$, then there are many precise results in the literature.

The spontaneous healing can be seen as a counter measure that the operator has installed in the system. However, we consider such a model not as the most suitable one for describing this; rather we would like to introduce a third possible state in the model: to be healthy and immune for later attacks by the malware. This seems to lead to an IPS that has not yet been introduced to the literature yet.

In Section 5.1 we introduce continuous-time Markov chains on finite sets and explain how jump rates characterize the process. The contact process is introduced in Section 5.2, and in Section 5.3 we summarize some of its most important properties in the simpler situation on the grid $\mathbb{Z}^d$ in place of $\mathbb{X}$. Finally, in Section 5.4 we give an outlook on related models that might be useful in the modeling of propagation of malware in telecommunication systems and still are waiting for their exploration.

5.1 Markov chains in continuous time

Let us explain how a Markov chain in continuous time is characterized and constructed generally in terms of waiting times and jump rates. We do this here only for a finite set Ω as the state space. Let $(\omega_t)_{t\in[0,\infty)}$ denote a time-homogeneous Markov chain in continuous time on Ω with starting site $\omega_0 = a$. Such a process must be necessarily a *jump process*. Let $\tau_1 = \inf\{t > 0 \colon \omega_t \neq a\}$ denote the first jump time, i.e., the first time at which the particle leaves its initial state. Then, by the Markov property and time-homogeneity,

$$\mathbb{P}^a(\tau_1 > t + h | \tau_1 > t) = \mathbb{P}^a(\tau_1 > t + h | \omega_t = a) = \mathbb{P}^a(\tau_1 > h), \qquad t, h \in (0, \infty).$$

Hence, $\mathbb{P}^a(\tau_1 > t+h) = \mathbb{P}^a(\tau_1 > h)\mathbb{P}^a(\tau_1 > t)$, but this is a functional equation that can only be solved by the exponential function. Thus, $\mathbb{P}^a(\tau_1 > t) = \exp(-\lambda_a t)$, for some $\lambda_a \in (0, \infty)$, which means that τ_1 must be exponentially distributed. The parameter λ_a can be recovered via

$$-\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} \mathbb{P}^a(\tau_1 > h) = \lambda_a.$$

The decision of the process at the jump times is described by a stochastic matrix $p = (p_{a,b})_{i,j\in\Omega}$, where $p_{a,b} \in [0, 1]$ is the probability to jump from state a to state b. We can neglect $p_{a,a}$ and assume that $p_{a,b}$ be positive only for $a \neq b$. This is simply the transition matrix of a discrete-time Markov chain that never stands still. But then, for $b \neq a$ we have that $\mathbb{P}^a(\omega_{\tau_1}) = b) = p_{a,b}$. We introduce the *jump rates* of the process as

$$c(a, b) = -\frac{\mathrm{d}}{\mathrm{d}h}\Big|_{h=0} \mathbb{P}^a(\omega_h = b) = \lambda_a p_{a,b} \in [0, \infty). \tag{5.1.1}$$

The collection of all the jump rates is called the *transfer matrix*, *Q-matrix* or *generator* of the process

$$L = (c(a, b))_{a,b\in\Omega}.$$

Certainly, the further propagation of the Markov process X is done (in accordance with the Markov property) by iteration of the above described procedure, starting at time τ_1 with the process at its new location after one jump.

The transfer matrix, together with the initial distribution, characterizes the distribution of the process completely. The expected waiting time for a jump is given by $\lambda_a = \sum_{b\in\Omega\setminus\{a\}} c(a, b)$ and by definition $c(a, a) = -\lambda_a$. In particular, the row sums of L are all equal to zero.

A common way to describe the jump mechanism is in terms of the asymptotics

$$\mathbb{P}^\nu(\omega_{t+h} = b \mid \omega_t = a) = 1 - c(a, b)h + o(h), \qquad h \downarrow 0,$$

for initial distribution ν on Ω and transfer matrix L.

