Crowd-networking : modelling Device-to-Device connectivity on street systems using percolation theory, and economic consequences

Quentin Le Gall

PSL
UNIVERSITÉ PARIS

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL

Préparée à l'Ecole Normale Supérieure et à l'INRIA
Dans le cadre d'une convention CIFRE avec Orange

Crowd-networking : modèles de percolation pour la connectivité Device-to-Device en environnement urbain, et conséquences économiques

City crowd-networking: modelling Device-to-Device connectivity on street systems using percolation theory, and economic consequences

Soutenue par
Quentin LE GALL
Le 06 octobre 2020

École doctorale n°386
Sciences Mathématiques de Paris Centre

Spécialité
Mathématiques appliquées

ENS | PSL
Inria INVENTEURS DU MONDE NUMÉRIQUE
orange

Composition du jury :

Philippe MARTINS Professeur, Télécom Paris	*Président*
David COUPIER Professeur, IMT Lille-Douai	*Rapporteur*
Laurent DECREUSEFOND Professeur, Télécom Paris	*Rapporteur*
Catherine GLOAGUEN Ingénieure de recherche, Orange Labs	*Examinatrice*
Justin SALEZ Professeur, Université Paris Dauphine	*Examinateur*
Marie THÉRET Professeure, Université Paris Nanterre	*Examinatrice*
Bartłomiej BŁASZCZYSZYN Directeur de recherche, INRIA-ENS	*Directeur de thèse*
Élie CALI Ingénieur de recherche, Orange Labs	*Encadrant de thèse*
Taoufik EN-NAJJARY Ingénieur de recherche, Orange Labs	*Invité*

N° National de Thèse : XXX

THÈSE DE DOCTORAT

en vue de l'obtention du grade de

Docteur de l'École Normale Supérieure,
dans le cadre d'une convention CIFRE avec Orange S.A.

Discipline : **Mathématiques Appliquées**

Département d'informatique de l'ENS et INRIA (UMR 8548)

École Doctorale n°386 : Sciences Mathématiques de Paris Centre

Présentée et soutenue publiquement le 06 octobre 2020
par **M. Quentin LE GALL**

City crowd-networking: modelling Device-to-Device connectivity on street systems using percolation theory, and economic consequences

---

Directeur de Thèse : Professeur Bartłomiej Błaszczyszyn

Devant le jury suivant :

Professeur	Philippe Martins	*Télécom Paris*	Président
Professeur	David Coupier	*IMT Lille-Douai*	Rapporteur
Professeur	Laurent Decreusefond	*Télécom Paris*	Rapporteur
Docteure	Catherine Gloaguen	*Orange Labs Networks*	Examinatrice
Professeur	Justin Salez	*Université Paris Dauphine*	Examinateur
Professeure	Marie Théret	*Université Paris Nanterre*	Examinatrice
Professeur	Bartłomiej Błaszczyszyn	*ENS-INRIA*	Directeur de thèse
Docteur	Élie Cali	*Orange Labs Networks*	Encadrant de thèse
Docteur	Taoufik En-Najjary	*Orange Labs Networks*	Membre invité

Centre de Recherche INRIA Paris
Équipe DYOGENE
2, rue Simone Iff
CS 42112
75589 Paris Cedex 12
FRANCE

Orange Gardens
Modelling and Statistical Analysis
44, avenue de la République
CS 50010
92326 Châtillon Cedex
FRANCE

Remerciements

Il y a, paraît-il, une certaine constance chez les doctorants : l'écriture des remerciements est un moment d'extase ; en partie car il signifie la fin de l'écriture de leur thèse ; et très certainement car il s'agit du seul endroit d'un pavé de plusieurs centaines de pages où l'auteur peut se permettre d'être un tantinet moins professionnel que dans le reste dudit pavé. Je ne dérogerai donc pas à la règle pour goûter, à mon tour, à la saveur de ce moment.

Mes premiers remerciements vont tout naturellement à mon directeur de thèse Bartek Błaszczyszyn ainsi qu'à mes encadrants de thèse Élie Cali et Taoufik En-Najjary. Ils ont su me mettre en confiance et être disponibles en tout temps durant ces trois années, tant pour des discussions mathématiques que plus récréatives. C'est grâce à leur encadrement bienveillant et leur (infinie !) patience que cette thèse a pu voir le jour.

Je suis profondément honoré que David Coupier et Laurent Decreusefond aient accepté d'être les rapporteurs de ma thèse et je les remercie pour leur relecture attentive de mon manuscrit ainsi que pour leurs remarques d'experts qui ont permis d'améliorer son contenu.

J'ai l'immense chance d'avoir un jury formé de grands experts et je remercie très chaleureusement Catherine Gloaguen, Philippe Martins, Justin Salez et Marie Théret d'avoir accepté de donner leur temps pour en faire partie. Leur présence dans mon jury m'honore.

Cette thèse a été effectuée dans un environnement tri-partite, entre Orange Labs à Châtillon, l'INRIA à Bercy et l'ENS, rue d'Ulm. Dans chacun de ces lieux, j'ai toujours pu bénéficier de conditions de travail exceptionnelles et d'un environnement épanouissant, tant scientifiquement qu'humainement.

Un grand merci à mes collègues de l'équipe MSA d'Orange Labs qui m'ont chaleureusement accueilli et ont partagé mon quotidien pendant mon doctorat : Aurélien, Catherine, Chaima, Élie, Marie-Pierre, Matthieu, Mikaël, Olfa, Stanislas, Stéphane, Taoufik, Rachel, Wafae, Yu et tous ceux dont le passage dans notre équipe a été plus bref. Ils ont, chacun à leur manière, contribué à l'aboutissement et à la réussite de mon doctorat.

Une thèse n'en est pas une sans quelques menues difficultés administratives qui vont avec. Toutefois, j'ai eu la chance de pouvoir compter sur des persones compétentes, efficaces et disponibles pour qui les rouages de l'administration de l'ENS, de l'Ecole Doctorale 386 et d'Orange n'ont aucun secret. Merci donc à Olfa, Régine, Mustapha, Sophie et Hélène pour leur aide si précieuse dans ces moments où la paperasse prenait le dessus sur la thèse et m'agaçait plus qu'autre chose.

Merci du fond du cœur aux doctorants et post-doctorants d'Orange, de l'Inria et de l'ENS avec qui j'ai partagé tout ou partie de cette thèse : Abdellatif, Adrien, Ahlam, Antoine, Arnaud, Ayman, Fatma, Georges, Iaad, Isaias, Marie, Paul, Raquel, Vincent, Wesley et Wilfried. C'est grâce à vous que l'ambiance au labo ou au deuxième étage du bâtiment 2A était aussi sympathique ! Je garde d'excellents souvenirs des discussions que nous avons partagées à titre professionnel ou au cours de mémorables pauses café qui ont grandement fait avancer le monde de la recherche.

Je suis également reconnaissant envers les professeurs qui ont croisé ma route d'ingénieur puis de mathématicien, d'un côté et de l'autre de la Manche. Un merci tout particulier à Célestin Rakotoniaina et Marie-Christine Pauzin, avec qui les relations ont été constantes au fil des années : de professeurs de Mathématiques Supérieures, vous êtes devenus collègues et amis !

Mes remerciements suivants vont à mes fort sympathiques élèves de la Prépa Médecine du lycée Stanislas. Leur goût (ou aversion, selon les cas ...) pour les mathématiques a très certainement fait de moi un meilleur enseignant, toujours en quête (plus ou moins réussie) de nouvelles manières de leur enseigner les réjouissances au programme de leur concours. Sans eux, cette thèse aurait très probablement comporté une centaine de pages supplémentaires ... mais leur entrain et leur humour (décapant pour certains !) font largement pencher la balance de leur côté.

Je voudrais également dire un grand merci à ma filleule et mes filleuls, qui me poussent sans cesse à donner le meilleur de moi-même. Merci pour votre confiance et tous ces beaux moments partagés ensemble.

Dans un registre plus personnel, je voudrais remercier ma famille et tout particulièrement ma mère et mon père, qui ont toujours cru en moi et m'ont soutenu lors de mes (longues) études. Trente ans après votre passage sur les bancs de l'ENS, c'est avec une grande émotion que je soutiens ma thèse devant vous, dans cette même école qui vous accueillis !

Je suis infiniment reconnaissant envers mes amis pour leur présence au cours de ces trois dernières années, leur soutien, leur écoute attentive (surtout quand je râlais à propos de ma thèse ou de mes simulations qui ne marchaient pas !) et leur disponibilité pour m'aider à penser à autre chose que la thèse quand j'en avais besoin. J'espère que chacun d'entre vous se reconnaîtra dans ces lignes : les amis du scoutisme, de la prépa, les Centraliens, les Cantabrigiens... et tous ceux qui ont leur propre catégorie ! Un merci tout particulier à Pierre S. pour sa présence constante au long des années et parce qu'il est lui, à Mathieu qui voit toujours le verre à moitié plein plutôt qu'à moitié vide, à Pierre C. pour son humour à toute épreuve et à Mathias pour son attention à ma santé.

Pour finir, j'aimerais remercier mon épouse Charlotte pour sa présence quotidienne à mes côtés et son indéfectible soutien. Il n'y aura jamais assez de mots pour dire tout ce que je lui dois. Je conclurai donc avec ceci : cette thèse est tout autant la sienne que la mienne.

« Oh ! reprit-il, je donnerais bien cent sous au mathématicien qui me démontrerait par une équation algébrique l'existence de l'enfer. Il jeta une pièce en l'air en criant :
— Face pour Dieu !
— Ne regarde pas, dit Raphaël en saisissant la pièce, que sait-on ? Le hasard est si plaisant. »

Honoré de Balzac, *La peau de chagrin*

A mon épouse Charlotte et notre enfant à naître.

Contents

List of Figures

List of Tables

List of Notations and Abbreviations

Abbreviations

ANFR	**A**gence **N**ationale des **Fr**équences (French National Frequency Agency)
a.s.	**a**lmost **s**urely
AWGN	**A**dditive **W**hite **G**aussian **N**oise
BS	**B**ase **S**tation
CAPEX	**Cap**ital **Ex**penditure
CF	**C**ash **F**low
CR	**C**umulated **R**evenue
D2D	**D**evice-**To**-**D**evice
eNB	**E**volved **N**ode **B**
Hz	**Hertz**
i.i.d.	**i**ndependent and **i**dentically **d**istributed
IEEE	**I**nstitute of **E**lectrical and **E**lectronics **E**ngineers
ISM	**I**ndustrial **S**cientific and **M**edical radio band
ITU	**I**nternational **T**elecommunication **U**nion
LCSCH	**L**ocally **C**ompact **S**econd **C**ountable **H**ausdorff
LOS	**L**ine-**O**f-**S**ight
LTE	**L**ong **T**erm **E**volution
m	**m**etres
m.s^{-1}	**m**etres per **s**econd
MG	**M**anhattan **G**rid
NB	**N**ode **B**

NLOS	**N**on-**L**ine-**O**f-**S**ight
NoSha	**No Sha**dowing
OPEX	**Op**erational **Ex**penditure
PDT	**P**oisson **D**elaunay **T**essellation
PLT	**P**oisson **L**ine **T**essellation
PPP	**P**oisson **P**oint **P**rocess
PSTN	**P**ublic **S**witched **T**elephone **N**etwork
PVT	**P**oisson **V**oronoi **T**essellation
ROI	**R**eturn **O**n **I**nvestment
RR	**R**adio **R**egulations
SAP	**S**elf-**A**voiding **P**ath
SINR	**S**ignal-to-**I**nterference-plus-**N**oise **R**atio
SIR	**S**ignal-to-**I**nterference **R**atio
SNR	**S**ignal-to-**N**oise **R**atio
UE	**U**ser **E**quipment
W	**W**att
1G	**First G**eneration of cellular networks
2G	**Second G**eneration of cellular networks
3G	**Third G**eneration of cellular networks
4G	**Fourth G**eneration of cellular networks
5G	**Fifth G**eneration of cellular networks

Notations

Some of the notations recurrently used in this thesis are given below.

Spaces

$\mathbb{R}$	Real numbers
$\mathbb{R}_+$	Non-negative real numbers
$\mathbb{N}$	Set of positive integers of $\mathbb{R}$: $\mathbb{N} = \{1, 2, \dots, \}$
$\mathbb{N}_0$	$\mathbb{N} \cup \{0\}$

$\mathbb{Z}$	Set of integers of $\mathbb{R}$: $\mathbb{Z} = \{\ldots, -1, 0, 1, \ldots\}$
$\mathbf{E}$	denotes an LCSCH space
$\mathcal{B}$	denotes the Borel σ-algebra of an LCSCH space $\mathbf{E}$
$\mathcal{B}_c$	denotes the set of bounded Borel subsets of an LCSCH space $\mathbf{E}$
$\mathbb{M}$	denotes the space of counting measures of an LCSCH space $\mathbf{E}$
$\mathcal{M}$	denotes the evaluation σ-algebra on $\mathbb{M}$
$\mathbf{M}$	denotes the space of Borel measures on $\mathbb{R}^2$
$(\Omega, \mathcal{A}, \mathbb{P})$	denotes a probability space

Sets

$\#(\ldots)$	cardinal of the set $\ldots$
$\setminus$	Set difference: $A \setminus B := \{x \in A,\, x \notin B\}$
A^c	Complementary of A: $A^c := \{x,\, x \notin A\}$
$\lvert A \rvert$	Lebesgue measure of $A \subset \mathbb{R}^d$
$\mathrm{cl}(A)$	Topological closure of the set $A \subset \mathbb{R}^d$
$\mathring{A}$	Topological interior of the set $A \subset \mathbb{R}^d$
∂A	Topological boundary of the set $A \subset \mathbb{R}^d : \partial A = \overline{A} \setminus \mathring{A}$

Probability and measures

$\mathbb{1}\{\cdot\}$	denotes the indicator function equal to 1 if true and 0 if false
$\perp\!\!\!\perp$	denotes the independence of two random variables
$\overset{d}{=}$	means equality in distribution
$\mathbb{P}$	Probability measure
$\mathbb{E}$	Expectation with respect to $\mathbb{P}$
$\mathbb{P}_X$	probability distribution of the random variable X under the probability measure $\mathbb{P}$
$\mathbb{P}^0$	Palm probability
$\mathbb{E}^0$	Palm expectation
$\mathcal{U}[a, b]$	Uniform distribution on the interval $[a, b]$
$\mathcal{N}(\mu, \sigma^2)$	Normal distribution with mean μ and variance σ^2
$\overset{i.i.d.}{\sim}$	denotes i.i.d. random variables

δ_x	Dirac measure at x
μ_A	denotes the restriction of the measure μ to a set A: $\mu_A(\cdot) := \mu(A \cap \cdot)$
$\mathrm{supp}(\mu)$	denotes the support of a measure μ
$L^1(\mu)$	Set of functions that are integrable with respect to the measure μ

General notation

$\square$	denotes the end of a proof
$\min(a, b)$	Minimum of the real numbers a and b
$\max(a, b)$	Maximum of the real numbers a and b
$\lfloor x \rfloor$	Lower integer part of the real x: $\lfloor x \rfloor := \max\{k \in \mathbb{Z},\, k \leq x\}$
$\lceil x \rceil$	Upper integer part of the real x: $\lceil x \rceil := \min\{k \in \mathbb{Z},\, k \geq x\}$
$\|\cdot\|_1$	Taxicab norm on $\mathbb{R}^d$: if $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $\|x\|_1 := \sum_{i=1}^{d} \lvert x_i \rvert$
$\|\cdot\|_2$	Euclidean norm on $\mathbb{R}^d$: if $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $\|x\|_2 := \sqrt{\sum_{i=1}^{d} \lvert x_i \rvert^2}$
$\|\cdot\|_\infty$	Supremum norm on $\mathbb{R}^d$: if $x = (x_1, \ldots, x_d) \in \mathbb{R}^d$, $\|x\|_\infty := \max_{i=1,\ldots,d} \lvert x_i \rvert$
$a \ll b$	informally means that a is much smaller than b
$a \gg b$	informally means that a is much bigger than b
$a \propto b$	means that a and b are proportional
dim	Dimension of a linear subspace of a vector space
deg	Vertex degree in a graph
V	Vertex-set of a graph $G = (V, E)$
E	Edge-set of a graph $G = (V, E)$
$\mathbb{L}^d$	d-dimensional Euclidean lattice
$Q_n(x)$	d-dimensional cube of side length $n \in \mathbb{N}$ and centre $x \in \mathbb{R}^d$: $Q_n(x) := x + [-n/2, n/2]^d$
Q_n	$Q_n(0)$
$\mathcal{B}(x, r)$	d-dimensional closed Euclidean ball with centre $x \in \mathbb{R}^d$ and radius $r > 0$: $\mathcal{B}(x, r) := \{y \in \mathbb{R}^d,\, \|x - y\|_2 \leq r\}$.
$\mathscr{B}(x, r)$	d-dimensional closed ball with centre $x \in \mathbb{R}^d$ and radius $r > 0$ for the supremum norm : $\mathcal{B}(x, r) := \{y \in \mathbb{R}^d,\, \|x - y\|_\infty \leq r\}$.

Introduction générale

La cinquième génération de réseaux mobiles (5G) devra relever de nombreux défis. Être en mesure de servir un nombre jamais vu d'équipements avec une qualité de service satisfaisante sera l'un d'entre eux. En effet, d'après une étude d'Ericsson [3], le nombre total d'appareils connectés dans le monde était d'environ 15 milliards en 2015 et devrait atteindre 30 milliards en 2022. Cette explosion du nombre d'équipments ainsi que l'émergence de nouvelles applications consommant de plus en plus de données (consultation de sites internet, jeu en ligne, téléchargement et visionnage de vidéos en haute résolution, internet des objets ...) ont conduit à la problématique suivante : comment faire face à cette croissance du trafic tout en maintenant une qualité de service acceptable ?

Pour répondre à cette problématique, il faut considérer de nouvelles manières de penser les réseaux de demain. Une des principales technologies étudiées dans la littérature pour répondre à ce défi est le *device-to-device* (D2D). Le D2D permet l'établissement de communications directes et de courte portée entre deux appareils, sans que le signal ne doive passer par des infrastructures supplémentaires dans le réseau. Les applications du D2D pour les réseaux cellulaires sont nombreuses et très prometteuses.

Un cas d'usage d'intérêt pour les opérateurs est celui de l'übérisation des réseaux. Si un nouvel opérateur n'ayant pas (ou presque pas) d'infrastructures réseaux arrivait à construire un réseau reposant uniquement sur des terminaux mobiles, il aurait à sa disposition un réseau fonctionnel sans avoir eu besoin d'investir dans des infrastructures réseau. Ceci pourrait donc constituer une menace économique sérieuse pour les opérateurs historiques. A contrario, la possibilité de construire un réseau sans devoir effectuer de tels investissements pourrait être une opportunité pour les opérateurs historiques dans des pays où ils ne sont pas encore implantés.

Dans la littérature, l'écrasante majorité des articles de recherche concernant le D2D s'intéresse soit aux applications qui peuvent en être faites, soit à des problématiques techniques très particulières (sécurité, consommation énergétique, distribution des ressources radio, gestion des interférences ...). Cependant, ces travaux ne s'intéressent que rarement au problème de la modélisation de réseaux D2D à grande échelle. La connectivité de tels réseaux est alors souvent supposée établie et aucune condition de faisabilité n'est mentionnée.

Contexte de la thèse et objectifs de recherche

Dans cette thèse, nous nous intéressons à de nouveaux modèles de réseaux D2D en environement urbain qui nous permettent d'étudier la problématique de leur connectivité à grande échelle, ainsi que les conséquences économiques de la faisabilité de tels réseaux D2D à grande échelle pour les opérateurs. Plus particulièrement, nous essayons de répondre aux questions suivantes :

- Comment prendre en compte la distribution spatiale des utilisateurs, la topologie de l'environnement urbain couvert par le réseau et des conditions de connectivité suffisamment réalistes entre les nœuds du réseau ?

- Quels sont les paramètres minimaux (par exemple la densité minimale d'utilisateurs) permettant une bonne connectivité du réseau et comment les estimer ?

- Comment quantifier les investissements requis pour mettre en place un réseau D2D fonctionnel ?

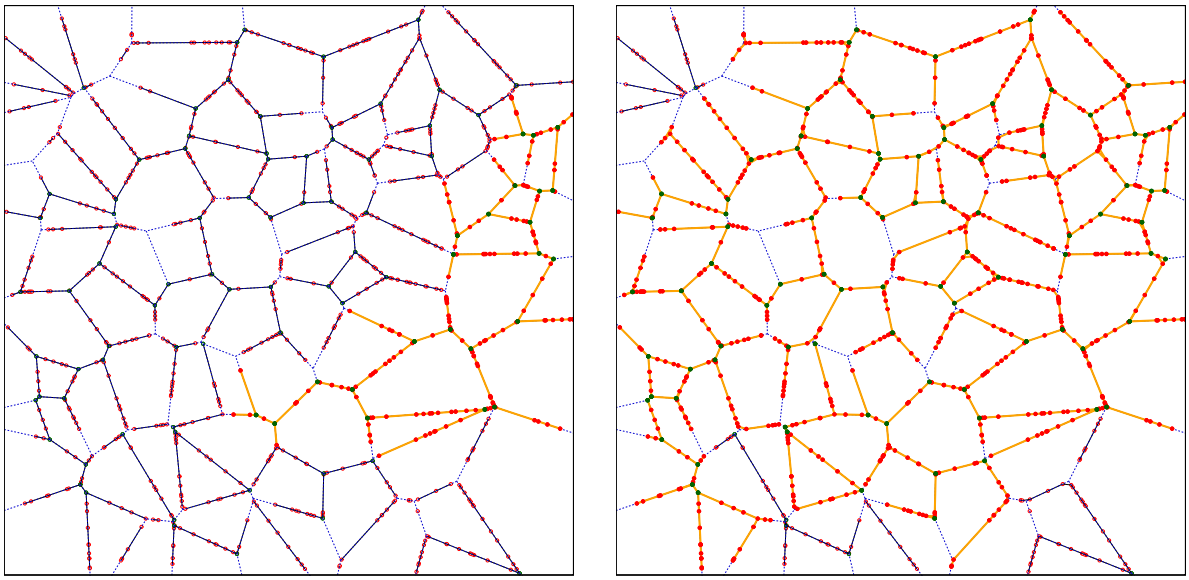
- Quelles pourraient être les conséquences économiques du déploiement de réseaux basés sur le D2D pour les opérateurs ?

Pour répondre aux questions précédentes, nous proposons de nouveaux modèles stochastiques pour les réseaux D2D, basés sur des outils empruntés à la géométrie stochastique et à la théorie de la percolation.

Contributions de la thèse

Dans notre travail, nous introduisons de nouveaux modèles mathématiques de réseaux D2D en environnement urbain. En résumé, notre approche de modélisation est comme suit. Un *système de rues* S modélisé par une *mosaïque de Poisson-Voronoi* (abrégé en PVT, sigle de l'anglais *Poisson-Voronoi tessellation*) est le support de deux processsus ponctuels : un *processus de Cox* modélisant les utilisateurs du réseau répartis sur les arêtes de S et un *processus de Bernoulli* modélisant des relais répartis sur les sommets de S. En notant Z la superposition de ces deux processus ponctuels, nous modélisons le réseau par un *graphe aléatoire de connectivité* dont les sommets sont les atomes de Z et où une arête est tracée entre deux points de Z selon certaines règles données par des *conditions de connectivité* entre les nœuds du réseau. Nous interprétons la *percolation* de ce graphe aléatoire, c'est à dire une probabilité positive d'existence d'une composante connexe infinie, comme signe d'une bonne connectivité du réseau modélisé.

Nous étudions dans un premier temps un cas particulier où les seules connexions possibles sont celles en *vision directe* (abrégé en LOS, sigle de l'anglais *line-of-sight*), le long de la PVT S et dans une certaine portée. Autrement dit, deux nœuds du réseau sont connectés par une arête dans le graphe de connectivité si et seulement si ils sont suffisamment proches l'un de l'autre et sont situés sur la même arête de la PVT S. Un exemple d'un graphe de connectivité du réseau avec de telles hypothèses est donné par la Figure (a). L'une des principales contributions mathématiques de cette thèse est l'étude du



(a) Simulation d'un graphe de connectivité quand seules les connexions LOS (vision directe) sont possibles.

(b) Simulation d'un graphe de connectivité dans le cas où les connexions supplémentaires en NLOS (vision indirecte) sont prises en compte.