The paths of the Markov chain $(\omega_t)_{t\in[0,\infty)}$ are right-continuous everywhere and constant strictly between the jump times and have left limits at the jump times. In other words, the paths lie in the space of *càdlàg paths*

$$D[0, \infty) = \{f \colon [0, \infty) \to \Omega \colon f \text{ is right-continuous and has left-limits everywhere}\}.$$

To fix some measurability structure on $D[0, \infty)$, we at least want to be able to evaluate the process at any given time on the Borel-sigma-algebra on Ω, i.e., we equip $D[0, \infty)$ with the

sigma algebra generated by the events $\{\omega_t \in A\}$ for all $t \geq 0$ and $A \in \mathcal{B}(\Omega)$, which also provides us with a filtration.

In the simplest case where $\Omega = \mathbb{N}$, $\omega_0 = 0$ and $c(a,b) = q(\mathbb{1}\{b = a+1\} - \mathbb{1}\{b = a\})$ for all $a \in \mathbb{N}_0$, the process of jump times is a PPP on $[0,\infty)$ with intensity q. The Markov process $(\omega_t)_{t\in[0,\infty)}$ increases by one after independent exponential times with parameter q. Such a process can be used as the counting process for another (time-homogeneous) Markov chain in continuous time; it 'rings' at all the times at which that process must make a jump.

5.2 The contact process

In this section, we consider a particular continuous-time Markov chain with relevance for the random propagation of malware over a system of devices, the *contact process*. We first make some general remarks on interacting particle systems as continuous-time Markov chains with infinite state space.

Let us keep a locally finite point configuration $x = \{x_i \colon i \in I\}$ fixed, which can now be countably infinite. We assign to each x_i a certain property $\omega_t(x_i) \in E$ at time t, where E is a finite set of all the states in which the users x_i can be. Rates $c(\zeta,\xi)$ are then *a priori* defined for any pair of states $\zeta, \xi \in \Omega$. There is no problem to construct a Markov process, as long as the set x of devices is finite. However, if x is infinite, some serious mathematical problems in the construction arise, since the state space $\Omega = E^x$ is then even uncountably infinite. The transition rates can not be simply collected in a transfer matrix anymore, which necessitates the use of Markov generators and their associated semigroups, a theory which we will not enter here. The rates can still be extracted from the distribution via (5.1.1). In order to go the other way around, namely to construct a process from a given set of rates, one has to employ deep results from functional analysis, in particular the Hille-Yoshida theorem, for details see [L85]. Existence of a process can be guaranteed if the rates satisfy a number of conditions, for example that $\sup_\zeta \sum_{\xi\in\Omega} c(\zeta,\xi) < \infty$, which prevents the process from having waiting-time distributions with infinite parameter. Let us note that many IPS can also be constructed by a graphical representation without the use of generator theory.

In the cases that we want to consider here, we keep things much simpler (and more natural) and requires that the only steps that can be made at a given time are the flips of any of the states of one x_i. Hence, the only non-zero rates $c(\zeta,\xi)$ are the ones for ζ and ξ that are distinct in just one site x_i. This is already an ad-hoc definition of an *interacting particle system (IPS)* in the sense of [L85] (even though the term 'interacting particle system' is sometimes understood in a much wider sense or is used for certain processes of a rather different nature).

One simple way to determine rates for the flip of the state of a given x_i from $a \in E$ to $b \in E \setminus \{a\}$ is to fix it as a constant, not depending on anything else. This is a good choice for describing a spontaneous healing of an infected device, when a is the infected state and b a healthy state. However, we also want to describe more interesting mechanisms, where an infection comes from neighbors. The best well-known and perhaps most studied process that describes this is the *contact process*, which we introduce now.

The contact process is defined on $\Omega = E^x$ for $E = \{0,1\}$ ('1' standing for 'infected' and '0' for 'healthy') and a general countably-infinite site space $x = \{x_i \colon i \in I\}$ with some neighborhood structure. More precisely, we need a notion of two sites being directly connected, usually represented by an edge between the points. Think for example of the Boolean model, i.e., the in-

fection can jump from device to device if their spatial distance is smaller than a given threshold. More refined models of the underlying graph can also be considered such as the SINR model introduced in Section 3.3.