Figure : Deux exemples de simulations d'un graphe de connectivité du réseau. Les lignes pointillées bleues représentent le système de rues (PVT S), les points rouges représentent les utilisateurs du réseau (processus de Cox), les points verts représentent les relais (processus de Bernoulli). La plus grande composante connexe de la fenêtre de simulation est mise en valeur par des arêtes orangées et des nœuds coloriés dans leurs couleurs respectives. Les arêtes des autres composantes connexes sont noircies. Dans les deux cas, les réalisations du système de rues, du processus de Cox et du processus de Bernoulli sont les mêmes.

graphe de connectivité du réseau sous ces hypothèses. Nous démontrons que des transitions de phase entre différents régimes de connectivité ont lieu. Deux régimes particulièrement intéressants sont mis en évidence. Dans le premier d'entre eux (appelé régime *relay-limited*, de l'anglais "limité par les relais"), la connectivité du réseau peut être assurée uniquement par les relais (points du processus de Bernoulli) tandis que dans le second régime (appelé régime *relay-and-user-limited*, de l'anglais "limité par les relais et les utilisateurs"), la présence d'utilisateurs (points du processus de Cox) est essentielle. A l'aide de simulations numériques, nous estimons également les frontières de ces différents régimes. Connaître ces frontières est en effet important pour des questions de planification économique et permet également de donner des conditions pour que le réseau ainsi modélisé puisse fournir une bonne connectivité.

Dans le cas particulier décrit ci-dessus, la présence de relais (points du processus de Bernoulli) aux carrefours des rues (sommets de la PVT S) est cruciale car les seules connexions possibles dans le graphe de connectivité ont lieu le long des rues (arêtes de la PVT S). Une quantité insuffisante de relais peut donc constituer un obstacle à la connectivité du réseau. Nous proposons deux variantes du précédent modèle pour pallier

ce problème.
D'une part, nous laissons le modèle géométrique tel quel mais nous apportons quelques modifications au modèle de connectivité. Dans cette nouvelle approche, les connexions en vision directe sont conservées et des connexions suplémentaires en *vision indirecte* (abrégé en NLOS, sigle de l'anglais *non-line-of-sight*) sont égalements envisagées. Deux nœuds du réseau pourront ainsi être connectés par une arête dans le graphe de connectivité s'ils sont situés sur deux arêtes de la PVT S qui sont incidentes à un même sommet v et si la somme des distances respectives de chacun des deux nœuds à v est inférieure à un certain seuil. Ces connexions supplémentaires peuvent être vues comme étant dues à des réflexions ou à la diffusion du signal par des obstacles physiques présents aux carrefours (sommets de la PVT S). La Figure (b) donne une illustration d'un graphe de connectivité prenant en compte ces connexions NLOS supplémentaires. L'autre principale contribution mathématique de cette thèse est l'étude de la percolation du nouveau graphe de connectivité résultant de la prise en compte des connexions NLOS.
D'autre part, nous proposons une approche plus appliquée où les arêtes du système de rues sont "élargies". Les carrefours, qui sont donnés par les intersections des rues, ne sont donc plus ponctuels et ont une surface positive. Ce faisant, nous sommes en mesure de revoir nos estimations des seuils délimitant les frontières entre les différents régimes de connectivité du réseau.

Nous introduisons ensuite un *modèle de coûts* afin de compléter les précédentes estimations numériques et d'étudier de potentiels *exemples d'übérisation* des réseaux de télécommunications où un "néo-opérateur" souhaite construire un réseau entièrement fonctionnel en se basant seulement sur le D2D. Nous établissons, de manière quantitative, des relations fondamentales entre des paramètres du réseau (densité d'utilisateurs et de relais) et la possibilité d'établir des réseaux D2D à grande échelle. Ce faisant, nous donnons des outils d'aide à la décision pour les opérateurs : il s'agit là de la principale contribution opérationnelle de notre travail.

En dernier lieu, les simulations numériques permettant d'estimer les paramètres critiques des modèles de percolation étudiés dans cette thèse ont été réalisées à l'aide d'algorithmes originaux. Ceci constitue la principale contribution de notre travail d'un point de vue algorithmique et pourrait être appliqué à d'autres problèmes de percolation.

Travaux en lien avec le sujet de la thèse

Avant d'en venir au cœur du sujet, nous introduisons quelques travaux fondateurs en lien avec cette thèse et qui aideront à comprendre l'approche de recherche adoptée dans notre travail. Une revue plus exhaustive de la littérature sera faite dans les prochains chapitres de ce document.

Les modèles mathématiques de réseaux de télécommunications fondés sur une approche en lien avec la percolation ne sont pas tout à fait une idée récente. Au début des années 60, dans l'article fondateur du domaine [55], Gilbert a eu l'idée de modéliser les nœuds d'un réseau de télécommunications sans fil par les atomes d'un processus ponctuel de Poisson dans le plan $\mathbb{R}^2$. En notant par $R > 0$ la portée (déterministe) des connexions,

le réseau est modélisé par un graphe de connectivité où deux points du précédent processus de Poisson planaire sont reliés par une arête s'ils sont à une distance inférieure ou égale à R l'un de l'autre. Pour $N \geq 1$, Gilbert s'est ensuite intéressé à la probabilité $P(N)$ qu'il existe une composante connexe de ce graphe contenant au moins N points. A l'aide de simulations numériques, Gilbert mit en évidence une transition de phase pour le réseau : lorsque la surface de couverture $E = \pi R^2$ d'un nœud du réseau devient supérieure à une certaine valeur critique E_c, la probabilité $P(\infty) := \lim_{n\uparrow\infty} P(N)$ qu'il existe une composante connexe infinie dans le réseau devient strictement positive. L'approche de Gilbert a donné naissance à la théorie moderne de la percolation continue. De nos jours, la théorie de la percolation (continue ou non) reste un domaine de recherche très actif ayant de nombreuses applications dans divers domaines des sciences et en particulier en télécommunications.

Des modèles stochastiques pour représenter de vrais systèmes de rues ainsi que des applications aux réseaux de télécommunications ont été étudiés dans [57–60]. L'idée principale développée dans ces travaux est qu'un système de rues peut être remplacé par une mosaïque aléatoire lui étant statistiquement équivalente, c'est à dire ayant les mêmes propriétés géométriques moyennes. Cette idée permet ainsi de considérablement simplifier des problèmes de planification économique des réseaux ou des problèmes d'optimisation liés au dimensionement de réseau (par exemple le problème du câblage dans un réseau d'accès). Des procédures statistiques visant à trouver le meilleur modèle, parmi différentes classes de mosaïques, pour représenter une extraction d'une carte urbaine ont été étudiées dans la thèse [32].

En se basant sur l'approche précédente et l'idée initiale de Gilbert, la suite logique fut de considérer des modèles plus complexes de réseaux de télécommunications où les nœuds du réseau ne peuvent pas être situés n'importe où dans le plan mais sont plutôt contraints à être localisés sur un support aléatoire tel qu'une mosaïque aléatoire. En gardant une dispersion poissonienne des nœuds du réseau, on en vient donc à être confronté à l'étude de processus ponctuels de Poisson doublement stochastiques, souvent appelés processus de Cox dans la littérature. La percolation de tels processus de Cox a été étudiée pour la première fois dans [70] et une application à la modélisation de réseaux de télécommunications a été traitée dans [29].

Dans la plupart des problèmes de percolation, des expressions analytiques des seuils critiques délimitant les transitions de phase du système sont souvent hors de portée. Le simple fait de trouver des bornes pour ces seuils peut s'avérer très difficile. Ainsi, il faut quasi systématiquement recourir à des simulations numériques pour estimer ces seuils de percolation. A cet égard, les tout premiers travaux ayant considéré des problèmes de percolation d'un point de vue numérique sont issus de la physique et concernent des problèmes de conduction [111,118]. Depuis ces travaux et avec l'avènement de l'informatique moderne, des algorithmes et méthodes de simulations plus efficaces ont été étudiés. Plus particulièrement, une technique ayant retenu notre attention est empruntée là encore à la physique [106] : des algorithmes "union-find" sont utilisés pour suivre efficacement l'évolution des composantes connexes d'un graphe issu d'un problème de percolation.

Dans une moindre mesure, nos travaux sont aussi en lien avec [13], où des outils de géométrie stochastique sont utilisés pour poser et résoudre des problèmes d'optimisation liés à la planification économique de réseaux de télécommunication, en considérant certaines fonctionnelles de processus ponctuels.

Structure de la thèse

Cette thèse est divisée en six chapitres. Elle débute par deux chapitres introductifs : dans le premier chapitre, nous présentons les outils mathématiques (principalement des concepts de géométrie stochastique et de théorie de la percolation) qui s'avèreront utiles dans le reste de la thèse ; le chapitre 2 est quant à lui dédié à une présentation générale des communications sans fil et D2D dans les réseaux cellulaires. Un lecteur familier avec l'un ou l'autre de ces domaines pourra sauter le chapitre adéquat et commencer sa lecture du document par le chapitre 3, qui pose les bases de notre travail de manière plus concise et introduit notre approche pour modéliser et étudier la connectivité des réseaux D2D en environnement urbain. Dans les chapitres 4 et 5, nous utilisons notre modèle stochastique pour établir des conditions de faisabilité de connectivité à grande échelle pour des réseaux D2D en environnement urbain. Dans le dernier chapitre, nous introduisons notre modèle opérationnel de coûts pour analyser les conséquences économiques du déploiement du D2D dans les réseaux cellulaires. Nous appliquons ensuite ce modèle à l'étude d'un exemple d' übérisation des réseaux de télécommunications par l'utilisation du D2D.

General introduction

The fifth generation of mobile networks (5G) will have to face various challenges. Providing coverage and a satisfying quality of service for an unprecedented number of devices over large areas will be one of them. Indeed, according to a study from Ericsson [3], the total number of connected devices in the world reached circa 15 billion in 2015 and is expected to reach 30 billion in 2022. In particular, this explosion of the number of connected devices and the emergence of new and more and more data-intensive applications (internet browsing, online gaming, high-resolution video streaming and downloading, internet of things ...) have led to the following question: how to deal with this surge in data traffic, while keeping up with a satisfying quality of service?

In this regard, new ways of thinking the future cellular networks must be considered. One of the main technologies investigated in the literature to address the previous question is *device-to-device* (D2D). D2D consists in the possibility of short-range direct communications between two devices in a telecommunications network, without the need for the signal to be routed through additional network infrastructure. Applications of D2D in cellular networks are not only numerous but also very promising [8, 52].

A use case of particular interest for operators is the one of *uberisation* of networks. If a new telecommunications service provider having no (or very few) network infrastructure could build a network only relying on mobile devices, it would have at its disposal a complete network without having invested massively in network infrastructure. This could be a serious threat for traditional operators. Conversely, the possibility of building a network without making such investments could be a great opportunity for traditional operators in countries where they are not already present.

In the literature, the overwhelming majority of research articles concerning D2D either focus on possible applications or on particular technical issues (e.g. security, energy consumption, radio resources distribution, interference management ...). However, they rarely deal with the problem of modelling large-scale D2D networks. Good connectivity of such networks is often assumed to be established and no feasibility conditions of this connectivity are mentioned.

Thesis overview and research objectives

In this thesis, we are interested in new models of D2D networks in urban environments which allow us to study the question of their connectivity at large-scale as well as the economic consequences of the feasibility of large D2D networks for operators. In particular, we attempt to answer the following questions:

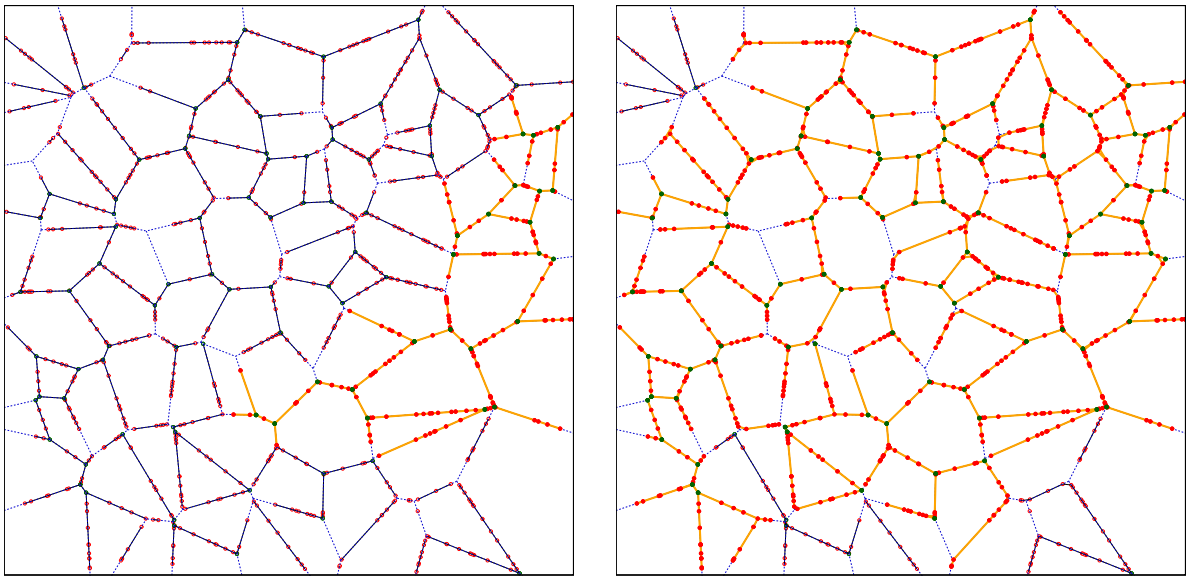
- How to take into account the spatial distribution of network users, the topology of the urban environment covered by the network as well as appropriate connectivity conditions between network nodes?
- What are the minimal parameters (e.g. the density of network users) allowing for good connectivity of the network and how to estimate them?
- How to quantify the investments required to set up a functional D2D network?
- What could be possible economic consequences of the deployment of D2D-based networks for operators?

To address the previous questions, we build and propose new stochastic models of D2D networks relying on tools borrowed from *stochastic geometry* and *percolation theory*.

Contributions of the thesis

In our work, we introduce new mathematical models of D2D networks in urban environments. Briefly, our modelling approach is as follows. A *street system* S being given by a *Poisson-Voronoi tessellation* (PVT) is the support of two point processes: one *Cox process* of network users distributed on the edges of S and one *Bernoulli process* of relays distributed on the vertices of S. Denoting by Z the superposition of these two point processes, the network is then modelled by a *random connectivity graph* with nodes being the atoms of Z and where an edge is drawn between two points of Z according to certain *connectivity conditions* ruling how network nodes are connected. *Percolation* of the network connectivity graph, i.e. the existence of an infinite connected component with positive probability, is then interpreted as good connectivity of the modelled D2D network.

In a first approach, we study a case where only so-called *line-of-sight* (LOS) connections along the edges of the supporting PVT S within a certain range are possible. In other words, two network nodes are connected if and only if they are sufficiently close to one another and located on the same edge of the PVT S, as illustrated by Figure (a). As one of the main mathematical contributions of the thesis, we studied the percolation of the associated connectivity graph and proved that phase-transitions between different *connectivity regimes* exist. Two particularly interesting regimes are called *relay-limited* and *relay-and-user-limited*. In the relay-limited regime, percolation of the graph can solely be ensured by relays (Bernoulli points), while network users (Cox points) are essential in the relay-and-user-limited regime. Resorting to numerical simulations, we estimated the *frontiers* between these regimes: knowing them is important for economic planning and allows to state feasibility conditions of good connectivity of D2D networks.



(a) Example of a simulated network connectivity graph when only LOS connections are possible.

(b) Example of a simulated network connectivity graph when supplementary NLOS connections are considered.

Figure: Examples of simulated connectivity graphs. The blue dashed lines represent the PVT street system S, the red points represent network users (Cox process), the green points represent relays (Bernoulli process). The largest connected component in the simulation window is emphasised by orange edges and vertices filled in their respective colours. Edges of the smaller connected components are highlighted in black. In both cases, the realisations of the street system, of the Cox process and of the Bernoulli process are the same.

In the case described above, since connections between network nodes are only possible along straight lines (edges of the PVT S), the additional Bernoulli relays at the crossroads (vertices of S) are crucial and constitute a bottleneck for percolation of the connectivity graph. We thus propose two variants of the previous model for tackling this issue.
On the one hand, we leave the geometric nature of the model unchanged but we refine the connectivity model. More precisely, in this new approach, line-of-sight connections are preserved and we add supplementary *non-line-of-sight* (NLOS) connections, allowing two network nodes to be connected if they are located on edges that are incident to a common vertex v of S and if the sum of their distances to v is less than a given threshold. This is supposed to model scattering and/or reflections on buildings at crossroads of the streets, see Figure (b). As the other main mathematical contribution of this thesis, we study the percolation of the connectivity graph arising from these new connectivity conditions.
On the other hand, we propose a more applied approach where the street system is thought to be "enlarged" and where intersections between streets (edges of the PVT S) are no longer punctual but have a positive surface. Doing so, we revise previous numerical estimates of the thresholds delimiting the connectivity regimes of the network graph.

We then go on to introduce a supplementary *cost model* to complete the previous numerical estimates and study possible *uberisation scenarios* of telecommunications net-

works where a "neo-operator" is willing to set up a fully functional network by relying on D2D only. Predicting and quantifying fundamental relations between crucial network parameters (density of users and relay points) and the possibility of deploying large-scale D2D networks, as well as providing examples of decision making tools based on these fundamental relations is the main operational contribution of this thesis.

Last but not least, numerical estimations of the critical parameters of the percolation models studied in this thesis are conducted via original algorithms, which are the main computer science contribution of our work and could be used in other percolation problems.

Related works

Before coming to our own work, we briefly introduce milestone works related to this thesis, which help to understand the research approach adopted therein. A more exhaustive review of the literature will be given in the subsequent chapters.

Mathematical models of telecommunications networks based on a percolation approach actually go a long way back. At the beginning of the 1960s, in his seminal work [55], Gilbert already had the idea of modelling nodes of a wireless telecommunications network by the atoms of a homogeneous Poisson point process in the plane $\mathbb{R}^2$. Denoting the (deterministic) range of communications by $R > 0$, the network was then modelled by a connectivity graph where any two points of the previous planar Poisson process are joined by an edge if they are within distance R of each other. For $N \geq 1$, Gilbert then studied the probability $P(N)$ that there exists a connected component of this graph containing at least N points. Resorting to numerical simulations, Gilbert underlined the following phase transition of the network: above a critical value E_c for the coverage zone $E = \pi R^2$ of a network node, the probability $P(\infty) := \lim_{n\uparrow\infty} P(N)$ that there exists some infinite connected component in the network becomes positive. Gilbert's approach gave birth to the modern theory of continuum percolation. Today, (continuum) percolation theory remains a very active field of research with numerous applications to various areas of science and, in particular, telecommunications.

Stochastic models of real street systems with applications to telecommunications networks have been investigated in [57–60]. The main idea developed in these works is that real street systems can be replaced by statistically equivalent random tessellations having the same average geometric properties. This in turn allows to considerably simplify economic planning of networks and optimisation problems related to network design (for instance, how to spread cables in an access network). Statistical procedures aiming to find the best fit for an extraction of a real-world city map among several tessellation models have been investigated in the PhD thesis [32].

Building on the previous approach and on Gilbert's initial idea, it is no wonder that refined geometric models of telecommunications networks appeared. In such models, nodes of the network cannot be located anywhere in the plane but are rather constrained to be supported by a random support such as a random tessellation. Keeping a Poissonian

spreading of network nodes, one is thus confronted to doubly stochastic Poisson point processes, also known as Cox processes in the literature. Percolation of such Cox processes has been studied for the first time in [70] and an application to models of telecommunications network has been investigated in [29].

In most percolation problems, analytical expressions of critical thresholds delimiting the phase transition of the system are out of reach, while finding theoretical bounds can turn out to be technically challenging. As a matter of fact, one must quasi-systematically resort to numerical simulations to estimate such percolation thresholds. In this regard, the very first works considering percolation from a numerical perspective concerned conduction problems in Physics [111, 118]. Since then and with the advances of computer science, more efficient simulation methods and algorithms have been proposed. In particular, a technique which caught our attention is borrowed from the Physics literature [106], where so-called union-find algorithms allow to efficiently keep track of the connected components of a graph resulting from a percolation problem.

To a lesser extent, our work is also related to [13], where stochastic geometry tools are used to pose and solve optimization problems related to economic planning of telecommunications networks by considering appropriate functionals of point processes.

Structure of the thesis

This thesis is divided into six chapters. It starts with two general introductory chapters: in the first one, we present the mathematical tools (mainly concepts of stochastic geometry and percolation theory) used in the rest of the thesis ; the second one is dedicated to a general overview of wireless and device-to-device communications in cellular networks. A reader familiar enough with any of these domains can skip the appropriate introductory chapter and start reading from Chapter 3, which more straightforwardly lays the foundations of the work done during this thesis and introduces our approach for modelling and studying the connectivity of D2D networks in urban environments. In Chapters 4 and 5, we use our stochastic model to derive feasibility conditions of large-scale connectivity of D2D networks in urban environments. In the final and sixth chapter, we introduce our operational cost model for analysing the economic consequences of D2D deployment in cellular networks. As an example of application, we study a particular scenario of uberisation of a telecommunications network through D2D.

Chapter 1

Mathematical foundations

The definition of a good mathematical problem is the mathematics it generates rather than the problem itself.

Andrew Wiles

In this chapter, we review fundamental mathematical tools which will be needed throughout the thesis. As explained in the introduction, we will be interested in studying telecommunications network models using random connectivity graphs and random spatial patterns. It is therefore essential to introduce a formal and rigorous framework for the study of such objects. This framework is established in four different sections.

First, we begin with introducing general concepts of *stochastic geometry*, the branch of probability theory studying random spatial patterns. Then, these general concepts will bring us to introduce a class of objects that play a key role both in this thesis and in stochastic geometry: stationary point processes. Such stationary processes exhibit interesting properties that make them particularly amenable to analytical study. We then go on with introducing the general theory of objects called *random tessellations*. Such tessellations naturally occur in many applications where the *typology* of a territory plays a critical role and has to be modelled in a particular way. Finally, as we will be interested in the *connectivity* of random graphs in the rest of this thesis, we explore this notion from a mathematical point of view. In the literature, this problem has been formalised around a branch of probability called *percolation theory*. Foundations of this theory, with emphasis on some particular models, will be presented in the fourth section of this chapter.

1.1 Preliminaries from stochastic geometry

1.1.1 What is stochastic geometry?

Geometrical patterns surround us pretty much everywhere in nature, both at microscopic and macroscopic scales. In many areas of science and technology, these patterns appear as data sets, which need to be studied and analysed in greater detail. Common examples include:

- Networks in Telecommunications
- Crystal structures in Chemistry
- Biological tissues in Biology
- Material structures in Physics or Engineering

The complexity of such structures often requires a statistical analysis, which can only be performed using relevant mathematical models and methods. Conversely, one may be interested in generating random patterns which accurately reproduce geometric features of a pattern observed in nature. These simulated patterns may then be used to better understand the original pattern and its physical properties.

A suitable theory for both of these purposes is proposed by a whole branch of mathematics and probability theory studying random spatial patterns and structures: *stochastic geometry*. In the foreword to the first edition of the standard reference textbook [31], Kendall traces the origin of the discipline back to the early 1960s. Going even further back, one may relate the story of stochastic geometry to early problems of classical geometrical probability, (e.g. Buffon's needle, addressed in the 18th century).

Since then, the discipline has developed and remains a very active field of research. Textbooks introducing the reader to stochastic geometry are now numerous. The book [31] has been re-edited twice and presents a comprehensive and global overwiew of the field. More theoretical textbooks which also include recent developments are [82, 117]. A comprehensive and recent approach of the field can be found in [12]. Regarding the applications of stochastic geometry in telecommunications, many textbooks are available. Standard references on such applications are [10,11,64]. Newer textbooks, such as [21,73], underline that stochastic geometry applied to wireless network modelling still concentrates a great research effort.

1.1.2 General framework

Space of points

In stochastic geometry, one is interested in observing (possibly random) patterns of points located in some observation space $\mathbf{E}$, called *space of points*. Most of the time, this space is sufficiently "nice" and "regular" to avoid pathological issues when dealing with the properties of point processes defined thereon. A typical example encountered in applications is $\mathbb{R}^d$, for $d \geq 1$.

However, the need for new models and theoretical purposes have led mathematicians to look at point processes sometimes defined on rather abstract spaces (e.g. the set of closed subsets of $\mathbb{R}^d$ for the Euclidean topology). Though being abstract, these spaces still verify a certain number of properties and are often referred to as *LCSCH spaces* in the literature:

Definition 1.1.1 (LCSCH space)**.** A topological space $\mathbf{E}$ is called LCSCH if the following properties are satisfied:

- $\mathbf{E}$ is locally compact, i.e. every point of $\mathbf{E}$ has a compact neighbourhood
- $\mathbf{E}$ is second countable, i.e. its topology has a countable base

- **E** is Haudorff, or separated, i.e. distinct points of **E** have disjoint neighbourhoods

By standard toplogy arguments, it is easily seen that a consequence of the above properties is the following:

Corollary 1.1.1.1. *Every LCSCH space is a Polish space, i.e. has a countable dense subset and admits a complete metric.*

Proof. See [81, Theorem 5.3]. □

Note that the metric making **E** complete needs not be unique, but the framework we will be presenting is not based on any particular choice of one.

A typical example of LCSCH space is the d-dimensional Euclidean space $\mathbb{R}^d, d \geq 1$. From now onwards, we let **E** be an LCSCH space, endowed with its Borel σ-algebra $\mathcal{B}$, that is, the σ-algebra generated by the open sets of the topology of **E**. Elements of $\mathcal{B}$ will be called *Borel sets* or *Borel subsets* of **E**. A subset $B \subset \mathbf{E}$ is called *bounded* if it is *relatively compact*, i.e. the closure $\mathrm{cl}(B)$ of B is compact. We denote by $\mathcal{B}_c$ the set of all bounded Borel subsets of **E**.

Another consequence of the definition is the following one:

Corollary 1.1.1.2. *Every LCSCH space* **E** *can be covered by countably many relatively compact open sets. In other words,* **E** *can be covered by countably many sets of* $\mathcal{B}_c$*.*

Proof. See [117, Theorem 12.1.1] or [12, Lemma 1.1.4]. □

Counting measures

In the general theory of point processes, countable patterns of points in an LCSCH space **E** are identified with particular measures called *counting measures* on $(\mathbf{E}, \mathcal{B})$. Informally, letting δ_x denote the Dirac measure for $x \in \mathbf{E}$, this is done by considering any countable pattern of points as a sequence $(x_i)_{i\geq 1}$ of points of **E** and then identifying it with some measure $\mu = \sum_i k_i \delta_{x_i}$ on $(\mathbf{E}, \mathcal{B})$, where the $k_i \geq 0$'s are non-negative coefficients representing a possible multiplicity of points.

In a more rigorous way, this identification, with a proper definition of counting measures, is done as follows. First, we consider all measures μ on $(\mathbf{E}, \mathcal{B})$ which are *locally finite*, i.e. such that $\mu(B) < \infty$ for all bounded Borel sets $B \in \mathcal{B}_c$. Then, we define a counting measure as follows:

Definition 1.1.2 (Counting measure)**.** A counting measure μ is a locally finite measure taking only values in $\{0, 1, 2, \ldots, \infty\}$ (this latter value being only allowed on unbounded sets). The set of all counting measures on $(\mathbf{E}, \mathcal{B})$ is denoted by $\mathbb{M}$.

A convenient consequence of the previous definition is the following:

Corollary 1.1.2.1. *Any counting measure* μ *on* **E** *is* σ*-finite.*

Proof. This is a consequence of Corollary 1.1.1.2 and the fact that a counting measure, being locally finite, is finite on bounded sets. □

This theoretical framework now allows for a rigorous statement of the aforementioned identification between a countable pattern of points in $\mathbf{E}$ and a counting measure:

Proposition 1.1.3. *Any counting measure μ on an LCSCH space $\mathbf{E}$ can be written as:*

$$\mu = \sum_{i=1}^{N} \delta_{x_i}, \tag{1.1.1}$$

where $N \in \mathbb{N} := \{1, \dots, \infty\}$ and $(x_i)_{i=1,\dots,N}$ is a sequence of points of $\mathbf{E}$ without accumulation points.

Proof. See [78]. A stronger version of this result is proven in [117, Lemma 3.1.3]. □

A counting measure μ as represented by (1.1.1) is called *simple* if the x_i's are distinct. If not, then μ can be written as $\mu = \sum_{i=1}^{N'} k_i \delta_{x_i'}$, where the scalars $k_i \in \{1, 2, \dots\}$ are called the multiplicity of points and the x_i''s are distinct. Identifying a counting measure μ with a subset of $\mathbf{E}$ as done in (1.1.1), we will often use the notation $x \in \mu$ to mean $\mu(\{x\}) \geq 1$. In such a case, we shall say that x is an *atom* of μ.

Space of configuration of points

Since point processes can be seen as *random* patterns of points, we need to define a σ-algebra on the set of counting measures $\mathbb{M}$. We will consider the *evaluation σ-algebra* $\mathcal{M}$ generated by the *projections* (or *evaluation maps*):

$$\begin{aligned} \mathbb{M} &\to \{0, 1, 2, \dots, \infty\} \\ \mu &\mapsto \mu(B) \end{aligned}$$

with $B \in \mathcal{B}$ (we can even restrict ourselves to $B \in \mathcal{B}_c$ by a monotone class argument). In other words, $\mathcal{M}$ is the smallest σ-algebra making the above mappings measurable. The space $(\mathbb{M}, \mathcal{M})$ endowed with this evaluation σ-algebra is often called *space of configuration of points*. Note that this specific choice of σ-algebra allows for the following:

Proposition 1.1.4. *Let μ be a counting measure written as in (1.1.1). Then, the enumeration of the atoms x_i of μ can be chosen so as to make the mappings: $\mathbb{M} \ni \mu \mapsto x_i \in \mathbf{E}$ measurable for all i.*

Proof. See [117, Lemma 3.1.3]. □

1.1.3 Point processes

General definition

The most basic objects dealt with in stochastic geometry are *point processes*. Informally, such point processes are random patterns of points in an observation space $\mathbf{E}$. The theoretical general framework exposed in the previous subsection will now allow us to properly define those point processes.

From now on and throughout the rest of this chapter, unless stated otherwise, we adopt the general framework exposed in the previous subsection and let $\mathbf{E}$ be an LCSCH space endowed with its Borel σ-algebra $\mathcal{B}$. We also fix some probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

The definition of a point process is as follows:

Definition 1.1.5 (Point process)**.** A *point process* (PP) is a measurable mapping:

$$\Phi : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathbb{M}, \mathcal{M})$$

The *distribution* of Φ will be noted $\mathbb{P}_\Phi := \mathbb{P} \circ \Phi^{-1}$

From the definition, it is clear that a point process Φ can also be seen as a stochastic process $\Phi = \{\Phi(B)\}_{B\in\mathcal{B}}$ with state space $\mathbb{N}_0 := \{0, 1, \ldots\}$ and index B running through $\mathcal{B}$. Hence, by Kolmogorov's consistency theorem, the distribution of the point process Φ is characterised by the family of finite dimensional distributions $(\Phi(B_1), \ldots, \Phi(B_k))$, where $k \geq 1$ and $B_1, \ldots, B_k$ run over $\mathcal{B}$. Note once again that, thanks to a monotone class argument, it is sufficient to consider that $B_1, \ldots, B_k$ are bounded and so run over $\mathcal{B}_c$ (rather than $\mathcal{B}$).

Since a point process Φ is a random counting measure, the following definitions make sense:

Definition 1.1.6 (Fixed atom of a point process)**.** The point process Φ is said to have a *fixed atom* at $x_0 \in \mathbf{E}$ if $\mathbb{P}(\Phi(\{x_0\})) > 0$.

Definition 1.1.7 (Simple point process)**.** A point process Φ is said to be *simple* if $\Phi(w)$ is a simple counting measure almost surely. In other words:

$$\mathbb{P}(\forall x \in \mathbf{E}, \Phi(\{x\}) \leq 1) = 1 \tag{1.1.2}$$

Remark. Note that $\mathbf{E}$ is not necessarily countable and so the measurability of the event considered in (1.1.2) is not straightforward. This technical issue is dealt with by considering a system of partitions of $\mathbf{E}$ becoming finer and finer and whose diameter (for a metric making $\mathbf{E}$ complete) goes to 0. Such a system is called a *null-array of partitions* and its existence is in fact independent of the chosen metric, as long as it makes $\mathbf{E}$ complete. A formal definition and proof of this fact can be found in [78] or [12, Lemma 1.6.3].

Some characteristics of point processes

We now introduce some general characteristics of point processes. The first one is the *mean measure*:

Definition 1.1.8 (Mean measure)**.** The *mean measure* M of a point process Φ is defined by $\forall B \in \mathcal{B}, M(B) := \mathbb{E}[\Phi(B)]$. In other words, $M(B)$ is the expected number of points of Φ in B.

It is straightforward to check that M is indeed a well-defined measure on $(\mathbf{E}, \mathcal{B})$, which may be infinite on unbounded Borel subsets of $\mathbf{E}$.

The mean measure will play a great role in evaluating the expectation of integrals against Φ. Before coming to that point, we introduce two convenient characteristics of point processes:

Definition 1.1.9 (Void probability)**.** The *void probability* of a point process Φ is a set function $\nu_\Phi = \nu$ defined on $(\mathbf{E}, \mathcal{B})$ by $\nu(B) := \mathbb{P}(\Phi(B) = 0)$.

The void probability is of great interest thanks to the following fact:

Theorem 1.1.10 (Rényi's theorem)**.** *The probability distribution of a point process Φ is fully characterised by the family of void probabilities $\nu_\Phi(B)$ for all $B \in \mathcal{B}_c$.*

Proof. The proof is rather technical and can be found in [78, Theorem 3.3]. □

Another useful tool for calculations is the *Laplace functional*, defined as follows:

Definition 1.1.11 (Laplace functional)**.** The *Laplace functional* of a point process Φ is a functional $\mathcal{L}_\Phi = \mathcal{L}$ on the space of non-negative measurable functions $f : \mathbf{E} \to \mathbb{R}_+$ and defined by:

$$\mathcal{L}(f) := \mathbb{E}\left[e^{-\int_{\mathbf{E}} f d\Phi}\right] \tag{1.1.3}$$

This definition can be extended to all functions f for which the expectation in (1.1.3) is well-defined.

As for the void probability, we have the following:

Theorem 1.1.12. *The Laplace functional of a point process Φ fully characterises the distribution of Φ.*

Proof. See [10, Section 1.2]. □

We end up this general section on point processes with a very basic but useful formula for computing expectations of integrals against point processes: *Campbell's formula*.

Theorem 1.1.13 (Campbell's formula)**.** *Let Φ be a point process with mean measure M and let $f : \mathbf{E} \to \mathbb{R}$ be either non-negative or integrable with respect to M (i.e. $f \in L^1(M)$). Then the integral $\int_{\mathbf{E}} f d\Phi$ is almost surely well-defined and we have:*

$$\mathbb{E}\left[\int_{\mathbf{E}} f d\Phi\right] = \int_{\mathbf{E}} f dM$$

Proof. The proof is done by a monotone class argument. First, consider simple functions of the form $f = \sum_{i=1}^k a_i \mathbb{1}_{B_i}$ where $\forall\, 1 \leq i \leq k, a_i \geq 0$ and $B_i \in \mathcal{B}$. The formula holds for such functions. Then, for a general non-negative function f, the formula also holds by considering a sequence of simple functions converging to f and using the monotone convergence theorem. Finally, for $f \in L^1(M)$, write $f = f_+ - f_-$, where $f_+ = \max(f, 0)$ and $f_- = \max(-f, 0)$. The formula holds for both f_+ and f_-, so it holds for f. □

We now review one of the most celebrated class of point processes in the literature: the one of *Poisson point processes*.

1.1.4 Poisson point processes

Poisson point processes appear quite everywhere in the stochastic geometry literature, for many reasons.

On the theoretical perspective, Poisson point processes are sufficiently tractable to make lots of analytic computations doable by hand. Moreover, Poisson point processes feature good properties of stability against several geometric operations, such as superposition or thinning.

On a more applied perspective, Poisson point processes are easy to simulate numerically and provide good models for lots of geometrical patterns in nature. They also feature a fundamental property called *complete independence*, which informally means that points of Poisson point processes do not tend to interact with one another.
In the rest of this subsection, we only provide a global overview of the main properties of Poisson point processes. More details can be found in almost every stochastic geometry textbook ; some even being fully dedicated to this particular topic [85, 86].

Definition and first properties

Definition 1.1.14 (Poisson point process)**.** Let Λ be a (deterministic) locally finite measure on $(\mathbf{E}, \mathcal{B})$. A point process Φ is said to be a *Poisson point process* (PPP) *with intensity measure* Λ if the following conditions are satisfied:

1. For all $B \in \mathcal{B}$, $\Phi(B)$ is a Poisson random variable of parameter $\Lambda(B)$.
2. For all $k \in \mathbb{N}$ and all disjoint Borel sets $B_1, \dots, B_k \in \mathcal{B}$, the random variables $\Phi(B_1), \dots, \Phi(B_k)$ are independent. This condition is called the *property of complete independence*.

Clearly, the two conditions above suffice to characterise the finite dimensional distributions of a Poisson point process. Hence, by Kolmogorov's consistency theorem, they characterise the distribution of the Poisson point process itself.

A case of particular interest when $\Lambda(dx) = \lambda dx$ on $\mathbb{R}^d$ for some $0 < \lambda < \infty$, i.e. the intensity measure is a multiple of Lebesgue measure. In that case, we say that Φ is a *homogeneous Poisson point process*. The scalar λ is called the *intensity* of Φ.

The main characteristics of a Poisson point process are simple consequences of the definition above:

Proposition 1.1.15. *Let Φ be a Poisson point process with intensity measure Λ. Then we have the following:*

- *The mean measure of Φ is Λ.*
- *The void probability of Φ is defined by: $\forall B \in \mathcal{B}, \nu(B) = e^{-\Lambda(B)}$*
- *The Laplace functional of Φ is defined by:*

$$\mathcal{L}(f) = e^{-\int_{\mathbf{E}} (1 - e^{-f(x)}) \Lambda(dx)}.$$

for all f either non-negative or in $L^1(\Lambda)$.

Proof. The form of the mean measure and of the void probability are simple consequences of the definition. For the Laplace functional, the proof is done by a monotone class argument, as has been done in the proof of Theorem 1.1.13. See [31, Example 4.2] for more details. □

Another convenient property of Poisson point processes is the following:

Proposition 1.1.16. *Let Φ be a Poisson point process with intensity measure Λ. Let $B_1, \dots, B_k$ be disjoint and denote $W = \bigcup_{i=1}^k B_i$. Then, for all $n \in \mathbb{N}$ and $n_1, \dots, n_k$ such that $\sum_{i=1}^k n_i = n$, we have:*

$$\mathbb{P}\{\Phi(B_1) = n_1, \dots, \Phi(B_k) = n_k \,|\, \Phi(W) = n\} = \frac{n!}{n_1! \dots n_k!} \frac{1}{\Lambda(W)^n} \prod_{i=1}^k \Lambda(B_i)^{n_i}$$

Proof. This is a straightforward consequence of the definition. □

The above conditional distribution is a multinomial distribution. Thus, given that there are n points of a Poisson point process Φ in some observation window W, these points are i.i.d. distributed according to the distribution $\frac{\Lambda(\cdot)}{\Lambda(W)}$. This fact is very useful for two purposes: on the one hand, proving that Poisson point processes do exist and construct them; on the other hand, numerically simulate Poisson point processes. More details about the existence and construction of Poisson point processes can be found in [86, Section 3.2], while details about numerical simulations can be found in [31, Section 2.5].

Finally, one may wonder when a Poisson point process is simple or has a fixed atom. This can easily be deduced from simple conditions on the intensity measure:

Proposition 1.1.17. *Let Φ be a Poisson point process with intensity measure Λ. Then:*

1. *Φ has a fixed atom at $x_0 \in \mathbf{E}$ if and only if Λ has an atom at x_0, i.e. $\Lambda(\{x_0\}) > 0$.*
2. *Φ is simple if and only if Λ is non-atomic, i.e. $\forall x \in \mathbf{E}, \Lambda(\{x\}) = 0$.*

Proof. The first part is a consequence of the definition. The second part is proved using the conditional distribution of Poisson points. See [10, Proposition 1.1.3] for more details. □

A straightforward consequence of the previous result is that an homogeneous Poisson point process is simple.

Characterisations of Poisson point processes

In addition to the definition and the Laplace functional, Poisson point processes can be characterised by other means. The first one is based on the form of the void probability:

Theorem 1.1.18. *Let Φ be a simple point process on $\mathbf{E}$. Then Φ is a Poisson point process if and only if there exists a locally finite non-atomic measure Λ on $(\mathbf{E}, \mathcal{B})$ such that $\forall B \in \mathcal{B}_c, \mathbb{P}(\Phi(B) = 0) = e^{-\Lambda(B)}$.*

Proof. The direct part follows from the definition and Proposition 1.1.17. The converse part is a consequence of Rényi's theorem. □

In particular, a consequence of the previous result is that any simple point process Φ such that $\Phi(B)$ is a Poisson random variable for every $B \in \mathcal{B}_c$ is a Poisson point process.

There is a more subtle way to characterise Poisson point processes: the *property of complete independence*. Recall that a point process Φ is said to have the property of complete independence if for all disjoint bounded Borel sets $B_1, \dots, B_k$, the random variables $\Phi(B_1), \dots, \Phi(B_k)$ are independent.

Theorem 1.1.19. *Let Φ be a point process on $\mathbf{E}$ without fixed atoms. Then Φ is a Poisson point process if and only if Φ is simple and has the property of complete independence.*

The result above explains why Poisson point processes are often chosen in models where there are no interactions between points.

Proof. The direct part comes from the definition and the fact that if Φ is a Poisson point process without fixed atoms, its intensity measure Λ is non-atomic and so Φ is simple. For the converse part, see [37, Lemma 2.4.IV] □

Operations preserving Poisson point processes

As mentioned earlier, Poisson point processes are sufficiently tractable to do lots of analytic computations by hand. Moreover, Poisson point processes are stable under certain operations. In this section, we define such operations and give without proofs the statements that Poisson point processes are stable under them. For more details, the interested reader can refer to [86, Theorem 3.3], [86, Chapter 5] or [12, Section 2.2].

We begin with *superposition*. If $(\Phi_k)_{k\in\mathbb{N}}$ is a sequence of point processes on the same LCSCH space $\mathbf{E}$, their superposition is defined by $\Phi = \sum_{k\in\mathbb{N}} \Phi_k$, where the sum is to be understood as a sum of measures. Note that this superposition may not be a point process itself, and one may wonder when it is the case. This point is detailed in [12, Proposition 2.2.1] and in the subsequent corollaries.
Nevertheless, we have the superposition of independent Poisson point processes remains a Poisson point process:

Theorem 1.1.20 (Superposition theorem)**.** *The superposition of independent Poisson point processes $(\Phi_k)_{k\in\mathbb{N}}$ with intensity measures $(\Lambda_k)_{k\in\mathbb{N}}$ is a Poisson point processes with intensity measure $\sum_{k\in\mathbb{N}} \Lambda_k$ if and only if this latter measure is locally finite.*

The second operation worth considering is *thinning*. This consists in suppressing some of the points of a point process. The most basic kind of thinning encountered in the stochastic geometry literature is *independent thinning*, where the decision of erasing or retaining a point is taken independently from all the other points.
More formally, independent thinning can be defined as follows. Consider some point process $\Phi = \sum_{k\in\mathbb{N}} \delta_{X_k}$ defined on some LCSCH space $\mathbf{E}$, $(U_k)_{k\in\mathbb{N}}$ i.i.d. random variables that are independent of Φ and uniformly distributed on $[0,1]$ and some measurable function $p : \mathbf{E} \to [0,1]$ called *retention function*. The independent thinning of Φ by p is defined by:

$$\tilde{\Phi} := \sum_{k\in\mathbb{N}} \mathbb{1}\{U_k \leq p(X_k)\}\delta_{X_k}$$

In other words, each point X_k of Φ is kept independently from all other points with probability $p(X_k)$.

The independent thinning of a point process remains a point process: its mean measure and Laplace functional can be computed via the ones of the original un-thinned process, see [12, Proposition 2.2.6]. More importantly, Poisson point processes are stable under independent thinning:

Theorem 1.1.21. *The independent thinning of a Poisson point process of intensity measure Λ with retention function p is a Poisson point process of intensity measure $\tilde{\Lambda}(dx) := p(x)\Lambda(dx)$.*

The last operation we consider is *random transformation of points*, which informally consists in a random and independent displacing of each point of a point process to some new location, possibly in some different LCSCH observation space $\mathbf{E}'$, according to some probability kernel p.

Recall from [79, Section 1.3] that a probability kernel p from some LSCSH space $(\mathbf{E}, \mathcal{B})$ to another LSCSH space $(\mathbf{E}', \mathcal{B}')$ is a function $p : \mathbf{E} \times \mathcal{B}' \to [0, 1]$ such that for all $x \in \mathbf{E}$, $p(x, \cdot)$ is a probability measure on $(\mathbf{E}', \mathcal{B}')$ and for all $B' \in \mathcal{B}'$, $p(\cdot, B')$ is a non-negative measurable-function on $(\mathbf{E}, \mathcal{B})$. The transformation of a point process $\Phi = \sum_{k\in\mathbb{N}} \delta_{X_k}$ by the probability kernel p is a point process Φ^p on $\mathbf{E}'$ given by:

$$\Phi^p = \sum_{k\in\mathbb{N}} \delta_{Y_k},$$

where the (Y_k) are points of $\mathbf{E}'$, independent given Φ, with conditional distribution $\mathbb{P}(Y_k \in B' \,|\, \Phi) = p(X_k, B')$ whenever $B' \in \mathcal{B}'$. We assume that this indeed defines a point process on $\mathbf{E}'$ (see [12, Lemma 2.2.14] for a proof of this fact).

Poisson point processes also are stable under transformation by probability kernels: this is referred to as *displacement theorem* in the literature.

Theorem 1.1.22 (Displacement theorem)**.** *Let Φ be a Poisson point process on some LSCSH space $\mathbf{E}$, with intensity measure Λ and let p be a probability kernel from $(\mathbf{E}, \mathcal{B})$ to another LSCSH space $(\mathbf{E}', \mathcal{B}')$. Let Λ' be the measure on $\mathbf{E}'$ defined by:*

$$\forall B' \in \mathcal{B}', \Lambda'(B') = \int_{\mathbf{E}} p(x, B')\Lambda(dx),$$

If Λ' is a locally finite measure on $\mathbf{E}'$, then the random transformation of Φ by the probability kernel p is a Poisson point process on $\mathbf{E}'$ with intensity measure Λ'.

1.1.5 Other celebrated point processes

Poisson point processes are not the only ones considered in the stochastic literature. In this subsection, we review two other celebrated types of point processes which will appear, at some point, in this thesis.

One-point and binomial point processes

Let X be an $\mathbf{E}$-valued random variable with probability distribution $\mathbb{Q}$, i.e. $\mathbb{P}(X \in B) = \mathbb{Q}(B)$ whenever $B \in \mathcal{B}$. Then $\Phi = \delta_X$ is a point process called *one-point process*. One-point processes arise often in geometric probability, especially in the case where $\mathbf{E} = W$ is some compact of $\mathbb{R}^d$ and $\mathbb{Q}$ is the uniform distribution on W.

A natural extension of a one-point process is done by considering the superposition of independent one-point point processes with common underlying probability distributions $\mathbb{Q}$ for their points. This is the so-called *binomial point process*:

Definition 1.1.23 (Binomial point process)**.** Let $n \in \mathbb{N}$ and let $X_1, \dots, X_n$ be i.i.d. $\mathbf{E}$-valued random variables with common distribution $\mathbb{Q}$. The superposition $\Phi := \sum_{k=1}^{n} \delta_{X_k}$ of the one-point processes given by the X_k's is called a *binomial point process*. $\mathbb{Q}$ is sometimes called the *sampling distribution* of Φ in the literature.

More details about binomial point processes can be found in [31, Section 2.2]. Such binomial point processes are closely related to Poisson point processes when the number of points n becomes random (see [86, Section 3.2]).

Cox point processes

Cox point processes are extensions of Poisson point processes when the intensity measure Λ is no longer deterministic but random. For this reason, they are often referred to as *doubly stochastic* Poisson point processes. Refining the simpleness of the Poisson assumption, Cox-based models play a critical role in many areas such as telecommunications, finance and even biology.

We only give a formal definition, the form of the mean measure and of the Laplace functional. The interested reader can refer to [86, Section 13] or [31, Section 5.2] for more detailed accounts of properties of Cox point processes.

Definition 1.1.24 (Cox point process)**.** Let Λ be a random measure on some LCSCH space $(\mathbf{E}, \mathcal{B})$. A point process Φ defined on the same probability space as Λ is said to be a *Cox point process* directed (or driven) by Λ if the conditional distribution of Φ given Λ is the one of a Poisson point process with intensity measure Λ. Λ is called the *directing measure* or the *driving measure* of Φ.

Proposition 1.1.25 (Mean measure and Laplace transform of a Cox point process)**.** *Let Φ be a Cox point process on some LCSCH space $(\mathbf{E}, \mathcal{B})$ with driving measure Λ. Then we have the following:*

1. *The mean measure M of Φ is given by: $\forall B \in \mathcal{B}, M(B) = \mathbb{E}[\Lambda(B)]$.*
2. *The Laplace functional $\mathcal{L}$ of Φ is defined by $\mathcal{L}(f) = \mathbb{E}\left[e^{-\int_{\mathbf{E}}(1-e^{-f(x)})\Lambda(dx)}\right]$ for all f either non-negative or such that the right-hand side of the previous equality is finite when f is replaced by $|f|$.*

Proof. This is a simple consequence of the definition and of Proposition 1.1.15. □

1.2 Stationary point processes

Stationary point processes, i.e. point processes whose distribution is translation-invariant, play a key role in stochastic geometry. Indeed, the stationarity assumption may allow one to shift all the atoms of a point process by a common vector, chosen in a convenient way. Going further, one may be interested in *point conditioning* questions: for instance, one may want to study a point process Φ given that Φ has an atom at some fixed location, say the origin 0. In other words, one may want to study events of the kind

$$\mathbb{P}(\Phi \in \cdot \,|\, 0 \in \Phi). \tag{1.2.1}$$

If 0 is an atom of Φ, the conditioning event in the former probability makes sense. But what happens when this is not the case? For instance, in the simple case where Φ is a homogeneous Poisson point process on $\mathbb{R}^d$ with intensity $\lambda > 0$, we have $\mathbb{P}(\Phi(\{0\}) > 0) = 0$, and so the conditional probability in (1.2.1) seems to make no sense.