Then, the contact process is characterized by jump rates on Ω of the form $c(\zeta, \zeta^i)$, where the configuration ζ^i equals ζ except at the site x_i where $\zeta^i_{x_i} = 1 - \zeta_{x_i}$. Then, for all $x_i \in x$,

$$c(\zeta, \zeta^i) = \begin{cases} 1, & \text{if } \zeta_{x_i} = 1, \\ \lambda \sum_{j \in I \colon x_j \sim x_i} \mathbb{1}\{\zeta_{x_j} = 1\}, & \text{if } \zeta_{x_i} = 0, \end{cases} \tag{5.2.1}$$

and $\sim$ denotes the neighborhood relation. All other rates $c(\zeta, \xi)$ are equal to zero, i.e., only single-site states are changed in a jump. An infected site becomes spontaneously healthy with fixed rate one, independent of all the other sites. On the other hand, a healthy site becomes infected with a rate that is λ times the number of infected neighbors. The parameter $\lambda > 0$ in the model allows us to tune the strength of the infection. The model has a natural interpretation as to represent a random spread of a disease with an additional self healing component. Note that a healthy site can again be infected later; there is no possibility of immunization in this model.

The contact process defined by the rates in (5.2.1) does belong to the class of models for which an infinite-space version is well-defined at least for situations where the underlying graph is a lattice or regular fixed geometry.

5.3 The contact process on $\mathbb{Z}^d$

The contact process, although appearing innocent, shows many interesting features already on the lattice $\mathbb{Z}^d$ with the usual neighborhood structure. Most prominently, it has a phase transition with respect to its class of invariant measures.

Let us first present some general notions for IPSs. We denote by $\mathbb{P}^\nu$ the measure under which the IPS is defined with the initial distribution ν on Ω.

Definition 5.3.1 (Invariant measures)**.** *A probability measure ν on Ω is called* invariant *for the IPS if*

$$\mathbb{P}^\nu \circ \omega_t^{-1} = \nu, \qquad \text{for all } t \geq 0.$$

(We use the measure-theoretic notation $\mathbb{P} \circ X^{-1}$ for the distribution of a random variable X under a probability measure $\mathbb{P}$.) Moreover, the IPS is called ergodic *if it has a unique invariant measure ν that satisfies that, for all starting measures ν',*

$$\lim_{t \to \infty} \mathbb{P}^{\nu'} \circ \omega_t^{-1} = \nu$$

in the weak topology.

Invariant measures represent initial distributions that do not change under the evolution of the process. Note further that the set of invariant measures is always a convex set and hence it can be represented by the subset of extremal invariant measures, which are the ones that can not be represented by a convex combination of other invariant measures.

Now we turn to the special case of the contact process on $\mathbb{Z}^d$ (i.e., with the state space $\Omega = \{0,1\}^{\mathbb{Z}^d}$) and with parameter $\lambda \in (0, \infty)$ as in (5.2.1). Obviously, the delta measure δ_0,

which puts mass one on the configuration ω with $\omega_{x_i} = 0$ for all $x_i \in x$, is invariant under the evolution of the contact process. We have the following theorem.

Theorem 5.3.2. *Consider $x = \mathbb{Z}$, then there exists $\lambda_{\mathrm{cr}} \in (0, \infty)$ such that for $\lambda < \lambda_{\mathrm{cr}}$, the contact process is ergodic. For $\lambda > \lambda_{\mathrm{cr}}$ the process is not ergodic. In particular, there exist infinitely many additional invariant measures.*

Remark 5.3.3 (Properties of the contact process)**.**

1. In the super-critical regime, the set of invariant measures is given by convex combinations of δ_0 and another invariant measure ν_λ, usually referred to as the *upper invariant measure*. It can be defined as $\lim_{t\to\infty} \mathbb{P}^{\delta_1} \circ \omega_t^{-1} = \nu_\lambda$. In this regime, for some starting configurations, the malware can survive for all times.
2. Approximations for the critical rate suggest that $\lambda_{\mathrm{cr}} \approx 1.6494$, see [L85, page 275].
3. Similar to the percolation probability θ, we can define $\rho(\lambda) = \nu_\lambda(\zeta_0 = 1)$, the probability that the device at the origin is infected under ν_λ. Again, close to the critical value, this quantity should behave like a power law with some critical exponent. ◇

5.4 Other IPS for telecommunications

The contact process, although defendable as a model for simple telecommunication systems, is certainly not the most accurate description of malware propagation in such a network. Another classical model that has an interpretation in telecommunications is the *voter model*, see [L85]. Here, $q \in \mathbb{N}$ different opinions, or viruses, are in the system and the updating is given by

$$c(\zeta, \zeta^{i,a}) = \lambda \sum_{j \in I:\, x_j \sim x_i} \mathbb{1}\{\zeta_{x_j} = a\}.$$

In words, the configuration ζ flips site i into state a with a rate equal to λ times the number of neighboring sites that have the opinion a.