To tackle such issues, a more general approach based on measure-theoretic arguments has been proposed for stationary point processes and allows one to define a *Palm probability* under which the origin 0 can be seen as a *typical point* of Φ. We now present a global overview of this approach.

1.2.1 Stationary point processes and stationary framework

From now onwards, we consider the case where the observation space is the d-dimensional Euclidean space for some $d \geq 1$, i.e. $\mathbf{E} = \mathbb{R}^d$. We endow $\mathbb{R}^d$ with its Borel σ-algebra $\mathcal{B} = \mathcal{B}(\mathbb{R}^d)$. As stated earlier, this is an example of an LCSCH space.

Stationary processes

We will now be interested in *stationary* (point) processes. Recall that a stochastic process $\{X(x) : x \in \mathbb{R}^d\}$ is said to be *stationary* if its distribution is translation-invariant, which, by Kolmogorov's consistency theorem, is equivalent to the translation invariance of its finite dimensional distributions. In other words:

$$\forall n \in \mathbb{N}_0, \forall y \in \mathbb{R}^d, (X(x_1 + y), \ldots, X(x_n + y)) \overset{d}{=} (X(x_1), \ldots, X(x_n)), \tag{1.2.2}$$

where $\overset{d}{=}$ denotes equality in distribution.

Since a point process Φ can be seen as a stochastic process $\Phi = (\Phi(B))_{B \in \mathcal{B}}$, the former definition also applies to point processes.

When studying a family of point processes, the following terminology is often encountered in the literature:

Definition 1.2.1 (Jointly stationary point processes)**.** Say a family of point processes $(\Phi_i)_i$ are *jointly stationary* if the joint distribution of these point processes is invariant under the respective translation by any vector of $\mathbb{R}^d$.

Shifting on measures and functions

The traditional approach to the study of (jointly) stationary point processes is done by considering a *stationary framework* on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. This approach is a flow-based one and induces a notion of shifting on measures.

Therefore, we begin by defining a *shift* on the set of measures μ on $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ by:

$$\forall t \in \mathbb{R}^d, \forall B \in \mathcal{B}(\mathbb{R}^d),\ S_t\mu(B) \coloneqq \mu(B + t)$$

Given a measurable space $(\mathbb{K}, \mathcal{K})$, for every mapping $f : \mathbb{R}^d \to \mathbb{K}$, we denote:

$$\forall t \in \mathbb{R}^d,\ S_t f : x \mapsto f(x + t)$$

Remark.
- Let μ be any measure on $\mathbb{R}^d$. Then:

$$\forall t \in \mathbb{R}^d, \forall B \in \mathcal{B}(\mathbb{R}^d), \int_B f(x) S_t\mu(dx) = \int_{B+t} f(x-t)\mu(dx)$$

- If $\mu = \sum_i k_i \delta_{x_i} \in \mathbb{M}$, then $\forall t \in \mathbb{R}^d$, $S_t\mu = \sum_i k_i \delta_{x_i - t}$

In light of the last point, we can give an equivalent definition for the stationarity of point processes:

Definition 1.2.2 (Stationary point process)**.** Let Φ be a point process on $\mathbb{R}^d$. For $t \in \mathbb{R}^d$, denote by $S_t\Phi$ the point process resulting from the translation of all atoms of Φ by $-t$. We say that Φ is a *stationary point process* if:

$$\mathbb{P}_\Phi \overset{d}{=} \mathbb{P}_{S_t\Phi}, \qquad \text{whenever } t \in \mathbb{R}^d \tag{1.2.3}$$

In other words, for all $t \in \mathbb{R}^d$, the distribution of $S_t\Phi$ is the one of Φ.

We now introduce the concept of flow:

Definition 1.2.3 (Flow)**.** A family $\{\theta_t\}_{t\in\mathbb{R}^d}$ is called a *flow* on the measurable space $(\Omega, \mathcal{A})$ if the following properties are satisfied :

- $\forall t \in \mathbb{R}^d$, $\theta_t : \Omega \to \Omega$ is bijective
- $\forall s, t \in \mathbb{R}^d$, $\theta_{t+s} = \theta_t \circ \theta_s$ (hence $\theta_0 = \mathrm{id}_{\mathbb{R}^d}$ and $\forall t \in \mathbb{R}^d$, $\theta_t^{-1} = \theta_{-t}$)
- The mapping $(\mathbb{R}^d, \Omega) \ni (t, \omega) \mapsto \theta_t(\omega)$ is $\mathcal{B} \otimes \mathcal{A}$-measurable

To properly define a stationary framework, we need to make sure that the objects we will consider are somehow going to be compatible with the underlying flow. This induces the following definitions:

Definition 1.2.4 (Compatible stochastic process)**.** A stochastic process $X = \{X(t) : t \in \mathbb{R}^d\}$ is said to be *compatible* with the flow $\{\theta_t\}_{t\in\mathbb{R}^d}$ if:

$$\forall t \in \mathbb{R}^d,\ X \circ \theta_t = S_t X$$

Note that since a point process Φ can be seen as a stochastic process $\Phi = (\Phi(B))_{B\in\mathcal{B}}$, this definition also applies to point processes. Thus, if Φ is compatible with the flow $\{\theta_t\}$ we have:

$$\Phi(\theta_t(\omega))(B) = S_t\Phi(\omega)(B) = \Phi(\omega)(B + t)$$

for all $\omega \in \Omega, B \in \mathcal{B}, t \in \mathbb{R}^d$.

Definition 1.2.5 (Preserving flow, Stationary framework)**.** We say the flow $\{\theta_t\}_{t\in\mathbb{R}^d}$ is $\mathbb{P}$-*preserving* if:

$$\forall t \in \mathbb{R}^d, \mathbb{P} \circ \theta_t^{-1} = \mathbb{P}$$

i.e. $\forall t \in \mathbb{R}^d, \forall A \in \mathcal{A}$, $\mathbb{P}(\{\omega \in \Omega, \theta_t(\omega) \in A\}) = \mathbb{P}(A)$
In that case, we say that $(\Omega, \mathcal{A}, \{\theta_t\}_{t\in\mathbb{R}^d}, \mathbb{P})$ is a *stationary framework*.

1.2.2 Palm probability in the stationary framework

Throughout the rest of this subsection, we assume that $(\Omega, \mathcal{A}, \{\theta_t\}_{t\in\mathbb{R}^d}, \mathbb{P})$ is a stationary framework, i.e. $\{\theta_t\}_{t\in\mathbb{R}^d}$ is a $\mathbb{P}$-preserving flow on $(\Omega, \mathcal{A})$.

Consider some stationary point process Φ which is compatible with the flow $\{\theta_t\}_{t\in\mathbb{R}^d}$. By measure-theoretic arguments based on Haar's theorem [109, Theorem 2], we have the following:

Proposition 1.2.6. *If Φ is a stationary point process on $\mathbb{R}^d$, the mean measure M of Φ is proportional to Lebesgue measure, i.e. :*

$$\exists\, 0 \leq \lambda \leq \infty; \forall B \in \mathcal{B}, \mathbb{E}\left[\Phi(B)\right] := M(B) = \lambda|B| \tag{1.2.4}$$

The scalar λ is called the intensity *of Φ. It represents the mean number of points of Φ in an observation window of Lebesgue measure* 1.

Proof. This is just a consequence of Haar's theorem, which implies that every translation-invariant measure on $\mathbb{R}^d$ is a multiple of Lebesgue measure. □

Remark.

- A homogeneous Poisson point process is stationary
- A Poisson point process on $\mathbb{R}^d$ is actually stationary if and only if it is homogeneous (see [12, Proposition 6.1.18]).
- A stationary point process on $\mathbb{R}^d$ with finite intensity does not have fixed atoms (see [12, Lemma 6.1.19]).

From now onwards, we will only consider stationary point processes with positive and finite intensity, i.e. $0 < \lambda < \infty$.

We now introduce a measure on $\mathbb{R}^d \times \Omega$, as follows:

Definition 1.2.7 (Campbell-Matthes measure)**.** We define the Campbell-Matthes measure $\mathcal{C}$ on $(\mathbb{R}^d \times \Omega, \mathcal{B}(\mathbb{R}^d) \otimes \mathcal{A})$ by defining it on rectangles:

$$\forall (B, A) \in \mathcal{B}(\mathbb{R}^d) \times \mathcal{A}, \mathcal{C}(B \times A) := \mathbb{E}\left[\int_{\mathbb{R}^d} \mathbb{1}_{\{x\in B\}} \mathbb{1}_{\{\theta_x(\omega)\in A\}} \Phi(dx)\right]$$

Note that some work is required to extend the above definition to the whole product space $\mathbb{R}^d \times \Omega$, see [12, Proposition 6.1.20] for more details.

Nevertheless, it is easily checked that the Campbell-Matthes measure is invariant under translations on its first argument, so that $\forall (B, A) \in \mathcal{B}(\mathbb{R}^d) \times \mathcal{A}, \mathcal{C}(B \times A) = constant(A)|B|$. This induces the following definition:

Definition 1.2.8 (Palm probability)**.** We can define a probability measure $\mathbb{P}^0$ on $(\Omega, \mathcal{A})$, called *Palm probability* of the point process Φ, given by:

$$\forall A \in \mathcal{A}, \forall B \in \mathbb{R}^d \text{ such that } 0 < |B| < \infty, \mathbb{P}^0(A) = \frac{\mathcal{C}(B \times A)}{\lambda|B|} = \frac{constant(A)}{\lambda}$$

$\mathbb{E}^0$ will denote the expectation with respect to Palm probability.

We can thus summarise the definition of the Palm probability of Φ as follows: the Palm probability of Φ is the *unique* probability measure $\mathbb{P}^0$ on $(\Omega, \mathcal{A})$ given by

$$\mathbb{P}^0(A) = \frac{1}{\lambda |B|} \mathbb{E}\left[\int_{\mathbb{R}^d} \mathbb{1}\{x \in B\} \mathbb{1}\{\theta_x \in A\} \Phi(dx) \right]$$

whenever $A \in \mathcal{A}$ and $B \in \mathcal{B}$ is a Borel set of finite non-null Lebesgue measure $|B|$. The definition of $\mathbb{P}^0$ does not depend on the choice of B.

Remark. In a non-stationary context with a general (non-stationary) point process Φ, similar ideas can be developed but the framework is different. Indeed, relaxing the stationary assumption on Φ does not allow to introduce a unique Palm probability on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ anymore. One can however introduce the *Palm distributions* of Φ, which form a family of probability measures $\{\mathbf{P}_x(\cdot)\}_{x \in \mathbf{E}}$ on $(\mathbb{M}, \mathcal{M})$. These Palm distributions are also very much linked to *point conditioning questions* and can be seen, under certain assumptions, as the conditional distribution of Φ, given that Φ has an atom at some given location $x \in \mathbf{E}$. More details on Palm calculus in the non-stationary context can be found in [12, Chapter 3] or [31, Section 4.4].

1.2.3 Key properties of Palm probability

For the sake of completeness, we shall now mention three key properties of the Palm probability of a point process and omit their proofs. The interested reader can refer to [12, Chapter 6] for the proofs.

The most remarkable property of Palm probability is the following:

Proposition 1.2.9. *Under Palm probability, Φ has an atom at 0 almost surely, i.e.*

$$\mathbb{P}^0(\Phi(\{0\}) > 0) = 1$$

The former induces the notion of *typicality*: under Palm probability, the origin 0 can be considered, for ergodicity reasons, to be a representative point of the point process Φ, which will be called a *typical point* of Φ. Further references about ergodic theory for point processes are [10, Section 1.6] and [38, Section 12].

Another key result of Palm calculus in the stationary context is the following analogue of Campbell's formula:

Theorem 1.2.10 (Campbell-Little-Mecke-Matthes (CLMM))**.** *Let $(\Omega, \mathcal{A}, \{\theta_t\}, \mathbb{P})$ be a stationary framework and let Φ be a stationary process on $\mathbb{R}^d$, compatible with the flow and with finite positive intensity $0 < \lambda < \infty$. Let $\mathbb{P}^0$ be the Palm probability of Φ. Then for any non-negative measurable function f on $\mathbb{R}^d \times \Omega$, we have:*

$$\mathbb{E}\left[\int_{\mathbb{R}^d} f(x, \theta_x) \Phi(dx) \right] = \lambda \int_{\mathbb{R}^d} \mathbb{E}^0 \left[f(x, \omega) \right] dx \tag{1.2.5}$$

This result extends to all functions f when either of the two sides of the above equality is finite when f is replaced by $|f|$.

We end up with another useful application of Palm stationary calculus: having defined a Palm probability on $(\Omega, \mathcal{A})$ allows one to jointly analyse two different point processes on the same probability space. This gives rise to the so-called *mass-transport formula*:

Theorem 1.2.11 (Mass-transport formula)**.** *Let $(\Omega, \mathcal{A}, \{\theta_t\}, \mathbb{P})$ be a stationary framework and let Φ, Φ' be stationary processes on $\mathbb{R}^d$, compatible with the flow and with finite positive intensities λ, λ' respectively. Let $\mathbb{P}^0$ and $\mathbb{P}^{0'}$ be the respective Palm probabilities of Φ and Φ'. Then for any non-negative measurable function g (not necessarily compatible with the flow) on $\mathbb{R}^d \times \Omega$, we have:*

$$\lambda \mathbb{E}^0 \left[\int_{\mathbb{R}^d} g(y, \omega) \Phi'(dy) \right] = \lambda' \mathbb{E}^{0'} \left[\int_{\mathbb{R}^d} g(-x, \theta_x) \Phi(dx) \right] \tag{1.2.6}$$

The terminology can be justified by interpreting $g(y, \omega)$ as the amount of mass sent from the atoms of Φ to the atoms of Φ'. The atoms of Φ' receive this mass, hence the negative sign on $g(-x, \theta_x)$. Note that if we apply the formula to $\Phi' = \Phi$ we obtain a *conservation of mass*:

$$\lambda \mathbb{E}^0 \left[\int_{\mathbb{R}^d} g(y, \omega) \Phi(dy) \right] = \lambda \mathbb{E}^0 \left[\int_{\mathbb{R}^d} g(-x, \theta_x) \Phi(dx) \right]$$

In other words, in a stationary framework, the mass leaving Φ is equal to the mass entering Φ. Stationarity is actually often explained in this way in the Physics literature.

Palm stationary calculus is a must-go in every theoretical stochastic geometry textbook, and many more key formulae and results are available. In particular, the most celebrated are the *inversion formula* (allowing one to recover the stationary expectation $\mathbb{E}f$ of a functional f in terms of its Palm expectation) and the *Neveu exchange formula* (relating the expectations of a same functional f under the Palm probabilities of two different point processes). For more details on such matters, the interested reader can refer to [12, Chapter 6] or [10, Chapter 4].

1.2.4 Palm probability of Poisson and Cox point processes

When it comes to Palm probability, the case of Poisson point processes is of particular interest. This is one of the most celebrated results of Palm theory, due to Slivnyak. We shall state it without proof:

Theorem 1.2.12 (Slivnyak)**.** *A stationary point process Φ with positive and finite intensity on $\mathbb{R}^d$ is a Poisson point process if and only if its distribution under the Palm probability is equal to the distribution of the original process Φ with an extra atom at the origin 0. In other words, Φ is a Poisson point process if and only if:*

$$\mathbb{P}_\Phi \overset{d}{=} \mathbb{P}^0_{\Phi + \delta_0}$$

In a stationary context, the case of Cox point processes is also of particular interest. First, note that stationarity of a Cox point process can easily be deduced from the stationarity of its driving measure, as summarised by the following result:

Proposition 1.2.13. *Let Φ be a Cox point process on $\mathbb{R}^d$ with driving measure Λ. Then Φ is stationary if and only if Λ is a stationary random measure, i.e. the probability distribution of Λ is invariant under all translations of $\mathbb{R}^d$.*

Proof. This is a simple consequence of the fact that for all $t \in \mathbb{R}^d$, $S_t\Phi$ is a Cox process driven by $S_t\Lambda$. □

Remark. In the same way, a Cox point process is isotropic if and only if its driving measure is isotropic (i.e. its probability distribution is invariant under all rotations of $\mathbb{R}^d$), see [120, Theorem 5.5].

Since Cox point processes are doubly stochastic Poisson point processes, it is no wonder that an analogue of Slivnyak's theorem is available in the Cox case. A bit more work is however required, as one needs to define a *Palm version* of the driving measure, see [120, Equation (5.3)] for more details on this matter. The equivalent of Slivnyak's result for Cox point processes is the following theorem:

Theorem 1.2.14. *Let Φ be a stationary Cox point process on $\mathbb{R}^d$ with driving measure Λ. Denote by Λ^0 a Palm version of Λ. Then the distribution of Φ under its Palm probability is that of a Cox point process driven by Λ^0 with an additional atom at the origin 0.*

1.3 Tessellations and random tessellations

In the previous subsections, we reviewed some of the theory of point processes. These point processes can be used to model particle systems or point systems (such as location of users in a place) but do not say much about the *topology* of the environment where these systems live. In many applications, the topological configuration of an environment where some system lives is of great importance in performance assessment or in the behaviour of the system itself. To study the topological configuration of an environment, one needs to resort to suitable mathematical models: this is done by considering objects called *tessellations* or often referred to as *mosaics*, the latter term being more suggestive. In a stochastic geometry context, it then makes sense to consider *random* mosaics as models for theoretical study and spatial statistics.

1.3.1 Tessellations

A vast mathematical literature is available on (random) tessellations, from books entirely dedicated to the topic [102, 103, 108] to chapters of more general (stochastic) geometry books [31, 117]. Though the general definition varies a bit here and there, authors seems to globally agree on what a generic framework should be: a division of a space into polyhedra. We formalise this as follows:

Definition 1.3.1 (Tessellation)**.** Let $d \geq 1$ and let $S \subseteq \mathbb{R}^d$ be a bounded region of $\mathbb{R}^d$ or $\mathbb{R}^d$ itself. A *tessellation* of S is a countable collection $(C_i)_{i\geq 1}$ of convex compact connected subsets of $\mathbb{R}^d$ such that:

- $S = \bigcup_{i\geq 1} C_i$, i.e. the C_i's form a covering of S.

- $\forall i \geq 1, \dim \mathring{C}_i = d$, i.e. C_i is full-dimensional.
- $\forall i \neq j, \mathring{C}_i \cap \mathring{C}_j = \varnothing$ i.e. the C_i's are non-overlapping.

Remark.
- In the stochastic geometry literature, the cells of the tessellation are most of the time assumed to be *convex*, but tessellations with non-convex cells are also of noticeable importance in the literature. We won't need to resort to such tessellations in this thesis and thus include the convexity assumption in the definition.
- In the literature, other conditions may be added to the previous definition. The one appearing most often is that the collection $(C_i)_{i\geq 1}$ has to be *locally finite*, i.e. for all $B \subset \mathbb{R}^d$ bounded, $\#\{i \geq 1,\, B \cap C_i \neq \varnothing\} < \infty$.

Assuming, as described above, that the collection $(C_i)_{i\geq 1}$ is locally finite, one can derive from the tessellation itself a set of *vertices* (also called *0-faces* in the literature) and of *edges* (also called *1-faces*), thus allowing to see the tessellation as a (possibly infinite) graph. More details on this can be found in [102].

Seeing the tessellation as a graph can induce another tessellation called the *dual tessellation*. It is defined in the following way:

Definition 1.3.2 (Dual tessellation)**.** Let $\Xi = (C_i)_{i\geq 1}$ be a tessellation. To each cell C_i of Ξ is associated a *centroid* c_i (usually the barycentre of the cell). Then create an undirected graph with vertex set $V = \{c_i, i \geq 1\}$ and setting an edge $(c_i, c_j), j \neq i$ if and only if C_i and C_j share a boundary edge. The tessellation derived from this graph is called the *dual tessellation* of Ξ.

1.3.2 Random tessellations

In stochastic geometry, one may be interested in *random tessellations*. The question is: how to consider such objects in a rigorous mathematical framework? Several approaches are possible but all of them, at some point, refer to the theory of *random closed sets*. Indeed, it makes sense to think that a random tessellation is just a tessellation whose cells are *random*. Since the cells are compact in $\mathbb{R}^d$, they are closed, which makes them *random closed sets*.

The general theory of random (closed) sets has extensively been studied in the literature is a must-go in almost every stochastic geometry textbook. Results from random set theory will not be needed in the rest of this thesis. For more details on such matters, the interested reader can refer to [12, 31, 98, 101, 108, 117].

As mentioned above, a random tessellation can be simply be defined as a tessellation whose cells are random closed sets:

Definition 1.3.3 (Random tessellations)**.** Let $S \subseteq \mathbb{R}$ be a bounded region of $\mathbb{R}^d$ or $\mathbb{R}^d$ itself ($d \geq 1$). Consider a collection $\Xi = (C_i)_{i\geq 1}$ of random closed sets. Assume that the $C_i's$ satisfy the conditions of the definition of a tessellation almost surely. Then the tessellation $\Xi = (C_i)_{i\geq 1}$ is said to be a *random* tessellation (with cells being the C_i's.)

Remark. The former definition is sufficient for most applications and, in particular, is sufficient for this thesis. More theoretical approaches on random tessellations are available in the literature, e.g. in [31, Chapter 9].

In the literature, we often find the following terminology:

- If $d = 2$, a tessellation of $\mathbb{R}^2$ is said to be *planar*.
- If $d = 3$, a tessellation of $\mathbb{R}^3$ is said to be *spatial*.
- Sometimes, the terminology *spatial tessellation* is encountered for referring to tessellations of $\mathbb{R}^d$ when $d > 3$.

Even though some general results are available for all kinds of tessellations, the vast majority of results in the literature concern *stationary* and *isotropic* tessellations, defined as follows:

Definition 1.3.4 (Stationary, Isotropy, Motion-invariance)**.** A random tessellation is said to be *stationary* (respectively *isotropic*) if its distribution is invariant by all translations (respectively rotations) of $\mathbb{R}^d$. A stationary and isotropic random tessellation is said to be *motion-invariant*.

1.3.3 Processes related to random tessellations

A random tessellation induces other random processes[1]:

- The point process of vertices of the cells
- The point process of edges' midpoints
- The point process of centroids of the cells
- The segment process[2] of edges of the cells

If the tessellation is stationary, one can show (e.g. see [31]) that all of the above processes are. Therefore, one may define their respective intensities: λ_0 (intensity of vertices), λ_1 (intensity of edges' midpoints), λ_2 (intensity of centroids of cells) and L_A (intensity of the line process of segment, also called *line density*, representing the average edge-length in an observation window of Lebesgue measure 1)[3].
Recalling that the intensity of a stationary point process is its mean number of points per unit volume, we can have the following interpretation:

- λ_0 is the mean number of vertices of the tessellation per unit volume.
- Since each edge of the tessellation almost surely has a unique midpoint, λ_1 is the mean number of edges of the tessellation per unit volume.
- Since each cell of the tessellation almost surely has a unique centroid, λ_2 is the mean number of cells of the tessellation per unit volume.

[1] Actually, one can rigorously define these point processes only if the cells of the tessellation are convex. Since we imposed the convexity assumption in the definition of a tessellation, we do not worry about this technicality anymore.

[2] Segment and line processes can actually be seen as point processes on abstract sets of lines. See [31, Chapter 8] for a rigorous approach.

[3] L_A is sometimes also denoted λ_3 in the literature, but the notation L_A refers to a more general context of the intensity of a line process. We choose to keep the latter notation.

Depending on the dimension and the type of tessellation, many formulae are available for the previous quantities in the literature (e.g. see [31, Chapter 9] or [132] for more advanced formulae). We shall only give some of those without proofs, in very particular cases of interest for this thesis. Before coming to that point, we introduce two of the most celebrated tessellations models: *Voronoi* and *Delaunay* tessellations.

1.3.4 Voronoi and Delaunay tessellations

Let $P = \{x_i : i \geq 1\}$ be a locally finite set of points in $\mathbb{R}^d$, i.e. for every bounded $B \subset \mathbb{R}^d$, the number of points of P located in B is finite. For each point $x_i \in P$ we define its *Voronoi cell* $\mathrm{Vor}(x_i \| P)$ as follows:

$$\mathrm{Vor}(x_i \| P) := \{x \in \mathbb{R}^d \, : \, \|x - x_i\|_2 \leq \inf_{x_j \in P, x_j \neq x_i} \|x - x_j\|_2\}, \tag{1.3.1}$$

where $\|\cdot\|_2$ denotes the Euclidean norm. In other words, the Voronoi cell of x_i is the set of points of $\mathbb{R}^d$ that are closer to x_i than to any other point of P.
Note that the local finiteness assumption on P ensures that the infimum in the definition is actually a minimum and is reached. Moreover, it is easy to show that the Voronoi cells are closed, compact and convex, hence the following definition:

Definition 1.3.5 (Voronoi tessellation)**.** The set of Voronoi cells $\{\mathrm{Vor}(x_i \| P) \, : \, i \geq 1\}$ is a tessellation of $\mathbb{R}^d$ called the *Voronoi tessellation* (generated by P).

Delaunay tessellations

Delaunay tessellations, also referred to as *Delaunay triangulations* can be defined in two different ways. The easiest, but more abstract, is the following:

Definition 1.3.6 (Delaunay tessellation)**.** Let $P = \{x_i : i \geq 1\}$ be a locally finite set of points in $\mathbb{R}^d$. The *Delaunay* tessellation or *Delaunay triangulation* generated by P is the dual tessellation (in the sense of Definition 1.3.2) of the Voronoi tessellation generated by P.

It can be shown that the cells of the Delaunay tessellations are triangles, hence the terminology "triangulation". The Delaunay tessellation can also be defined in a more concrete way by a direct construction:

Proposition 1.3.7. *Let $P = \{x_i : i \geq 1\}$ be a locally finite set of points in $\mathbb{R}^d$. One can recover the Delaunay tessellation generated by P in the following way: for all triples of points $(x_i, x_j, x_k) \in P^3$ with $x_i \neq x_j \neq x_k$, the triangle with vertices x_i, x_j and x_k is drawn if and only if the circumcircle of that triangle does not contain any other point of P than the three vertices of the triangle.*

1.3.5 Poissonian random tessellations

In stochastic geometry, Poissonian random tessellations, i.e. random tessellations generated by Poisson point processes, play a great role. Such tessellations have extensively been studied, often exhibit remarkable properties, and many formulae are available for

their characteristics.

We first begin with introducing three of the most studied Poissonian random tessellations. The first two examples can be introduced straightforwardly thanks to the previous definitions:

Definition 1.3.8 (Poisson-Voronoi and Poisson-Delaunay tessellations (PVT & PDT))**.** Let Φ be a simple Poisson point process on $\mathbb{R}^d$. The Voronoi (respectively Delaunay) tessellation generated by the atoms of Φ is a random tessellation called *Poisson-Voronoi* (respectively *Poisson-Delaunay*) tessellation.

Note that this definition is consistent with the general definition of Voronoi and Delaunay tessellations, as the set of atoms a Poisson point process is locally finite[4].

Our third example is more subtle:

Definition 1.3.9 (Poisson line tessellation (PLT))**.** Consider $X = \sum_k \delta_{X_k}$ a Poisson point process of finite positive intensity ρ on the real line $\mathbb{R}$. Let ϕ be a probability measure on $[0, 2\pi)$. Consider a sequence of i.i.d. random variables $(\theta_k) \overset{\text{i.i.d.}}{\sim} \phi$. Then, associate to each point X_k of Φ the line H_k tangent to the circle of centre the origin and radius $|X_k|$ at the point $\mathfrak{R}_{\theta_k}(X_k)$ (the image of X_k by the rotation of centre the origin and angle θ_k). The set of lines $(H_k)_k$ is called a *Poisson Line Process* and induces a random tessellation of the plane called a *Poisson Line Tessellation* (PLT).

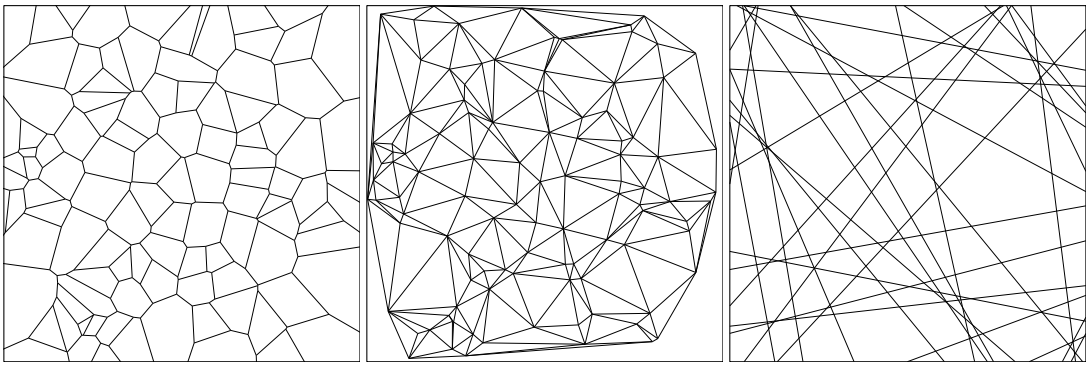
Remark. According to the form of ϕ, various tessellations can be obtained. For instance, if ϕ is the uniform distribution on $[0, 2\pi)$, the resulting PLT is isotropic. If ϕ consists in two atoms separated by $\frac{\pi}{2}$, the anisotropy is maximal. In general, the topology of the PLT is mainly constituted of degree 4-vertices. For more details on PLT and Poisson line processes, see [85, Section 7].

PVT and PDT are stationary and isotropic. A PLT is always stationary and isotropic when the angle distribution ϕ is uniform. For the sake of visualisation, Figure 1.1 shows particular realisations of a PVT, a PDT and a PLT in a 1-area window in $\mathbb{R}^2$.

As mentioned earlier, Poissonian tessellations have remarkable properties and their parameters $(\lambda_0, \lambda_1, \lambda_2, L_A)$ can often be expressed in terms of the intensity of the Poisson point process having generated the tessellation. For instance, consider the planar case, a homogeneous Poisson point process Φ of positive finite intensity $0 < \rho < \infty$ and $\phi \sim \mathcal{U}[0, 2\pi[$ uniformly distributed on $[0, 2\pi[$. Denote by PVT, PDT and PLT the respective Poisson-Voronoi, Poisson-Delaunay and Poisson line tessellations generated by Φ (respectively by Φ and a uniform angle distribution on $[0, 2\pi[$) for the PLT). Then the respective characteristics of these tessellations are given by Table 1.1.

Table 1.1 is just an example of mean formulae for Poissonian random tessellations. Many more results are available for different types of tessellations or in the spatial case (e.g. in [31, 108, 132]).

[4] This is even true for a more general point process Φ ; and due to the fact that Φ is a locally finite random measure, see Definition 1.1.5.



(a) PVT (b) PDT (c) PLT

Figure 1.1 – Examples of realisations of planar Poissonian random tessellations in a 1-area window. The PVT (Figure 1.1a) and the PDT (Figure 1.1b) have been generated by a homogeneous Poisson point process of intensity 100. The PLT (Figure 1.1c) has been generated with X of intensity 10 and ϕ the uniform measure on $[0, 2\pi[$.

Table 1.1 – Mean characteristics of planar Poissonian random tessellations as a function of the intensity ρ of the generating homogeneous Poisson point process.

	PVT	PDT	PLT
λ_0	2ρ	ρ	$\frac{\rho^2}{\pi}$
λ_1	3ρ	3ρ	$\frac{2\rho^2}{\pi}$
λ_2	ρ	2ρ	$\frac{\rho^2}{\pi}$
L_A	$2\sqrt{\rho}$	$\frac{32}{3\pi}\sqrt{\rho}$	ρ

More refined formulae are also of reach by considering the Palm probabilities of the point processes of vertices, edges midpoints and centroids of cells induced by the random tessellations. The respective typical points (i.e. the origin considered under Palm probability) of these point processes are called the *typical vertex*, *typical edge* and *typical cell* of the tessellation. The interested reader can for instance refer to [31, Sections 9.3 and 9.4] for more details.
Finally, the Voronoi tessellation generated by a general simple point process Φ plays a critical role in stationary Palm calculus. Indeed, the inversion formula of Palm calculus [12, Theorem 6.2.8] and the Neveu exchange formula [12, Theorem 6.3.7] are expressed in terms of Voronoi cells.

1.4 Discrete Percolation theory

1.4.1 A brief historical overview

Percolation theory originally arose from an applied problem introduced by Broadbent and Hammersley in [28, 68] and then developed by Frisch and Hammersley [49]. The problem was the following: assume that some liquid is poured over the top of a porous stone.

Will the liquid be able to make its way from pore to pore and reach the bottom of the stone? This physical question was mathematically modelled by Hammersley by using a three-dimensional network of, say, $n \times n \times n$ vertices called *sites* and in which the edges, called *bonds*, may be open, i.e. allowing the liquid through, with probability $p \in [0,1]$, independently from one another. The question therefore became the following one: for a given p, what is the probability that an *open path*, i.e. a path consisting of open bonds, connecting the top of the stone to its bottom exists? This problem is now called *Bernoulli bond percolation* has been extensively studied by physicists and mathematicians since its introduction. Standard references about percolation theory are [26, 62, 83, 95, 99, 121].

Over the years, percolation theory has gained much interest in the mathematical community. Though initial models for percolation problems only concerned discrete geometries, continuum models were also introduced[5] with the seminal work [55]. Nowadays, discrete and continuum percolation theory both remain very active fields of research and connections between the two are often made: indeed, continuum models are sometimes seen as the limit of discretised ones and many proofs of continuum models properties strongly rely on this limit argument.

In the next subsections, we shall review, in the case of the Euclidean lattice, the most basic model of (discrete) percolation known as *Bernoulli percolation* or *independent percolation*, where the state of bonds or sites are assumed to be independent. For the developments of this thesis, it will also be useful to put an emphasis on a more refined version of this model called *k-dependent percolation*, which is only briefly mentioned in some (not all!) percolation textbooks.

1.4.2 Bernoulli bond percolation

General framework

Consider an undirected graph $G = (V, E)$ with V the set of vertices (usually called *sites* in the context of percolation), assumed to be at most countable, and E the set of edges (called *bonds*). For $x, y \in V$, write $x \sim y$ if $\{x, y\} \in E$ and in that case, say that x and y are *connected by an edge*. A path from x to y is a sequence of sites $x = x_0 \sim x_1 \sim \cdots \sim x_k = y$ such that $\forall\, 0 \leq i \leq k-1$, $x_i \sim x_{i+1}$. Finally, we define the *degree* of a site $x \in V$ as $\deg x = \#\{y \in V, x \sim y\}$. We always assume the graph G to be *locally finite* i.e. $\forall\, x \in V, \deg x < \infty$.

In the simplest models of percolation having already been extensively studied, $G = \mathbb{L}^d$ is the *Euclidean lattice* for some $d \geq 1$, i.e. $V = \mathbb{Z}^d$ and

$$E = \left\{ \{x, y\}, x \in V, y \in V, \|x - y\|_1 := \sum_{i=1}^{d} |x_i - y_i| = 1 \right\}.$$

In other words, $x \sim y$ if and only if $\exists\, i \in \{1, \ldots, d\}$:

- $x_i = y_i \pm 1$
- $\forall j \neq i$, $x_j = y_j$

[5] We shall review them more precisely in the third chapter of this thesis. In this section, our main focus will be on discrete percolation.

Other lattices have been studied (in particular in dimension $d = 2$) [26].

In percolation, one randomises the states of bonds or sites. Bonds or sites can be *open* (i.e. allowing a flow[6] to get through) or *closed* (i.e. not allowing a flow to get through). If only the states of bonds are randomised, the model is called *bond percolation*. If we rather play with the states of sites, the model is called *site percolation*. Though mixed models with both the states of bonds and sites being randomised exist, we shall only be interested in site or bond percolation.

In both cases, the randomisation of the states of edges (respectively sites) is done by a parameter $p \in [0, 1]$, called *percolation parameter*. p will be the probability that a bond (respectively a site) is open in a bond (respectively site) percolation model.

Bernoulli bond percolation

For the sake of simplicity, we will assume from now onwards, unless stated otherwise, that $G = \mathbb{L}^d$ is the d-dimensional Euclidean lattice, for some $d \geq 1$.

In Bernoulli bond percolation model, we play with the states of the edges of G in the following way: for some percolation parameter $p \in [0, 1]$, we assign a state to each edge, independently from all other edges. An edge $e \in E$ may be *open* (probability p) or *closed* (probability $1 - p$). A convenient way to describe the states of all edges and hence of the system is the following one: consider some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a random field $(U_e)_{e \in E}$ of i.i.d. random variables uniformly distributed on $[0, 1]$. Declare the edge $e \in E$ to be *open* if $U_e \leq p$ and closed otherwise. The states of the edges are given by the variables U_e of the random field, and it is clear from this construction that each edge is open with probability p independently from all other edges[7].
The previously defined model is called *Bernoulli (or independent) bond percolation model* with parameter p. To insist on the dependence on the percolation parameter p, the probability measure $\mathbb{P}$ is often denoted by $\mathbb{P}_p$ and the related expectation by $\mathbb{E}_p$.

We now give some terminology which will be needed throughout the rest of this section:

Definition 1.4.1 (Open path)**.** A path $x_0 \sim \cdots \sim x_k$ is said to be an *open path* if all of its edges are open.

The state of an edge (open or closed) can be thought of as its presence or its absence, which, by abuse of terminology, asserts the following definition:

Definition 1.4.2 (Connected vertices)**.** Two vertices x and y are said to be *connected* (in the configuration ω) if there exists an open path from x to y. In that case, we denote $x \rightsquigarrow y$.

[6] Depending on the context and on the applications, a flow may be a flow of water, electrical current, signal, information... Nevertheless, the interpretation remains the same.

[7] A more rigorous construction often encountered in the literature is done by considering the sample space sample space of *configurations* $\Omega = \{0, 1\}^E$ and defining a proper probability measure $\mathbb{P}_p$ on it. We won't need to resort to this technicality in this thesis, but the interested reader can find details about this construction in [62].

Definition 1.4.3 (Connected sets)**.** Let $A, B \subset V$. Then A and B are said to be *connected* if $\exists\, x \in A, y \in B, x \rightsquigarrow y$. Similarly, this is denoted by $A \rightsquigarrow B$.

Percolation theory focuses on the study of existence of infinite connected components, which induces the following definition:

Definition 1.4.4 (Cluster)**.** Let $x \in V$. The cluster of x, denoted $\mathcal{C}(x)$ is the open connected component of x, i.e. the set of sites being connected (in the sense of Definition 1.4.2) to x:

$$\mathcal{C}(x) = \{y \in V, x \rightsquigarrow y\}$$

Finally, we denote by $|\mathcal{C}(x)|$ the cardinal of the cluster of x and write $x \rightsquigarrow \infty$ if $|\mathcal{C}(x)| = \infty$. The latter is equivalent to the existence of an infinite *self-avoiding* open path (meaning that no site is visited twice) starting from x.

Remark. Since G is the Euclidean lattice, by translation invariance of the Bernoulli bond percolation model, we have $\forall x \in V, |\mathcal{C}(x)| \overset{d}{=} |\mathcal{C}(0)|$. Therefore, $\mathcal{C}(0)$ is usually denoted $\mathcal{C}$.

Percolation or phase transition

Several quantities are of interest in bond percolation. The most important one is the *percolation function*:

Definition 1.4.5 (Percolation function)**.** The percolation function θ is defined by $\theta(p) = \mathbb{P}_p(|\mathcal{C}| = \infty) = \mathbb{P}_p(0 \rightsquigarrow \infty)$

By translation invariance of the model, we also have $\forall x \in V, \theta(p) = \mathbb{P}_p(|\mathcal{C}(x)| = \infty)$. *Percolation* of the model is defined in the following way:

Definition 1.4.6 (Percolation)**.** We say the model *percolates* if $\theta(p) > 0$.

It is easy to prove that θ is increasing as a function of p (indeed, the greater p is, the more connections are allowed in G and so the easier it is to percolate). Since $\theta(0) = 0$ obviously, the primary question that intially concerned Hammersley [68] is the existence and, if possible, value of a critical value for which θ is positive. This led to the following:

Definition 1.4.7 (Percolation threshold)**.** We define the *percolation threshold* $p_c = p_c(d)$ as: $p_c := \sup\{p \in [0,1] : \theta(p) = 0\}$

A consequence of the definition is that $\theta(p) = 0$ and $|\mathcal{C}| < \infty$ $\mathbb{P}_p$-a.s. for $p < p_c$, while $\theta(p) > 0$ and $|\mathcal{C}| = \infty$ with positive $\mathbb{P}_p$-probability for $p > p_c$.

The existence of a critical threshold makes Bernoulli bond percolation model one featuring a *phase transition*, i.e. the system undergoes a radical change when some parameter (here p) passes a certain critical threshold. This induces the following definition:

Definition 1.4.8. If $p < p_c$, we say that the the system is in the *subcritical phase*. If $p > p_c$, we say the system is in the *supercritical phase*.

The case of dimension 1 is quite trivial: indeed, if $p < 1$, there will be infintely many closed bonds to the right and to the left of the origin $\mathbb{P}_p$-a.s. Therefore, for $d = 1$, $p_c = 1$. The case of higher dimensions and of other lattices often remains an open question. Even

the simplest case of the square lattice (i.e. $d = 2$) remained an open question for years! For symmetry reasons, one might think that $p_c(2) = \frac{1}{2}$. This is indeed the case, but the proof is actually much more complicated. The result is due[8] to Kesten [84]:

Theorem 1.4.9. *If $d = 2$, i.e. the underlying graph of the model is the square lattice, then $p_c = \frac{1}{2}$.*

Since the original paper of Kesten [84], the proof has been refined and considerably shortened, but it still remains highly non-trivial.

Surprisingly though, it is relatively easy, by the means of duality arguments and the celebrated *Peierls argument* (used by Peierls to prove phase transition in the two-dimensional Ising model [110]) to prove that the phase transition is not trivial in dimension greater than or equal to 2. As the proof is very instructing, we shall state the result with its proof:

Theorem 1.4.10. *For all $d \geq 2$, we have $0 < p_c(d) < 1$.*

Proof. As is usually done in percolation theory, the proof is done in two parts. First we prove that $p_c > 0$ (existence of the subcritical phase) and then we prove that $p_c < 1$ (existence of the supercritical phase).

$\boxed{p_c > 0}$ We will use a so-called *path-count argument*. Fix some $n \geq 1$. A *self-avoiding path* γ of length n is a sequence of edges $e_1, \ldots, e_n$ with $e_i \neq e_j$ for $i \neq j$ and such that e_i and e_{i+1} share a common endpoint for every $1 \leq i \leq n-1$. Obviously, if 0 is connected to infinity, for all $n \geq 1$, there exists a self-avoiding path of length n, starting from 0 (i.e. $0 \in e_1$, which means that 0 is an endpoint of e_1) and made of open edges. Let A_n denote the former event. Then we have:

$$\theta(p) =: \mathbb{P}_p(0 \rightsquigarrow \infty) \leq \mathbb{P}_p(A_n)$$

Let SAP_n denote the set of self-avoiding paths of length n starting from the origin. We have:

$$\mathbb{P}_p(A_n) \leq \sum_{\gamma \in SAP_n} \mathbb{P}_p(\text{all edges of } \gamma \text{ are open}) = \#(SAP_n)p^n,$$

where we have used the union bound in the first inequality and the independence of edges in the equality, and where $\#(SAP_n)$ denotes the number of self-avoiding paths of length n starting from 0. Obviously, $\#(SAP_n) \leq (2d)^n$, so that, in all, we have:

$$\theta(p) \leq (2dp)^n, \quad \text{for all } n \geq 1$$

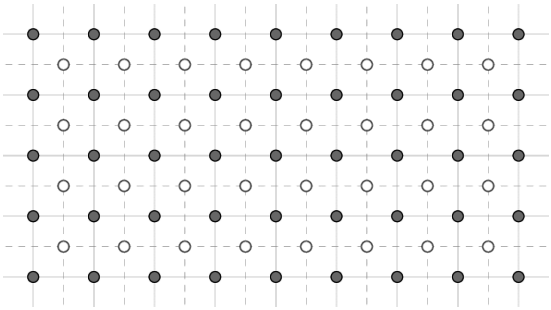
When $p < \frac{1}{2d}$, the quantity in the right-hand side converges to 0, and so $\theta(p) = 0$. Hence $p_c(d) \geq \frac{1}{2d} > 0$.

$\boxed{p_c < 1}$ The argument is a bit more subtle in the supercritical phase. First, note that percolation on $\mathbb{L}^d$ implies percolation on $\mathbb{L}^{d+1}$ for all $d \geq 2$, by considering $\mathbb{L}^d$ as a

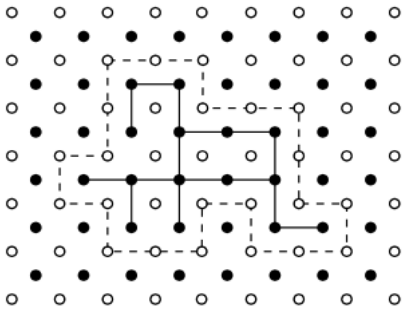
[8] Kesten actually used an earlier result from Harris [69] (who proved that $\theta(\frac{1}{2}) = 0$ when $d = 2$) and independently found results of Russo [115], Seymour and Welsh [119] about the role of the mean cluster size. Assembling all these arguments, he established the proof of Theorem 1.4.9.

subspace $\mathbb{L}^d \times \{0\}$ of $\mathbb{L}^{d+1}$. Therefore, we have $p_c(d+1) \leq p_c(d)$ whenever $d \geq 2$ and so it suffices to prove that $p_c(2) < 1$.

Call the square lattice $\mathbb{L}^2$ the *primal lattice* and consider the *dual lattice* $(\mathbb{L}^2)^*$ of $\mathbb{L}^2$ defined by the translation of the square lattice $\mathbb{L}^2$ by the vector $(1/2, 1/2)$. In other words, we put a dual vertex at the centre of every face of the primal lattice and draw a dual edge between two dual vertices if and only if the faces they belong to in the primal lattice share a common (primal) edge[9], see Figure 1.2a.



(a) The square lattice $\mathbb{L}^2$ and its dual lattice $(\mathbb{L}^2)^*$. The primary vertices are the black points, the dual vertices are the white ones. The primary edges are the plain segments, the dual edges are the dashed ones.



(b) Illustration of Peierls argument. Black points are the primal vertices, white points are the dual ones. A segment between two primal vertices (black points) represents an open edge in the primal lattice. A dashed segment between two dual vertices (white points) represents a closed edge in the dual lattice.

Figure 1.2 – Illustration of the dual lattice $(\mathbb{L}^2)^*$ and of the Peierls argument.

It is obvious that the square lattice and its dual are isomorphic. A lattice satisfying this property is called *self-dual*. Moreover, each edge of the dual lattice crosses a unique edge of the primal lattice. Hence, we can define a bond percolation model on the dual lattice $(\mathbb{L}^2)^*$ in the following way: declare a dual edge open (respectively closed) if and only if the unique primal edge it crosses is open (respectively closed). In doing so, it is obvious we have defined a percolation process on $(\mathbb{L}^2)^*$ that has the law $\mathbb{P}_p$ of Bernoulli bond percolation with parameter p. Now we use the celebrated *Peierls argument*: from a picture[10], it is clear that if the cluster of the origin is finite in the primal lattice, then it has to be surrounded by a closed dual circuit[11] in the dual lattice, see Figure 1.2b.

Therefore, we have:

[9] The terminology *dual* lattice comes from the fact that if we consider square lattice $\mathbb{L}^2$ as a tessellation of $\mathbb{R}^2$, the dual lattice $(\mathbb{L}^2)^*$ is just the dual tessellation of $\mathbb{L}^2$, in the sense of Definition 1.3.2.

[10] A full proof with complete rigour would be tedious to write but can nevertheless be found in [83].

[11] By *circuit* surrounding the origin we mean a self-avoiding loop, i.e. the endpoint is also the starting point.

$$\begin{aligned} 1-\theta(p) &=: \mathbb{P}_p(|\mathcal{C}| < \infty) \\ &\leq \mathbb{P}_p(\text{there exists a closed circuit surrounding the origin in the dual lattice}) \\ &\leq \mathbb{E}_p(\text{number of closed circuits surrounding the origin in the dual lattice}) \\ &= \sum_{n=4}^{\infty} \rho(n)(1-p)^n, \end{aligned}$$

where we have used Markov's inequality in the second inequality and the independence of edges in the dual percolation model in the last equality, and where $\rho(n)$ denotes the number of circuits surrounding the origin in the dual lattice. Obviously, the previous sum starts at $n = 4$ because a circuit surrounding the origin is at least composed of 4 edges.