Let us finally mention some extensions and generalizations that we believe should be investigated in the context of IPS for telecommunications.

Remark 5.4.1 (IPS for telecommunications)**.**

1. The set up of possible states should at least comprise a state of susceptibility, a state of infection and a state of immunization. The immune state then refers to a device on which an update has been installed.
2. For data protection reasons, the update can only be installed by the operator on an infected device if this device has made its infection obvious on its own, for example by attacking a device that possesses the healing capability.
3. Users with healing capability, sometimes called *white knights*, should be placed deterministically or randomly in the system. Their placement introduces another set of network parameters, which must be distinguished.
4. The setting of IPS should be extended to cover random geometries. For example, one should investigate IPS on realizations of the PPP with a neighborhood structure provided for example by the Boolean model. Here typically two scenarios are relevant: (1) the *annealed* setting, where the properties are investigated in expectation of the realizations of the Boolean model, and (2) the *quenched* setting, where a typical realization of the Boolean model is fixed in advance hopefully leading to results in the almost sure sense. ◇

Let us close this section by giving some illustration of the spread of some malware, starting from one infected user at the origin in the setting of a Boolean model for a Cox point process, see Figure 5.4.1.

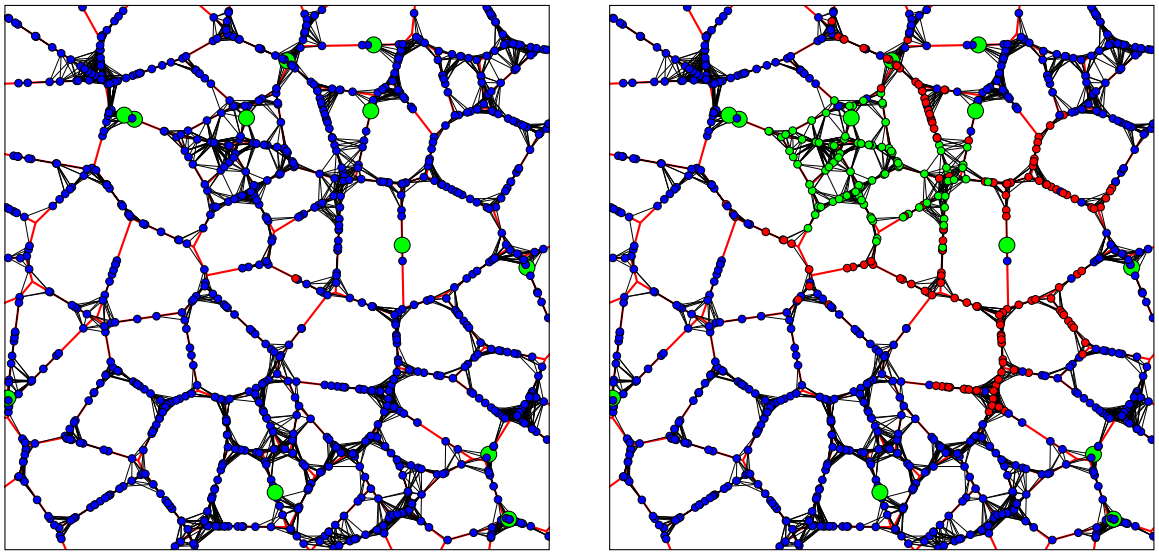

Figure 5.4.1: Realization of randomly placed users on a street system of Poisson-Voronoi tessellation type. The larger green discs indicate randomly placed white knights. Left: initial time with one infected user in red at the center. Right: some finite time snapshot in which the malware has started to infect surrounding users in red. The green regular size users are vaccinated by an software update initiated by some contact of the infection with a white knight.

References

[BR06] B. Bollobás and O. Riordan, *Percolation*, Cambridge University Press, New York (2006).

[Bi68] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York (1968).

[BB09a] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume I: Theory* Now Publishers Inc. (2009).

[BB09b] F. Baccelli and B. Błaszczyszyn, *Stochastic Geometry and Wireless Networks: Volume II: Applications* Now Publishers Inc. (2009).

[CHJ17] E. Cali, C. Hirsch and B. Jahnel, Continuum Percolation for Cox processes, *Preprint available at arXiv:1710.11407* (2017).

[CSKM13] S.N. Chiu, D. Stoyan, W.S. Kendall and J. Mecke, *Stochastic Geometry and Its Applications*, J. Wiley & Sons, Chichester (2013).

[DVJ03] D.J. Daley and D. Vere-Jones, *An Introduction to the Theory of Point Processes, Vol. I: Elementary Theory and Methods*, Second Edition, Springer (2003).

[DZ10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer (2010). Corrected reprint of the second (1998) edition.

[DBT05] O. Dousse, F. Baccelli and P. Thiran Impact of interferences on connectivity in ad hoc networks, *IEEE/ACM Transactions on Networking (TON)* **13**, 425–436 (2005).

[DFMMT06] O. Dousse, M. Franceschetti, N. Macris, R. Meester and P. Thiran Percolation in the signal to interference ratio graph, *Journal of Applied Probability* **43**, 552–562 (2006).

[FM08] M. Franceschetti and R. Meester, *Random Networks for Communication: from Statistical Physics to Information Systems*, Cambridge University Press, Cambridge (2008).

[G61] E.N. Gilbert, Random plane networks, *J. of the Society for Industrial & Appl. Math.* **9**, 533–543 (1961).

[G89] G.R. Grimmett, *Percolation*. New York: Springer-Verlag (1989).

[H12] M. Hänggi, *Stochastic Geometry for Wireless Networks*, Cambridge University Press, Cambridge (2012).

[HJKP18] C. Hirsch, B. Jahnel, P. Keeler and R. Patterson, Large deviations in relay-augmented wireless networks, *Queueing Systems* **88:3–4** (2018).

[J86] S. Janson, Random coverings in several dimensions, *Acta Math.* **156**, 83–118 (1986).

[K95] J. Kingman, *Poisson Processes*, Volume 3 of Oxford Studies in Probability, Oxford University Press, Oxford (1995).

[Kl06] A. Klenke, *Wahrscheinlichkeitstheorie*, Springer, Berlin, 2006.

[LP17] G. Last and M. Penrose, *Lectures on the Poisson Process*, IMS Textbook, Cambridge University Press, Cambridge (2017).

[L85] T. Liggett, *Interacting Particle Systems*, New York: Springer-Verlag (1985).

[M75] G. Matheron, *Random Sets and Integral Geometry.* London: John Wiley & Sons (1975).

[MR96] R. Meester and R. Roy. *Continuum Percolation*, Cambridge University Press, Cambridge (1996).

[OBSC00] A. Okabe, B. Boots, K. Sugihara and S.N. Chui. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, J. Wiley & Sons, Chichester (2000).

[P91] M.D. Penrose. On a continuum percolation model, *Adv. Appl. Probab.* **23:3**, 536–556 (1991).

[P03] M.D. Penrose. *Random Geometric Graphs*, volume 5 of *Oxford Studies in Probability.* Oxford University Press, Oxford (2003).

[QZ07] J.A. Quintanilla and R.M. Ziff. Asymmetry in the percolation thresholds of fully penetrable disks with two different radii, *Phys. Rev.* **E76**, 051–115 (2007).

[RS15] F. Rassoul-Agha and T. Seppäläinen, A course on large deviation theory with an introduction to Gibbs measures, *Graduate Studies in Mathematics* **162**, AMS (2015).

[Re87] S.I. Resnick, *Extreme Values, Regular Variation and Point Processes*, Applied Probability. A Series of the Applied Probability Trust 4. Springer, New York (1987).

[YCG11] C. Yao, G. Chen and T.D. Guo, Large deviations for the graph distance in supercritical continuum percolation, *J. Appl. Probab.* **48**, 154–172 (2011).

Index