A rather crude bound for $\rho(n)$ is $\rho(n) \leq n3^{n-1}$. Indeed, the circuit around the origin has to cross the half-line $\mathbb{R}_+$ at least once and must therefore intersect the set $\{(k+1/2, 0) : 0 \leq k \leq n/2\}$ (the bounds on k are due to the fact that the circuit is of length n). There are at most $1 + n/2 \leq n$ such possible crossing points. We can thus build any circuit surrounding the origin in the following way: once the (closest to the origin) crossing of the circuit with the half-line $\mathbb{R}_+$ has been placed, the rest of the circuit is built by adding edges one by one, each time making three possible choices: go ahead, turn left or turn right (going back is forbidden has the circuit has to be self-avoiding). There are thus 3^{n-1} possible choices. Hence $\rho(n) \leq n3^{n-1}$

In all, we get:

$$\mathbb{P}_p(|\mathcal{C}| < \infty) \leq \sum_{n=4}^{\infty} n3^{n-1}(1-p)^n$$

The series in the right-hand side converges to 0 when $p > 2/3$. Therefore, $\theta(p) > 0$ whenever $p > 2/3$ and so $p_c(2) \leq 2/3 < 1$. □

1.4.3 Bernoulli site percolation

In Bernoulli site percolation model, we play with the sites. In that case, we consider a random field $(U_v)_{v \in V}$ of i.i.d. random variables uniformly distributed on $[0, 1]$, now indexed by the set of sites of G.[12]. As before, a site $v \in V$ is said to be open if $U_v \leq p$ and closed otherwise. It is again clear from the construction that each site is open with probability p independently from all other sites. This model is called *Bernoulli (or independent) site percolation model* with parameter p.

All the definitions and theorems related to Bernoulli bond percolation can be adapted to Bernoulli site percolation. A more detailed account on site percolation can be found in [62].

Site percolation is somehow more general than bond percolation. This is due to the fact that a bond percolation model on any graph G can be seen as a site percolation model on a modified graph called the *line graph* of G, usually denoted $L(G)$. Indeed, consider

[12] Once again, a more rigourous construction is done by considering the sample space of *configurations* $\Omega = \{0, 1\}^V$ and defining a probability measure $\mathbb{P}_p$ on Ω. See [62] for more details.

a bond percolation model with parameter p on a graph G. $L(G)$ is obtained by placing a vertex v_e on each edge e of G and connecting two vertices v_e and $v_{e'}$ by an edge if and only if the corresponding edges e and e' share a common vertex in G. Then, declare a site v_e of $L(G)$ to be open (respectively closed) if and only if the corresponding edge e of G is open (respectively closed). Then the percolation process induced on $L(G)$ has the law of Bernoulli site percolation. Hence, we can view bond percolation as a particular case of site percolation.

In many percolation textbooks, the main focus is however rather on bond percolation models. This is mostly for historical reasons: bond percolation was the first model to be introduced and studied on the square lattice. Moreover, there may be a practical reason to that. In the proof of Theorem 1.4.10, we resorted to duality arguments to prove the non-triviality of the phase transition for bond percolation on Euclidean lattices ($p_c(d) < 1$). This duality argument strongly relies on the fact that the square lattice $\mathbb{L}^2$ is self-dual. For site percolation arguments, such duality arguments exist but are more subtle. Indeed, in site percolation models, we define the dual lattice as follows: for each face of the primal lattice, we keep the existing edges and add an edge between any two vertices that are not yet connected. In other words, the dual lattice is just the primal lattice with added *diagonals* in each face.

For such a type of duality, the square lattice is no longer self-dual but the *triangular lattice* is. Seeing $\mathbb{C}$ as being isomorphic to $\mathbb{R}^2$, the triangular lattice is easily defined as being a graph vith vertex set $V = \{x + ye^{i\pi/3}, (x, y) \in \mathbb{Z}^2\}$ and where an edge $z \sim z'$ is drawn if $|z - z'| = 1$. An example of realisation of a site percolation model on the triangular lattice is illustrated by Figure 1.3. We shall not go into further details on such matters, as it is not relevant for this thesis.

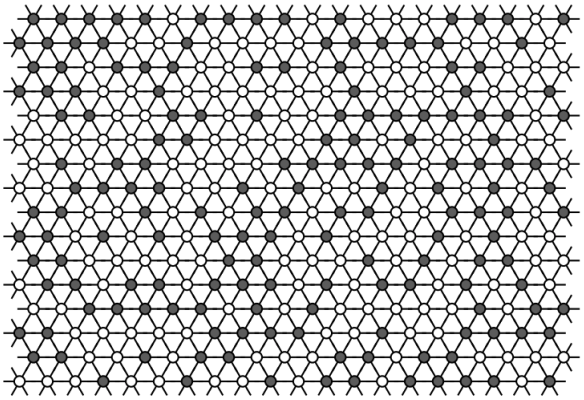


Figure 1.3 – An example of realisation of a site percolation model on the triangular lattice. Black points represent the open sites and white points represent the closed sites.

1.4.4 k-dependent site percolation

Consider Bernoulli site percolation on the Euclidean lattice $\mathbb{L}^d$ with parameter p. Each site $v \in \mathbb{Z}^d$ is open with probability p, independently from every other site. As before, wee see the distribution of the percolation process as the joint distribution of i.i.d. random variables $(U_v)_{v\in\mathbb{Z}^d}$ uniformly distributed on $[0, 1]$.

A natural extension to this model is done by loosening the independence assumption between the states of any two different sites. A particular case, which will be of interest in this thesis, is the case of k-dependent percolation, where the associated i.i.d. uniform random variables are no longer independent but form a *k-dependent random field*. The definition is as follows:

Definition 1.4.11 (k-dependent random field)**.** Let $k \geq 1$ and let $X = (X_z)_{z\in\mathbb{Z}^d}$ be a discrete random field. Denote by d a metric on $\mathbb{Z}^d$. The random field X is said to be k-dependent for the metric d if for all $q \geq 1$ and all $\{s_1, \dots s_q\} \subset \mathbb{Z}^2$ finite with the property that $\forall i \neq j, d(s_i - s_j) > k$, the random variables $(X_{s_i})_{1\leq i\leq q}$ are independent.

Informally, k-dependence of the field means that sufficiently distant random variables of the field will be independent. Note that the above definition obviously depends on the chosen metric d on $\mathbb{Z}^d$. In the literature, the definition is often given for the taxicab metric $\|\cdot\|_1$. Other choices of metric are possible, such as the classical Euclidean metric $\|\cdot\|_2$ or the supremum metric $\|\cdot\|_\infty$.

The model

Using the aforementioned analogy, it is easy to define the k-dependent site percolation model:

Definition 1.4.12 (k-dependent site percolation on the Euclidean lattice)**.** Consider the d-dimensional Euclidean lattice $G = \mathbb{L}^d$ and a k-dependent random field for the taxicab metric $d = \|\cdot\|_1$ $U = (U_z)_{z\in\mathbb{Z}^d}$, where all the underlying random variables are uniformly distributed on $[0, 1]$. Let $p \in [0, 1]$. Declare a site $z \in \mathbb{Z}^d$ to be *open* if $U_z \leq p$, and closed otherwise. This defines a site percolation process on $\mathbb{L}^d$ called *k-dependent site percolation*.

Remark. In the same way, one may define k-dependent bond percolation on the Euclidean lattice. k-dependent percolation models can also be defined for more general graphs, as long as the k-dependence property of the random field is ensured.

Phase transition

k-dependent site percolation, very much as Bernoulli site percolation, exhibits a non-trivial phase transition on Euclidean lattices:

Theorem 1.4.13. *Let $G = \mathbb{L}^d$ is the d-dimensional Euclidean lattice and let $p_{c,site}(k, d)$ denote the critical parameter for k-dependent site percolation on G. Then, for all $d \geq 2$ and $k \geq 1$, $p_{c,site}(k, d) \in (0, 1)$*

In the literature, the proof of Theorem 1.4.13 is most of the time done by resorting to a result of Liggett, Schonmann and Stacey [91, Theorem 0.0], stating that any $\{0, 1\}$-valued k-dependent random field can, under certain conditions, be stochastically dominated by a product measure. In terms of percolation models, this is equivalent to stochastically dominating the k-dependent model by an independent model, which is much easier to be work with.

In the next chapters of this thesis, we will again encounter k-dependent site percolation

models. An adapted *path-count argument*, very much as in the proof of Theorem 1.4.10, considerably simplifies the proof that $p_{c,site}(d,k) > 0$. We shall write it, for the sake of completeness.

Proof that $p_{c,site}(d,k) > 0$. We begin by proceeding exactly as in the proof of Theorem 1.4.10. Keeping the same notations (and replacing bonds by sites) as in the aforementioned proof, we still have:

$$\theta(p) \leq \mathbb{P}(A_n) \leq \sum_{\gamma \in SAP_n} \mathbb{P}(\text{all sites of } \gamma \text{ are open})$$

The subtlety in the k-dependent case is that the probability in the right-hand side is not equal to p^n anymore, as the states of the sites are no longer independent but k-dependent. To deal with this, write any self-avoiding path γ of length n as $\gamma = \{z_1, \dots, z_n\}$.

Since we will look at limits when $n \to \infty$, we can assume without loss of generality that n is large enough for the needs of our developments, say $n > (2k)^d$. We extract a subset γ' of sites of γ iteratively as follows. First, retain $z'_1 := z_1 = 0$ and remove from γ all those sites $z \in \gamma$ such that $\|z - z_1\|_1 \leq k$. Once we have removed all these sites from γ, retain the next available site of γ, say z'_2 and, again, remove all remaining sites z of γ such that $\|z - z'_2\|_1 \leq k$, and so on.
At each iteration j, the number of sites removed from γ is upper-bounded by the (discrete) volume of the (discrete) closed l^1-ball with centre z'_j and radius k. This volume is upper-bounded by $(2k)^d$: indeed, this l^1-ball is a subset of the l^∞-ball[13], whose volume is $(2k)^d$. Since the length of γ is n, the subset $\gamma' := \{z'_1, \dots, z'_m\}$ of sites of γ extracted by the previous construction contains at least $m \geq \lfloor \frac{n}{(2k)^d} \rfloor$ sites of γ.
Moreover, by construction, we have that $\forall x, y \in \gamma', \|x-y\|_1 > k$. Hence, by k-dependence of the random field $U = (U_z)_{z \in \mathbb{Z}^d}$, the states of the sites in γ' are independent. Thus, we get:

$$\begin{aligned}
\theta(p) &\leq \sum_{\gamma \in SAP_n} \mathbb{P}(\text{all sites of } \gamma \text{ are open}) \\
&\leq \sum_{\gamma \in SAP_n} \mathbb{P}(\text{all sites of } \gamma' \text{ are open}) \\
&= \sum_{\gamma \in SAP_n} \prod_{z' \in \gamma'} \mathbb{P}(z' \text{ is open}) \\
&\leq \sum_{\gamma \in SAP_n} p^{\lfloor \frac{n}{(2k)^d} \rfloor} \\
&= \#(SAP_n) p^{\lfloor \frac{n}{(2k)^d} \rfloor} \\
&\leq (2d)^n p^{\frac{n}{(2k)^d} - 1},
\end{aligned}$$

where we used the same bound as before $\#(SAP_n) \leq (2d)^n$ and the fact that $\lfloor \frac{n}{(2k)^d} \rfloor \geq$

[13] Recall that the l^1-distance and l^∞-distance between any two points $x = (x_1, \dots, x_d)$ and $y = (y_1, \dots, y_d)$ of $\mathbb{Z}^d$ are respectively defined by $\|x-y\|_1 := \sum_{i=1}^d |x_i - y_i|$ and $\|x-y\|_\infty := \max_{1 \leq i \leq d} |x_i - y_i|$. The closed l^1-ball with centre x and radius r is the set $\{y \in \mathbb{Z}^d : \|x-y\|_1 \leq r\}$ and the corresponding closed l^∞-ball is $\{y \in \mathbb{Z}^d : \|x-y\|_\infty \leq r\}$.

$\frac{n}{(2k)^d} - 1$ in the last line. As $n \uparrow \infty$, the last quantity converges to 0 when $p < \left(\frac{1}{2d}\right)^{(2k)^d}$, and so $p_{c,site}(d,k) \geq \left(\frac{1}{2d}\right)^{(2k)^d} > 0$.

□

The former technique can also be used to get back results that are known in the context of Bernoulli percolation (e.g. exponential bounds for the size of the cluster of the origin, see [26]) for k-dependent percolation models.

Chapter 2

Background on wireless communications

It is change, continuing change, inevitable change, that is the dominant factor in society today. No sensible decision can be made any longer without taking into account not only the world as it is, but the world as it will be...
This, in turn, means that our statesmen, our businessmen, our everyman must take on a science fictional way of thinking.

Isaac Asimov

In this chapter, we provide some background on wireless communications which will be needed to understand the networks models studied in the rest of this thesis. Rather than giving a complete introduction to the physical concepts at stake in wireless communications from a physicist's or an engineer's point of view, the goal of this chapter is to introduce basic concepts of information theory and digital communications that have their own place in all the mathematical models of wireless networks.
First, we present these concepts and provide further references for the interested reader. We then go on with introducing the concept of *Device-to-Device* (D2D) communications, thus giving rise to the so-called *Device-to-Device* networks, in which we will be interested throughout the rest of this thesis. We finally focus on a possible application of D2D in future networking scenarios consisting in taking advantage of the explosive growth of the number of connected devices. Such a way of thinking future cellular networks, resumed under the terminology *crowd-networking* in the literature, gives motivation for the work done in the rest of this thesis.

2.1 Wireless cellular networks

2.1.1 General terminology

In a first approach, a wireless network can be viewed as a collection of interconnected nodes located in some physical domain. By *wireless* networks, we mean that the connections between the nodes do not require the intervention of a physical medium to carry the signal such a copper wire or an optical fiber. A wireless link connecting two nodes of a wireless network is called a *wireless channel*. Depending on the type of network considered, the nodes may be of different kinds. For instance, a node may be a user equipped with a phone and trying to access the web, a fixed antenna or an infrastructure of the network, allowing the previous user to access the network and the internet. Each node of the network can be a transmitter or a receiver, and, depending on the type of network and technology used, some nodes can simultaneously transmit and receive a signal.

In wireless networks, the communications are based on *electromagnetic waves*, which are generated by the synchronised oscillations of an electric and a magnetic field. Two of the main characteristics of such waves are the *frequency* f, measured in Hertz (Hz) and its derived units; and the *wavelength* λ, measured in meters (m) and its derived units. These two quantities linked by the following relation:

$$f = \frac{c}{\lambda},$$

where $c = 3.00 \times 10^8$ m.s^{-1} is the speed of light in the vacuum. The frequency f can be interpreted as the number of oscillations of the wave per unit time, while the wavelength λ corresponds to the spatial period of the wave, i.e. the distance over which the shape of the wave repeats. The range of all possible electromagnetic waves that exist is called the *electromagnetic spectrum*. Not all frequencies of the electromagnetic spectrum are equivalent, and, according to where a frequency is located in the electromagnetic spectrum, the properties of a wave oscillating at that given frequency may hugely vary. In telecommunications networks, the frequencies of the waves propagating the signal belong to a part of the electromagnetic spectrum called the *radio spectrum*. The limits of the radio spectrum are subject to debate in the literature, but the most common point of view (e.g. developed in [126]) is that the radio spectrum roughly corresponds to frequencies in the range $[3\,\text{kHz}, 300\,\text{GHz}]$.

The frequencies belonging to the radio spectrum itself have been classified into different contiguous sections, also called *radio bands*, by the International Telecommunications Union (ITU) [1], a specialized agency of the United Nations charged with regulating and coordinating at an international scale the use of the different frequency ranges available. Different classifications done by other regulating bodies (e.g. the American Institute of Electrical and Electronics Engineers, IEEE) also exist, but the ITU classification is the most used internationally. Inside each radio band, more ITU's Radio Regulations (RR) define the applications and conditions under which a part of the band should be used.

All of the previous machinery is set up to ensure the best possible signal quality between transmitters and receivers, as well as to avoid signal degradation. At a national

scale, government agencies allocate the frequency ranges and define how they should be used or shared, while ensuring that international regulations are respected. However, the multiplication of users on a given frequency range as well as the variety of applications may result in a disruption of the quality of some signals[1]. As a matter of fact, regulating agencies sometimes authorise the use of a spectrum of frequencies only to a given actor and for a given use. This is called *licensing*. Radio bands thus fall into two main categories:

1. *Licensed bands*, which can be used after the acquisition and payment of a license to the governing body. The owner of the license is granted exclusive use of the band he paid for, with the guarantee that no parasite signals coming from another actor will be emitted in that band.

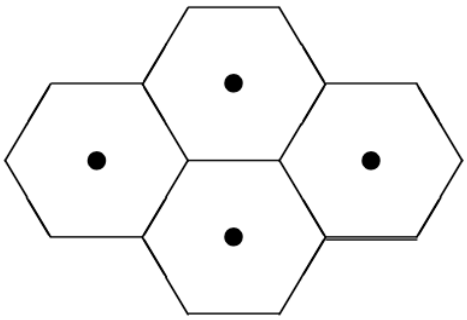
2. *Unlicensed bands*, whose use is not regulated by the acquisition of a license. This means that unlicensed bands can be used by anyone who wants to use them. However, the term *unlicensed* does not mean that anyone can do whatever he wants, as some general regulations must still be followed.

The acquisition of a license representing a considerable amount of money, it is no wonder that, most of the time, only telecommunications operators and companies have the financial ability to acquire licenses. This in turn ensures them that they are the only ones able to use a frequency band after having paid for it. By contrast, anyone may use unlicensed bands without paying for the right to do so, but has no guarantee he will be the only one to emit at a given frequency. Some parasite signals at the same frequency may also be emitted by other users of those unlicensed bands. Therefore, one of the main cons of using unlicensed spectrum is the difficulty of ensuring that the emitted signal will be received with a satisfying quality.

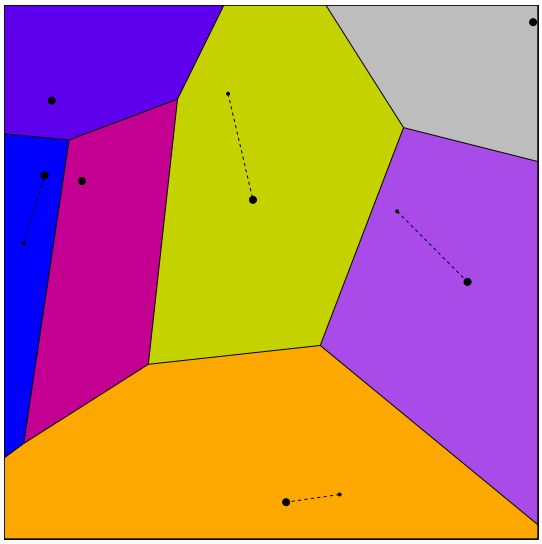
2.1.2 Cellular networks

A particular category of telecommunications networks are the well-known *cellular networks* or *mobile networks*. In such a network, there are many users, equipped with *cellphones* or *mobiles*, who communicate within a common network infrastructure. This common network infrastructure is organised and made of various equipment who ensure global connectivity of the network. Mobile networks were designed with the following thought: the user may not stay in a fixed location and move, thus asserting the terminology *mobile* networks. The original question in the design of such networks was to allow the conversation between two users of the network to be pursued even if the users are moving. Over the years and the advances in radio technology, mobile networks also brought new possibilities to their users, such as text messaging, video messaging and calling or internet browsing. The development of mobile networks is classified into several *mobile networks generations*, each representing a milestone compared to the precedent one [61, Section 1.4]. The most recent generation of mobile networks is the fourth one (4G) and the fifth generation of mobile networks (5G) is currently under development, even though its early deployment has already started in some parts of the world.

[1] This is called *interference*. We shall come back to this concept in Subsection 2.1.4.



(a) Hexagonal cells in a mobile network. Black points represents BSs. Figure taken from [126].



(b) A more realistic approach where BSs are not regularly placed in the network and where each cellphone connects to the closest BS available. The larger points represent BSs and the smaller ones represent users' cellphones. In this case, the coverage zone of a cell is given by its Voronoi cell, as defined in Definition 1.3.5.

Figure 2.1 – Schematic representation of coverage of an area by base stations and cells in a cellular network.

Each generation brought its own standards, regulations and network equipment. However, the same concept remains used in all the generations of cellular networks: the geographical area to be covered is divided into smaller land areas called *cells*, each being served by at least one (most of the time several) fixed-location equipment called *base station* (BS). Roughly speaking, a BS is what we would call "antenna" in the everyday language, though the actual BSs have a more complicated architecture including antennas but also several other communication equipment. A cell being the coverage zone of a BS, this explains the terminology of *cellular networks*. In the literature, the first proposed models for cells date back to 1947 [114], much before the early days of cellular networks, and made the case for regularly spread base stations with hexagonal cells, see Figure 2.1a. Though questionable, this initial model was refined over the years, with the hexagonal concept still in the line of thought as early as from 1979 [97], when standards from the first generation of mobile networks (1G) were investigated. Even today, models based on hexagonal cells still draw attention. Indeed, as underlined in [21, Section 2.1], one of the main advantages of a hexagon-based model is the following: regular hexagons can tile the plane while not overlapping one another. Equilateral triangles or squares would also verify the same property. However, for a fixed distance between the centre of a regular cell and its vertices, an hexagon has a much larger area than a square or a triangle. In reality however, for various reasons (spectrum management, interference, costs ...), BSs are not regularly placed in the plane and, when a user connects to the network, its cell-

phone usually connects to the closest BS available so as to ensure the best possible signal quality, see Figure 2.1b.

Regarding the architecture of cellular networks per se, BSs are not the only equipment intervening. An oversimplified view, as depicted in Figure 2.2, is the following:

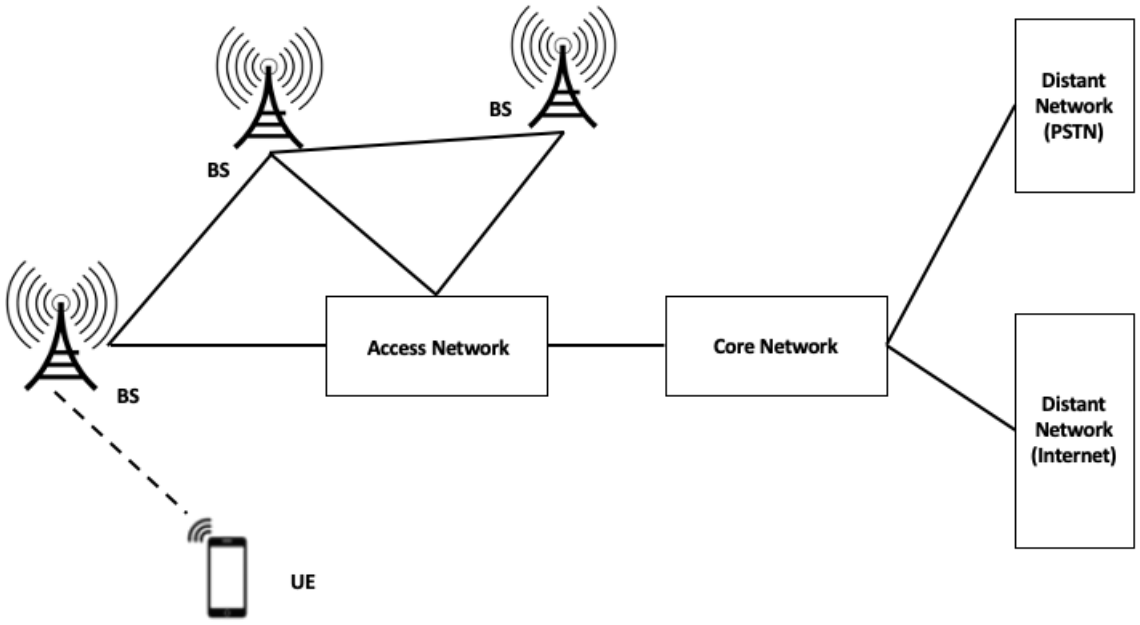


Figure 2.2 – An oversimplified view of a cellular network architecture. The plain lines correspond to wired connections, the dashed lines correspond to wireless connections.

- A user is equipped with a user equipment (UE), which is most of the time a cellphone. With the emergence of new applications, such as internet browsing or cloud-gaming, the UE may be a tablet, a computer... In any case, it designates the piece of equipment with which the user wants to connect to the network.

- A UE is connected by a wireless channel to a BS. In newer generations, the architecture of BSs has considerably changed and induced new terminology: in the third generation (3G) networks, BSs are designated by the term *Node B* (NB); in 4G networks, BSs are called *evolved Node B* or simply *eNode B* (eNB).

- BSs are interconnected by wired links to additional equipment constituting the *access network*. The main role of this access network is to identify a UE and give it access to the telecommunication service of the provider. The access network is itself made of differents parts.

- The different parts of the access network are interconnected by another network of wired links called the *core network*. Another function of this core network is to provide a gateway to distant networks, for instance the public switched telephone network (PSTN) for voice calls, or the internet network for data exchange.

2.1.3 Radio propagation effects

Consider a transmitter (say, a BS) transmitting a certain radio signal to a receiver (say, a UE) located at a certain distance from the transmitter. In actual applications, the path run by the radio wave to reach the receiver may include physical obstacles, such as hills, buildings, walls ... These obstacles will have significant effects on the signal transmitted by the radio wave. Moreover, the signal may also run a long distance before reaching the receiver: this also has some consequences on the signal received at destination. Such modifications between the signal emitted and the one received are due to the so-called *propagation effects*. Most of the time, the *power* emitted by the transmitter (representing the amount of energy necessary to push the signal over to the receiver) will be significantly different from the actual power received by the receiver, due to the possibly large distance run by the signal and the possible physical obstacles encountered in way. A low power at reception may result in an unstable connection (for instance a phone call with low voice quality, making it harder to understand the person on the phone). We now evoke some of these propagation effects.

Path loss

Path loss, or *path attenuation*, is the variation of the power of the radio wave due to distance between the transmitter and the receiver. This attenuation may be due to different factors such as diffraction or reflection on physical obstacles, as well as the nature of the medium in which the radio signal propagates. In the literature, a common modelling is as follows: assume that a transmitter emits a radio wave with power P_{tx} expressed in Watts (W) towards a receiver at distance d. Then, the power P_{rx} (also expressed in W) received by the receiver is proportional to $P_{tx}\ell(d)$ where the function $\ell : \mathbb{R}_+ \to [0, 1]$ is decreasing in d and called *path loss function*. The fact that ℓ is upper-bounded by 1 is explained by the fact that it is an attenuation factor of the power of the emitted signal. Common models used in the literature, as suggested in [11, Example 23.1.3], include the following:

$$\ell(d) = (\kappa \max(d_0, d))^{-\beta},$$
$$\ell(d) = (1 + \kappa d)^{-\beta},$$
$$\ell(d) = (\kappa d)^{-\beta},$$

for some constants $\kappa > 0$, $d_0 > 0$ and where $\beta \geq 2$ is called the *path loss exponent*. The constant κ depends on the geometry of the wireless link used for the transmission (mainly the heights of the antennas used for emission and reception) and on the frequency of the radio signal. The case $\beta = 2$ corresponds to a propagation model called *free-space propagation*, corresponding to a physical reality where the propagating medium of the wave is *free-space*, i.e. the emitter and the receiver are in an otherwise empty environment. In reality, however, the power received at distance d decreases much faster. A striking example is when a reflection of the signal on the ground occurs and the horizontal distance d between the transmitter and the receiver becomes considerably large compared to the heights of the antennas, see Figure 2.3. In that case, the received power decreases as d^{-4}, as explained in [126, Section 2.1.5].

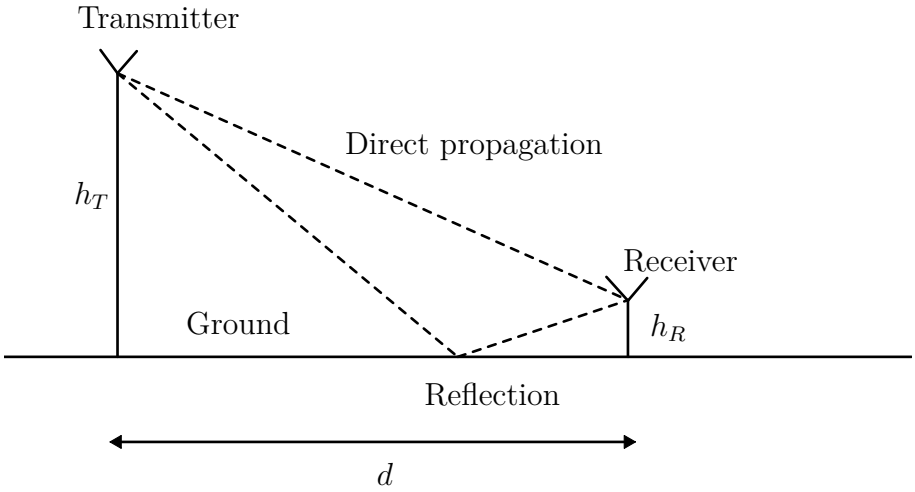


Figure 2.3 – Reflection of a radio wave on the ground. When $d \gg h_T$ and $d \gg h_S$, the power received at a distance d from the transmitter has a path loss function $\ell(d) \propto d^{-4}$.

The path loss exponent may also strongly depend on the radio frequency used for transmission, as suggested by numerous studies [42, 90, 112]. Typical values at 900 MHz and 1.9 GHz can be found in [61, Table 2.1] and the references therein. For higher range of frequencies, path loss exponents tend to be larger [41]. Finally, [61, 126] suggest that higher antennas result in smaller path loss exponents.

Shadowing

As we have previously seen, the power received decreases with the distance run by the signal due to path loss. However, path loss models only take into account the distance run by the signal and eventual changes of the medium of propagation. Another critical fact to take into account is the presence of physical obstacles on the path of the wave. Indeed, according to the laws of electromagnetism, when a wave encounters a physical obstacle, it may absorb some part of the signal and only transmit a reduced power of the original signal. Thus, the physical environment plays a critical role: indeed, imagine two antennas separated from a distance d. In a rural environment, there may be a hill between the antennas. Now, picture the scene in a densely populated urban area but with the same distance d between the antennas. For the same distance d, they may now be separated by dozens of buildings and hence the scenario is completely different, even though the ground distance between the transmitter and the receiver remains the same.

Most of the time, radio propagation happens in complex environments where the density of physical obstacles hugely varies. Indeed, the location, sizes, reflecting properties and material of all the physical obstacles encountered between the source and a destination are difficult to predict. As a matter of fact, one must often resort to random models to take into account the blockage of signals by obstacles. Hence, in addition to path loss, the radio signal also experiences random attenuation due to the presence of physical obstacles: this phenomenon is called *shadowing*[2]. A model often used to compute the received power

[2]The terminology comes from the fact that the attenuation of radio signals due to physical obstacles is similar to the effect of clouds partly blocking the sunlight in the sky.

while taking shadowing and path loss into account is the so-called *log-distance path loss model* [61, Section 2.8].

We end this subsection with some very useful terminology that will be used multiple times in the models studied in this thesis. Picture an emitter and a receiver, say two antennas. When there is no physical obstacle on the straight line between these two antennas, the radio signal between the emitter and the receiver can propagate on the straight line without being blocked: this is called *line-of-sight* (LOS) propagation. On the contrary, when the direct path between the two antennas is obstructed by physical obstacles, the resulting signal propagation is called *non-line-of-sight* (NLOS) propagation. Sometimes, in a context of NLOS propagation between two antennas, to secure a satisfying signal quality at reception, the path run by the radio signal must be diverted and go around the physical obstacles that are on the direct line between the emitter and the observer. To prevent signal degradation, network equipment called *relays* are often placed on the diverted path. The role of these relays consists in receiving the signal and emitting it again towards the diverted path, with a possible power amplification.

Fading

In wireless communications, *fading* is a general term designating an attenuation of a radio signal over a wireless channel. Fading can have various origins such as the geometry of the environment, the presence of physical obstacles, the conditions of propagation of the signal, the weather, a change of propagation medium or the propagation of the signal from the emitter to the receiver by differents paths (this is referred to as *multipath propagation*). Most models used in wireless systems engineering try to capture this phenomenon by random variables and random processes. More details on such models can be found in [11, 21, 64].

2.1.4 Reception of the signal, interference and noise

Interference

One of the key concepts in the Physics of waves and hence in radio communication is *interference*. Roughly speaking, interference is the phenomenon resulting in a superposition of electromagnetic waves when several waves propagate simultaneously towards a common a receiver. By a superposition principle, the signal received can be identified with the sum of all signals arriving at that point. The combination of all of these signals will result in a single wave with different characteristics than each one of the waves arriving at the receiver.

In wireless (cellular) networks, there is significant interference. Indeed, each transmitter-receiver pair communicates via a wireless channel over the air, and there is no physical medium for a communication. Furthermore, many users are communicating simultaneously and so many signals that are not relevant for a given communication can be received by a UE. As a matter of fact, dealing with interference is a central question in the design of cellular networks [66]. In cellular networks, interference may happen between different receivers communicating with a single transmitter, or between different transmitters

communicating with a common receiver, or even between different wireless pairs being in the same cell or not.

Noise

In addition to interference, all users of a cellular network are exposed to background parasite signals that are unwanted and have nothing to do with the network. In the literature, all of these signals are grouped under the terminology of *noise*. Noise can come from various sources, such as:

- Thermal motion of charged particles in electrical conductors: this is called *thermal noise*
- Statistical fluctuations of the electric current
- Parasite signals occurring over different frequency ranges than the radio spectrum
- Solar radiation

Due to its various origins and unpredictable nature, noise is most of the time modelled as a stochastic process [126].

Signal-to-interference-plus-noise ratio

Consider a cellular network and a transmitter-receiver pair. For instance, the receiver may be a UE and the transmitter its serving BS. Denote by S the signal power received by the user. Typically, as suggested in [21, Equation 2.8], S takes the form $S = PH_0\ell(d_0)$, where P denotes the transmit power of the BS, H_0 the fading coefficient of the wireless channel between the user and its serving BS, ℓ denotes the path loss function of the wireless channel and d_0 denotes the distance between the UE and the BS. Denote by I the interference power at the user's location and by N the power due to noise. A critical quantity used for performance assessments in wireless networks is the so-called *signal-to-interference-plus-noise ratio* (SINR), defined by:

$$\text{SINR} = \frac{S}{I + N} \tag{2.1.1}$$

The distribution of the SINR at a user's location is related to many quality metrics of the wireless channel but also of cellular networks in general [66]. A prominent example is the following: under general assumptions on the statistical model used for the wireless channel (e.g. see [21, Section 5.2]), the probability that a message is successfully transmitted on the channel is defined by the probability that the SINR exceeds some threshold θ. In other words, it is a function p_s of some threshold θ, defined by:

$$p_s(\theta) := \mathbb{P}(\text{SINR} > \theta) \tag{2.1.2}$$

p_s is called *transmission success probability*. The complement probability $1 - p_s$ is called *outage probability*. Depending on the stochastic modelling chosen, these quantities have various distributions. Examples of derivations of distributions for the SINR, the transmission success probability or the outage probability can be found in [10, 21, 64].

2.1.5 Capacity

Performance assessment of cellular networks is also done by the assessment of the wireless channels constituting the network. In the literature, a wireless channel is often modelled via so-called *input/output models*. In other words, the wireless channel is a black box transforming a sequence of information inputs $(X_i)_{i\geq 1}$ sampled at discrete time indices $i \geq 1$ into outputs $(Y_i)_{i\geq 1}$ by some transformation specific to the channel. Most of the time, the input signal is first encoded by the channel in binary language (i.e. into sequences of 0s and 1s called *bits*), then transformed by the channel, and decoded at the end of the channel, so as to be reconstructed as a proper signal (rather than a binary sequence which only the channel can work with). A critical quantity for evaluating the performance of a channel is its *capacity*, which is defined as the maximal rate at which information can be transmitted over the channel without error in encoding or decoding. In other words, capacity is a measure of how much the wireless channel can handle lots of simultaneous connections. At a more global scale, it is customary in engineering to talk about the *capacity of a (cellular) network* as a measure of how the network can handle traffic load, i.e. how many simultaneous connections the network can handle at the same time. The greater the capacity, the more people will be able to reliably communicate at the same time.

One of the most celebrated channels in the literature is the *additive white Gaussian noise* (AWGN) channel, defined in the following way:

$$Y_i = X_i + Z_i, \tag{2.1.3}$$

where the random variables $(Z_i)_{i\geq 1}$ are i.i.d. Gaussian random variables with 0 mean and variance N_0 corresponding to the *spectral density of noise*, i.e. the noise power per frequency unit. In the AWGN channel, there is no interference occurring and the channel does not necessarily operate at a single frequency f but may on frequencies belonging to some interval of the form $[f - B/2; f + B/2]$, where f is called the *central frequency* and B is called the *bandwidth*. For the AWGN channel, the celebrated Shannon-Hartley theorem [36, Chapter 10] states that the capacity C is given by:

$$C = B \log_2 \left(1 + \frac{S}{N_0 B}\right), \tag{2.1.4}$$

where S denotes the received signal power. Since there is no interference in the AWGN channel, the SINR ratio simplifies into the *signal-to-noise ratio* (SNR) given by

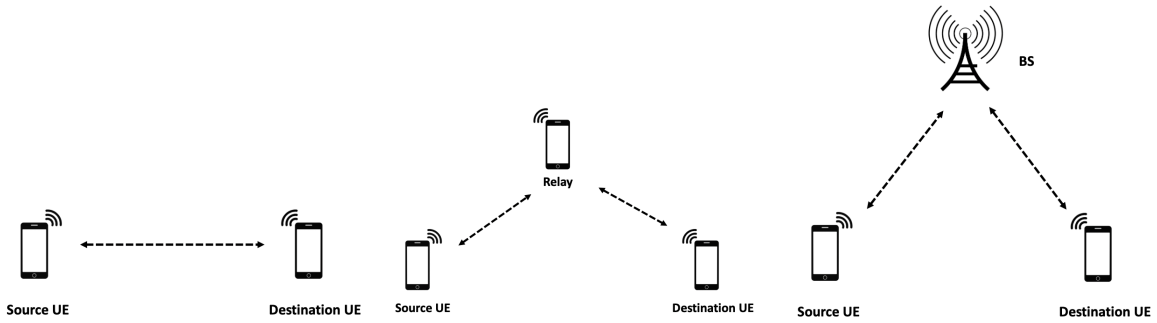
$$\mathrm{SNR} = \frac{S}{N} = \frac{S}{N_0 B},$$

where we have used the fact that the total noise power N is given by the product $N = N_0 B$, since the spectral density of noise N_0 is the noise power per frequency unit and the total useful range of frequencies of the channel is given by the bandwidth B. As a matter of fact, the capacity of the AWGN channel also writes as $C = B \log_2 (1 + \mathrm{SNR})$. This result shows another example of how the distribution of the S(I)NR is of great interest in the design of wireless networks.

2.2 Device-to-Device (D2D) communications

2.2.1 General principles

Device-to-Device (D2D) communications are defined as direct communications between two UEs (e.g. smartphones, tablets ...) without the need for the signal to go through the BSs or the core network [92], contrary to classical cellular communications, see Figure 2.4. Depending on the technology used for D2D communications (WiFi, Bluetooth, 4G ...)



(a) Single-hop D2D communciation

(b) Multihop D2D communication.

(c) Classical cellular communication.

Figure 2.4 – Oversimplified view of D2D communications (single hop and multihop) and cellular communications. In D2D communications, the signal is not routed through a BS or the core network. In classical cellular communications, the signal is routed through the BS and the core network (not represented here).

the range of a D2D link may vary but remains short compared to the range of a wireless cellular channel. As a matter of fact, two situations occur in D2D communications:

- The two devices communicating are sufficiently close to communicate directly via a single D2D link: this is called *single hop D2D communication*, see Figure 2.4a.
- The two devices communicating are at distance greater than the range of a D2D link. In that case, establishing D2D communications remains possible via a chain of consecutive D2D links where some devices located between the transmitter and the receiver intervene and relay the signal: this is called *multihop D2D communication*, see Figure 2.4b. Protocols are proposed in the literature [5, 107, 116] to establish multihop D2D and make sure that the intermediate devices, called *relays*, do not intercept the communication but only relay it. More advanced scenarios, where fixed antennas also play the roles of relays, have also been proposed [96, 133].

Initially proposed for network performance enhancements and extending coverage via multihop and relays [9, 131], D2D is also seen as a major component of the future 5G network, hence motivating an intense research activity [76, 125]. In the context of 5G networks, D2D is perceived as a way to meet the users' increasing data demand: over the past few years, the use of smart devices has been exponentially growing and the proportion of mobile users owning at least two devices (e.g. phone and tablet or phone and another connected object) is booming. Hence the need for capacity enhancements, network coverage extensions, better spectrum management, better quality of service ... The

picture is however much broader: the emergence of new use-cases taking advantage of users' proximity, such as content sharing in crowded environments, Machine-to-Machine communications (e.g. for autonomous cars) or ultra-low latency services, makes D2D a promising tool for the operators to enable new services on their networks [8, 80].

In the literature, D2D deployment scenarios in cellular networks fall into two main categories based on the frequency spectrum used by D2D:

- *Inband D2D*: In this scenario, licensed cellular spectrum is used by operators both for cellular communications and D2D communications.
- *Outband D2D*: In this scenario, cellular communications occur on the operators' licensed spectrum, while D2D communications occur on unlicensed spectrum.

We will now shortly review the pros and cons of each scenario and give further references for the interested reader.

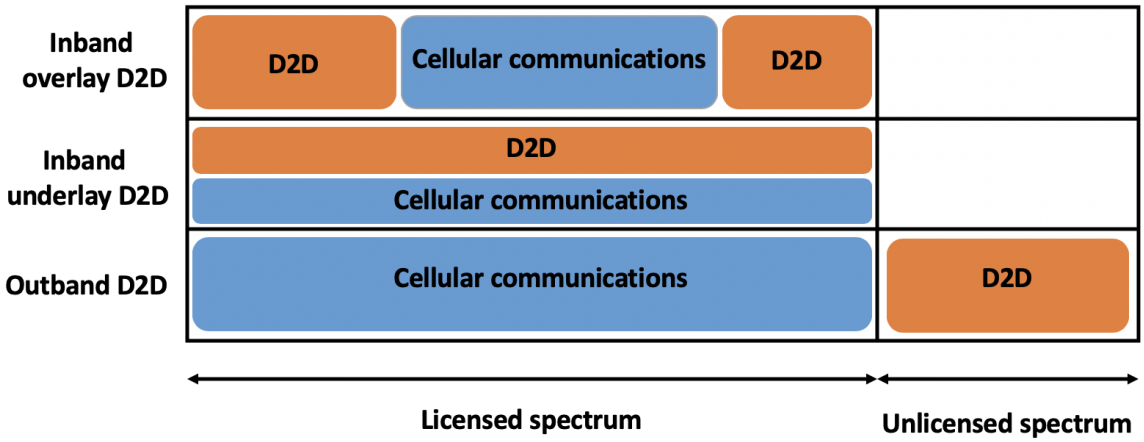


Figure 2.5 – Comparison between inband overlay, inband underlay and outband D2D.

2.2.2 Inband D2D

The literature on inband D2D contains the majority of the available references on D2D. The main advantage for operators in using inband D2D is that, by using their licensed spectrum for D2D communications, they have a good control on many key performance indicators [8]. However, by using the same frequency spectrum as cellular communications, inband D2D may cause interference between cellular communications and D2D communications and could harm the quality of service of the operators.
In order to solve the interference problems, two sub-scenarios for inband D2D have been considered in the literature:

- *Inband underlay D2D*: D2D and cellular communications operate on the operators' entire licensed spectrum and there is no part of the spectrum specifically dedicated to D2D communications, see Figure 2.5. The main challenge of this idea is of course the management of interference via clever algorithms [8, 92]. Observations are that the highest capacity gains are achieved with inband underlay D2D, since

the spectrum efficiency is maximised (all the available spectrum is reused for D2D in addition to cellular communications) [92].

- *Inband overlay D2D*: D2D and cellular communications still share the operators' licensed spectrum, but the available spectrum is partitioned: some part is dedicated to cellular communications and the other part to D2D communications, see Figure 2.5. This approach solves the problem of interference between cellular and D2D communications but reduces the resources available for cellular communications and is not the most efficient in terms of spectrum management [8, 92].

The literature on network performance enhancements via inband D2D is wide: the surveys [6,8,44,51,76] provide good summaries of the considered performance indicators and many further references.

2.2.3 Outband D2D

The literature on outband D2D is less substantial.
In outband D2D, D2D communications use unlicensed frequency bands, such as the industrial, scientific and medical (ISM) radio bands [8, 134], see Figure 2.5. The main advantage of doing so is of course avoiding the interference between cellular and D2D links. Moreover, D2D using different frequency bands, simultaneous D2D and cellular communications are possible in outband D2D. However, the use of unlicensed bands allows less control of the operators' key network performance indicators, hence a lower efficiency and reliability of D2D communications. Moreover, outband D2D is only feasible for devices having two wireless interfaces (e.g. a cellular interface and another technology such as WiFi or Bluetooth) [44] and hence is more energy consuming. The survey [8] proposes two subscenarios for outband D2D communications:

- *Outband Controlled D2D*: Though occurring on unlicensed spectrum, D2D communications are managed by the cellular network, so as to partially solve the reliability issue. This could however lead to network congestion. Outband Controlled D2D currently concentrates almost all of the research activity on outband D2D.

- *Outband Autonomous D2D*: There is no intervention of the cellular network in managing the D2D links. This is the easiest deployment scenario for outband D2D (as it does not require any changes in the core network) but the reliability of D2D links is absolutely not guaranteed.

A detailed summary of the research activity on outband D2D can be found in [8, 76].

2.2.4 Device-to-Device and crowd-networking

Connected devices, such as smartphones, tablets, smartwatches or intelligent sensors for domotics have been rapidly and massively adopted worldwide. In the mean time, subscriptions with a high amount of data have become available at reasonable costs. These two factors combined have resulted in an exponential growth of global mobile data traffic, as highlighted by numerous studies. For instance, Cisco calculated a 71% surge of the global mobile data traffic from 2016 to 2017 through its Visual Networking Index [2]. Between 2012 and 2017, this same traffic has been multiplied by 17. Meanwhile, the number

of connected devices undergoes an explosive growth too: Ericsson [3] calculated that the total number of connected devices in the world reached 15.1 billion in 2015 and will be above 30 billion by 2022.

An idea made possible by the explosion of the number of connected devices is the use of multihop D2D-enabled networks in so-called *crowd-networking* scenarios, where the different stakeholders of a network can collaborate to ensure a global connectivity and a better quality of service. More concretely, for an operator, this is equivalent to benefiting from the density of its existing customers for two purposes:

- Ensure a better quality of service by offloading some of the traffic from the BSs to D2D links.
- Ensure a better coverage in areas where less core network infracstructure has been deployed.

Crowd-networking, sometimes also called *cooperative networking*, is a relatively recent idea [93]. At the dawn of 5G networks, where D2D is seen as a key paradigm for solving many technical issues, cooperative networking using D2D communications has gained much interest in the literature [30, 63]. Indeed, thanks to the D2D paradigm, it seems relevant to take advantage of the spatial proximity of users in crowds to enhance the global performance of cellular networks. Especially in the outband mode, D2D is fairly accessible as it uses unlicensed spectrum, and is also relatively cheap to set up for operators, since there is no need to heavily invest in core network infrastructure and BSs. What is more, thanks to D2D and the possibility for a user of the network to serve as a relay between any transmitter and receiver in the network, operators have access to a vast number of somehow "cheap" relays: their own customers! On an economic point of view, one may then think of operators incentivising their customers (e.g. via a discount in the price of their mobile phone subscription) into serving as relays: such scenarios would be "win-win" on all perspectives. Thus, using D2D for crowd-networking scenarios seems of critical importance for operators and may lead to great economic opportunities.

Let us think even beyond: we can think of new actors willing to rely on D2D and a sufficiently large density of customers to set up fully functional networks while avoiding heavy infrastructure investments which had to be made by historical operators. If such perspectives are feasible, these "neo-operators" could be able to offer a service comparable to the one offered by traditional operators, at a much lower price! Their arrival would then deeply disturb the market and force traditional operators to re-invent themselves. Pretty much in the same way as Uber brought its own revolution in ride-hailing while not properly owning a single one of the cars used by its drivers, neo-operators relying on D2D could *uberise* the telecommunications market by offering to their customers a functional telecommunications service while not owning any particular network equipment.

Such scenarios show that D2D represents both particularly interesting economic opportunities but also threats for traditional operators, hence motivating a deeper analysis of the feasibility of large-scale D2D networks.

Chapter 3

Stochastic geometry and percolation for Device-to-Device networks modelling

The huger the mob, and the greater the apparent anarchy, the more perfect is its sway. It is the supreme law of Unreason. Whenever a large sample of chaotic elements are taken in hand and marshaled in the order of their magnitude, an unsuspected and most beautiful form of regularity proves to have been latent all along.

Francis Galton

In this chapter, we lay the foundations of the models that have been studied in this PhD thesis and then introduce these models.
While the historical approach for mathematically studying the connectivity of wireless networks has been based on percolation theoretic tools, more refined models have been considered over the years. In particular, stochastic geometry tools led to considering supplementary random models for the typology of the territory covered by a given network. This gave birth to *doubly-stochastic* percolation models where both the support of the network and the locations of network nodes are random. In this thesis, we have brought our own contributions to the field with the introduction and study of new models for addressing the question of large-scale connectivity of D2D networks.
First, we introduce some percolation models, related to our work, that have initially been proposed for modelling the connectivity of wireless networks in general. Then, we review more recent works where the underlying topology of the support of the network plays a critical role. In these models, the support of the network is most of the time given by a random tessellation. Finally, we introduce our own contributions by presenting the models and the research perspectives at stake in this thesis.

3.1 The percolation approach for studying the connectivity of wireless networks

3.1.1 From random graphs to Gilbert's graphs

From a mathematical point of view, a complex network, and, as such, a telecommunications network, can be seen as a collection of nodes, with interactions between these nodes representing possible communications on the network. Thus, modelling such networks by connectivity graphs makes sense. However, most of real-world networks are rather large, making it impossible to describe them in detail. Moreover, one may also need to be able to resort to tractable models that faithfully reproduce large-scale properties of actual networks. To deal with such technicalities and to accurately reproduce the complexity of actual networks, models based on *random graphs* have been considered.
Modelling complex networks and, in particular, wireless communications networks by random graphs is not exactly a recent idea: it dates back to the late 1950s and early 1960s. Though a few papers appeared in the mathematical literature before that time, the works from Erdős and Rényi [43] and Gilbert [54] are considered to have founded the field of random graph theory. In these papers, two approaches of the notion of *random graph* have been considered:

- In the model of Erdős and Rényi, for given numbers n of vertices and $N \leq \binom{n}{2}$ of edges, a graph is selected uniformly at random among the set of all graphs with n vertices and N edges, i.e. with probability $\frac{1}{C_{n,N}}$, where $C_{n,N} = \binom{\binom{n}{2}}{N}$. This model is usually denoted $G(n, N)$.

- In the model of Gilbert, a parameter $p \in [0, 1]$ is given and, for a number n of vertices, each of the possible $\binom{n}{2}$ edges is kept with probability p, independently from all other edges. This model is usually denoted $G(n, p)$.

Both of these models are referred to as *Erdős-Rényi* graphs in the mathematical literature, though Gilbert's variant, discovered independently but further studied by Erdős and Rényi, has gained more interest.

Seeing random graphs as models of wireless networks, one may think of the vertices representing nodes of the network (e.g. BSs or UE) and the edges symbolising possible connections. However, the Erdős-Réyni approach remains rather simple for modelling actual wireless networks, mainly for three reasons. First, the locations of the vertices of the graph is not taken into account, while, in real-world networks, the location of network nodes plays a critical role in large-scale properties of the network itself. Moreover, in Erdős-Rényi graphs, the probability that some edge exists is a common global constant for every edge[1]. In particular, this probability is independent of the nodes themselves. Finally, in real-world networks, a critical parameter that must be taken into account is

[1]Let $e = (v_i, v_j), 1 \leq i \neq j \leq n$ denote an edge. In the $G(n, p)$ model, the probability that e is present is by definition equal to p. In the $G(n, N)$ model, since all graphs with n vertices and N edges are equally likely, the probability that e exists is equal to ratio between the number of graphs with n vertices and N edges where e is present and the total number of graphs with n vertices and N edges. It is easy to see that this probability is just $N/\binom{n}{2}$.

the range of wireless communications, which is not considered in the Erdős-Rényi models.

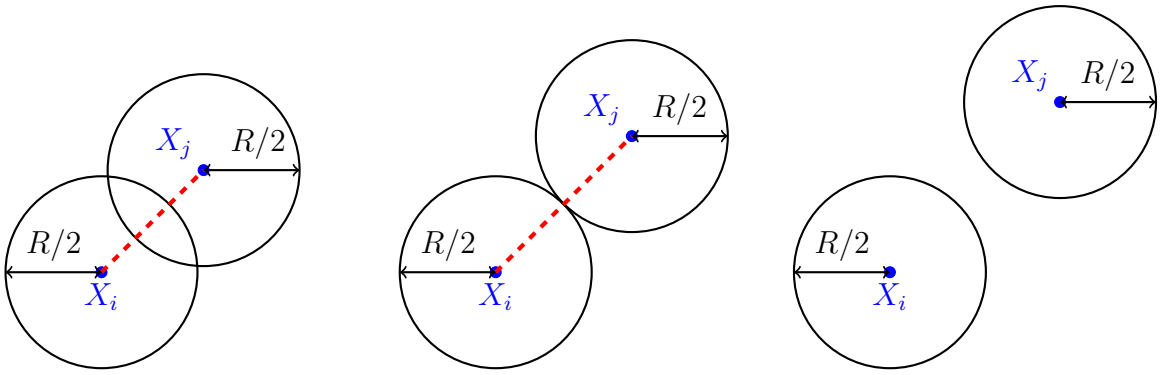
To tackle these issues, Gilbert published a seminal work [55] in 1961. In this paper, entitled *Random plane networks*, Gilbert precisely underlined the necessity to take into account the locations of BSs but also their range in the modelling. He considered a random network whose nodes are given by a homogeneous Poisson point process of positive finite intensity $\lambda > 0$ in the plane $\mathbb{R}^2$ and declared two nodes of the network to be connected by an edge if their mutual Euclidean distance is less than some threshold $R > 0$. The motivation of this work was to model a telecommunications network of BSs with communication range R and connections possible only along the edges of the random graph arising from the previous construction. Denoting by $E = \lambda\pi R^2$ the mean number of points in a circle of radius R (in other words the coverage zone of a station), Gilbert studied the behaviour of the probability $P(N)$ that some connected component containing at least N points exists. Resorting to numerical simulations, he noticed that the graph must have some *phase transition* in E and that the limit $P(\infty) := \lim_{N\to\infty} P(N)$ may not be zero in some cases. Finally, he proved the existence of some threshold E_c separating two regimes:

- when $E < E_c$ $P(\infty) = 0$ and hence the modelled network can only provide local communications.
- when $E > E_c$, $P(\infty) > 0$, so that arbitrarily long-range communications are possible with positive probability.

In the modern literature, Gilbert's model is often referred to as *Gilbert's graph* or *Gilbert's disk model*, in so far as an equivalent condition for two points X_i, X_j of the underlying Poisson point processes to be connected is that circles with radius $R/2$ respectively centered at X_i and X_j overlap, see Figure 3.1. Such a model for studying the connectivity of wireless networks answers to critics that could be made to models based on Erdős-Rényi graphs: the location of the stations, possibly spread over a wide area, as well as the range of stations are taken into account. Furthermore, the probability that two stations can communicate is no longer a global constant.

3.1.2 Continuum percolation and wireless networks

With the introduction of the aforementioned model, Gilbert gave birth to *continuum percolation theory*. Since then, the problem has been studied in several works [15, 67] and a comprehensive overview has been developed in the textbook [99]. In a more theoretical fashion than Gilbert's original idea, the main model at stake in continuum percolation is the so-called *Boolean model* (or *Poisson Boolean model*). This Boolean model can be thought of as an extension of Gilbert's model with random connection radii and is defined as follows: a homogeneous Poisson point process X in $\mathbb{R}^d$ with positive finite intensity $\lambda > 0$ is given, as well as some non-negative random variable ρ independent of X. Denoting by $(X_i)_{i\geq 1}$ the points of X, we consider i.i.d. copies $(\rho_i)_{i\geq 1} \sim \rho$ distributed like ρ. Each point X_i of X is the center of a closed Euclidean ball $\mathcal{B}(X_i, \rho_i)$ with random radius ρ_i, see Figure 3.2. We say that the Boolean model is *driven* by X and we denote the model by (X, ρ, λ).



(a) $\|X_i - X_j\|_2 < R$ and so the two circles of radius $R/2$ respectively centered at X_i and X_j overlap. Communication between X_i and X_j is possible.

(b) $\|X_i - X_j\|_2 = R$ and so the two circles of radius $R/2$ respectively centered at X_i and X_j are tangent. Communication between X_i and X_j is possible (limit case).

(c) $\|X_i - X_j\|_2 > R$ and so the two circles of radius $R/2$ respectively centered at X_i and X_j do not overlap. Communication between X_i and X_j is not possible anymore.

Figure 3.1 – Simplifying the conditions of mutual Euclidean distance into overlapping disks in Gilbert disk model.

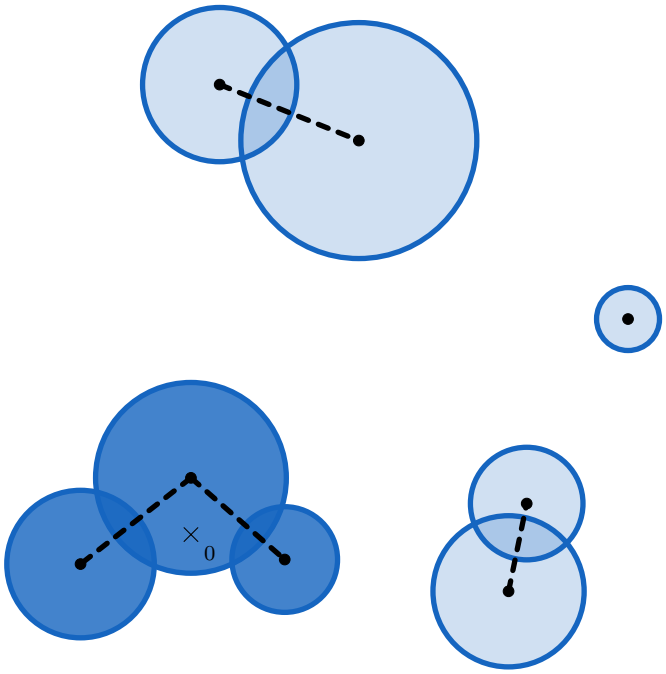


Figure 3.2 – A realisation of a Boolean model in the plane $\mathbb{R}^2$. The blue-shaded region is the occupied region $\mathcal{C}$. The occupied component $W =: W(\{0\})$ of the origin 0 is the darker blue-shaded region. The black points represents the atoms of the Poisson point process X and the cross represents the origin 0. The dashed edges represent all possible connections between the points of X. The figure is strongly inspired of [99, Figure 1.1] but reproduced by our own means.

Considering a realisation of a Boolean model in $\mathbb{R}^d$ is equivalent to partitioning the space $\mathbb{R}^d$ into two regions:

- The fraction of the space covered by the random balls $\mathcal{B}(X_i, \rho_i)$, $i \geq 1$ is called the *occupied* region and is denoted by $\mathcal{C}$.
- The complement of the occupied region is called the *vacant* region.

The connected components of the occupied region are called the *occupied components*. For a subset $A \subset \mathbb{R}^d$, we denote by $W(A)$ the union of all occupied components having non-empty intersection with A. We denote by $W := W(\{0\})$ the *occupied component of the origin 0*. Two points x and y belonging to the same occupied component are said to be connected[2]: this is denoted by $x \leftrightsquigarrow y$.

As in discrete percolation, one is interested in the probability of existence of an infinite occupied component. Denoting by $\mathbb{P}^0$ the Palm probability of the underlying Poisson point process, the *percolation function* is defined by $\theta_\rho(\lambda) := \mathbb{P}^0(W \text{ is unbounded})$. As in the discrete model, the Boolean model is said to *percolate* if $\theta_\rho(\lambda) > 0$. In the same fashion, the *critical intensity* $\lambda_c = \lambda_c(\rho)$ of the Boolean model is the smallest intensity of the underlying Poisson point process ensuring percolation of the Boolean model, i.e. $\lambda_c(\rho) := \inf\{\lambda > 0\,, \theta_\rho(\lambda) > 0\}$. The non-triviality of the phase transition in the Boolean model (and hence of λ_c) depends on the probability distribution ρ generating the random radii of the balls: general results can be found in the textbook [99] and the references therein. A more recent exposition of the model can be found in [10, Chapter 3]. In particular, two of the key results of continuum percolation theory are the following ones:

Proposition 3.1.1 ([99, Proposition 3.1]). *In a Boolean model in $\mathbb{R}^d, d \geq 1$, the whole space $\mathbb{R}^d$ is covered almost surely if and only if $\mathbb{E}(\rho^d) = \infty$.*

Theorem 3.1.2 ([99, Theorem 3.3]). *Consider a Boolean model in $\mathbb{R}^d, d \geq 2$, such that $\mathbb{E}(\rho^d) < \infty$. Then we have the following:*

1. *If $\mathbb{P}(\rho = 0) < 1$, then $\lambda_c(\rho) < \infty$, i.e. there exists a non-trivial supercritical phase for the percolation of the Boolean model.*
2. *If $\mathbb{E}(\rho^{2d-1}) < \infty$, then $\lambda_c(\rho) > 0$, i.e. there exists a non-trivial subcritical phase for the percolation of the Boolean model.*

As in the discrete case, the exact value of $\lambda_c(\rho)$ is far from being easy to get. Even in the simple context where ρ is almost surely constant, i.e. $\rho = r$ almost surely for some $r > 0$, the exact value of the resulting critical intensity $\lambda_c = \lambda_c(r)$, though being non-trivial, is out-of-reach by analytical means. As a matter of fact, numerical simulations are often used to get estimates and approximations for values of the critical intensity [17, 100, 105, 106]. Finally, in the context where the radii are almost surely constant, the Boolean model reduces to Gilbert's model and extensions to different point processes than Poisson have been considered, for example sub-Poisson [22, 23], Ginibre and Gaussian zero [53] and

[2] Actually, the exact terminology is *connected in the occupied region*. In some percolation textbooks, authors sometimes also define the concepts of vacant components and connectivity in the vacant region. This will not be useful for our developments and we thus adopt a simpler terminology.

Gibbsian [77, 122].

As far as wireless networks are concerned, mathematical models for connectivity based on a percolation approach have flourished since Gilbert's initial work. In mathematical models of wireless networks, the nodes of the networks are usually given by a point process, very much as in the classical Boolean model. The coverage zone of a network equipment is then given by some (possibly random) geometric shape, which can be interpreted as the radiation pattern of a radio equipment (a BS, a UE ...). The network is then represented by a random connectivity graph (two nodes of the network being joined if they respectively belong to the coverage zone of the other node) and percolation of this graph is interpreted as large-scale connectivity of the network. Finally, the critical intensity λ_c is interpreted as the minimal density of equipment (and/or network users) ensuring large-scale connectivity of the network.
Over the years and the research activity, various extensions of the Boolean model have been considered: for instance, patterns that are not circular have been considered in [17]. What is more, the percolation approach to wireless networks modelling has gone far beyond investigating the non-triviality of the critical user density λ_c: capacity of wireless networks has been investigated by percolation techniques in [46] and design questions have been considered in [45, 65]. Another beautiful example in the use of percolation techniques is the so-called *shape theorem* established in [135], investigating how many hops are needed to connect two points belonging to the same connected component of the Boolean model. More precisely, in the supercritical regime of the Boolean model, for two points $x \neq y$ of the infinite connected component[3] $\mathcal{C}$, the ratio between the number $N(x, y)$ of hops needed to connect x and y in $\mathcal{C}$ and the Euclidean distance $\|x - y\|_2$ converges in Palm probability to some constant $\mu = \mu(\lambda)$ only depending on λ as $\|x - y\|_2 \to \infty$. From an applied perspective, in the context of wireless networks modelling, this constant called *shape factor* simply represents a proportionality factor between the number of hops connecting two nodes of the network and the distance separating them and is of particular interest in the design of wireless networks.

3.1.3 Refined connectivity conditions: SINR percolation

As we have seen above, the percolation approach to the mathematical modelling of connectivity in wireless networks has gained much interest among the mathematical community. In continuum percolation models based on Gilbert's approach, the connectivity between two nodes of the network only depends on their mutual distance. In particular, the closer the transmitter and the receiver are, the more likely they will be connected. In real-world networks, however, this is not always the case. For instance, if there are a lot of other wireless pairs at close distance from a transmitter-receiver pair, a significant amount of interference may be present and make the transmission between the transmitter and the receiver unsuccessful, even if they are pretty close to one another. As a matter of fact, the connection between a transmitter and a receiver may not only depend on their mutual distance, but also on the locations of all the other nodes of the network that generate interference possibly disrupting the quality of the connection. This physical reality is

[3] It is a well-known fact that such a component, if it exists, is almost surely unique, for ergodicity reasons, e.g. see [99, Theorem 3.6].

not taken into account in classical continuum percolation models. Hence, refined models should be considered.

To tackle such questions, another variant of Gilbert's approach for Poisson point processes has been considered in [39, 40, 47]: the SINR graph. It is defined as follows: first, a homogeneous Poisson point process X in $\mathbb{R}^2$ with positive finite intensity $\lambda > 0$ is given. As in Gilbert's approach, X will represent the nodes of the network. Then, the following additional model parameters are considered:

- A continuous decreasing path loss function $\ell : \mathbb{R}_+ \to [0, 1]$. Recall from Chapter 2 that such a function describes the attenuation of the signal strength over the distance between a transmitter and a receiver. To ensure that the model is not degenerate, the two additional assumptions $\ell(0) > \tau N/P$ and $\int_0^\infty x\ell(x)dx < \infty$ are often made[4].
- All nodes of the network are assumed to transmit a signal at a common deterministic constant power $P > 0$.
- The power of the background noise is given by some deterministic parameter $N \geq 0$.
- A parameter $\gamma \in [0, 1]$, called *orthogonality factor* and given by the wireless technology of the system, represents the measures taken to mitigate interference in the network.

Using the aforementioned parameters, the SINR for the transmission from some node $x \in X$ to another node $y \in X$ of the network can be expressed as:

$$\mathrm{SINR}(x, y) := \frac{P\ell(\|x - y\|_2)}{N + \gamma \sum\limits_{z \in X, z \notin \{x,y\}} P\ell(\|z - y\|_2)}, \tag{3.1.1}$$

where, as before, $\|\cdot\|_2$ denotes the Euclidean distance. In (3.1.1), the numerator is equal to the power transmitted from x to y and, in the denominator, the sum is equal to the interference generated at y by all the other nodes of the network. The orthogonality factor γ represents the measures taken to mitigate interference in the network (e.g. orthogonality of signals, distinct wireless channels, beamforming ...). The case $\gamma = 0$ corresponds to a complete cancelling of interference while the case $\gamma = 1$ corresponds to the absence of measures taken to mitigate interference.
Connectivity is then defined in the following way: given some SINR threshold $\tau > 0$, two network nodes x and y, i.e. points of the Poisson point process X, are connected if the SINR for both transmissions (from x to y and from y to x) exceeds the SINR threshold. In other words: x and y are connected if and only if $\mathrm{SINR}(x, y) > \tau$ and $\mathrm{SINR}(y, x) > \tau$. The random graph arising from this construction, called the *SINR graph*, features arbitrarily long-range dependencies, contrary to Gilbert's graph: the existence of an edge between two nodes of the network now also depends on the location of all the other nodes of the network through the SINR ratio. Note that in the absence of interference, i.e. $\gamma = 0$, the SINR for both transmissions between two nodes x and y is equal to a common value:

$$\mathrm{SINR}(x, y) = \mathrm{SINR}(y, x) = \frac{P\ell(\|x - y\|_2)}{N}$$

[4]The scalar $\tau > 0$ represents the SINR threshold allowing for successful communications, see the next paragraph.

and so the condition $\mathrm{SINR}(x,y) > \tau$ is equivalent to $\|x-y\|_2 < \ell^{-1}(\tau N/P)$ where ℓ^{-1} denotes the inverse function[5] of ℓ. Thus, the SINR graph in the absence of interference is equivalent to a Gilbert graph with radius $r_B = \ell^{-1}(\tau N/P)$ or, equivalently, to a Boolean model with constant radii $r_B/2$.
Similarly to classical percolation, a connected path in the SINR graph is a sequence of Poisson points $x_1, \dots x_k$ such that x_i and x_{i+1} are connected by an edge whenever $1 \leq i \leq k-1$. A connected component or *cluster* is a maximal set of Poisson points $\{x_i : i \in I\}$ with the property that for all $i \neq j \in I$, x_i and x_j are connected by an edge. The SINR graph is said to *percolate* if it contains an unbounded cluster with positive probability. However, contrary to classical (continuum) percolation, the boundary between the subcritical regime (where the SINR graph does not percolate) and the supercritical regime (where the SINR graph percolates) in SINR percolation is not delimited by a single critical intensity λ_c for the underlying Poisson point process, but also by a limiting orthogonality factor given by some function $\gamma^*(\lambda)$. More precisely, the main result is as follows:

Theorem 3.1.3 ([40, Theorem 1]). *Denote by $\lambda_c = \lambda_c(r_B)$ the critical intensity for percolation of the equivalent Gilbert graph when $\gamma = 0$. Then, for all $\lambda > \lambda_c$, there exists an orthogonality factor $\gamma^*(\lambda) > 0$ such that, for all $\gamma \leq \gamma^*(\lambda)$, the SINR graph percolates.*

An example of a phase-transition diagram for the SINR graph is given by Figure 3.3. Under additional assumptions on ℓ, general bounds can be determined for γ, as has been done in [40]. Moreover, refined versions of the SINR graph have also been considered and a more detailed exposition with extensions of the model can be found in [10, Chapter 8]. In particular, the case where the transmitted power P is deterministic but not the same for every node has been considered, as has been the case where transmitted powers are no longer deterministic but random [94].

3.2 Taking the underlying geometry into account: towards doubly-stochastic models

3.2.1 Random tessellations as street systems models

In the previous section, we have reviewed several percolation models for mathematically dealing with the question of connectivity of wireless networks in terms of a reduced number of parameters. In such models, the users of the network are modelled by the atoms of a certain point process (Poisson, Ginibre, Gaussian zero, Gibssian ...) in the plane $\mathbb{R}^2$ or more generally in $\mathbb{R}^d$ for some $d \geq 2$. A direct consequence of that modelling is that users can be located everywhere in the underlying observation space and even though they can be spread over a wide area, as would be the case in actual networks, their specific locations are not constrained to a particular domain. However, in real-world networks, users cannot be located everywhere: some part of the environment where users are located may be occupied by buildings, trees, roads ... Moreover, a significant part of users, especially in urban areas, are rather located on the *streets* of a *street system*, which has a given

[5] Since $\ell : \mathbb{R}_+ \to [0,1]$ is continuous and decreasing with $\ell(0) > \tau N/P$, $\ell^{-1}(\tau N/P)$ exists and is well-defined.

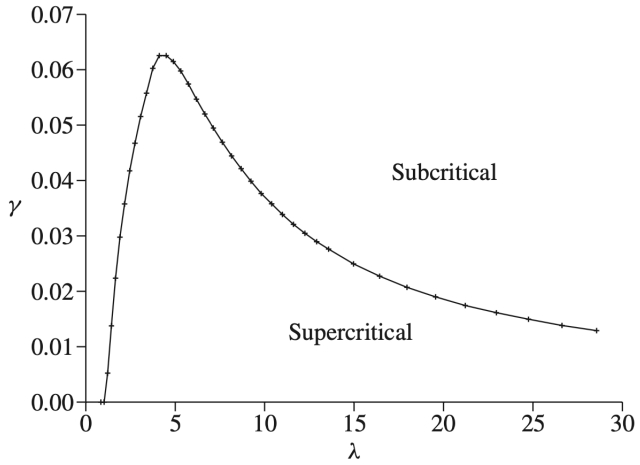


Figure 3.3 – An example of a simulated phase-transition diagram for the percolation of the SINR graph when the path loss function is given by $\ell(x) = \min(1, x^{-3})$. The curve represents the interference threshold $\gamma^*(\lambda)$ as a function of λ. Above this curve, the SINR graph is subcritical, i.e. it does not percolate. Below this curve, the SINR graph is supercritical, i.e. it percolates. This is very much different from the dichotomy $\lambda < \lambda_c$ (subcritical) versus $\lambda > \lambda_c$ (supercritical) of Gilbert's graph. Figure taken from [40].

topology. It thus makes sense to consider the topology of streets systems when modelling networks in general, and constrain network nodes, or at least network users, to be located on a restricted domain of the observation space.

Models of street systems have already been considered when solving telecommunications engineering problems and approaches for modelling street systems by statistically equivalent random tessellations have been proposed. By statistically equivalent, we mean that a given realisation of the chosen random tessellation may of course not look exactly like the modelled street system, but nevertheless has the same spatial characteristics. Examples of such characteristics are:

- The mean number of crossroads in the street system per unit area. Seeing the street system as modelled by a random planar tessellation, the mean number of crossroads per unit area in the street system corresponds to the intensity of vertices λ_0 of the tessellation.

- The mean number of streets per unit area. In the same way, it corresponds to the intensity of edges λ_1 of the tessellation.

- The mean street length per unit area. It corresponds to the line density L_A of the tessellation.

As has been shown in [59,60,130], large optimisation problems related to network planning and network deployment and planning can be solved in a much easier way by stochastic modelling and, in particular, by considering a random tessellation equivalent to the real-world street system at stake in the problem. The stochastic approach also allows to derive distributions of key quantities related to the design and analysis of real-world

networks, e.g. cable lengths required for deploying fixed networks. In [57], a general algorithm has been proposed to identity a map extraction of a real-world street system to a given random tessellation. Furthering this idea, statistical procedures have been proposed in [58] to determine the best possible fit for a given street system among several classes of tessellations models. Using the aforementioned tools and ideas, an application to wireless systems has been considered in [33], where the authors propose a procedure to analyse in detail the path loss function related to wireless propagation in various complex urban systems.

Considering that a whole street system can be modelled by a single random tessellation implies that the map from which the street system has been extracted can somehow be considered as a homogeneous whole. This is a highly counterintuitive assumption, especially in urban environments. Indeed, in a whole city, one can often distinguish several zones with different population densities, and hence different street systems. For instance, most major cities in Europe feature an hypercentre, then a less dense inner city and then, going further from the centre, suburban parts and rural fringes. Less populated areas, such as these rural fringes, typically feature sparser street systems than the centre of the city. The main question arising from this observation is therefore on how to segment a city map into several zones which can be considered as *morphologically homogeneous*. Each of these zones could then be modelled by a single random tessellation with given parameters, and different zones would be modelled either by the same type of random tessellation or not, with different parameters. Such issues have extensively been studied in the PhD thesis [35] and the associated works [32, 34]. In particular, a *city segmentation algorithm*, based on classical data clustering algorithms (K-means algorithms and spectral clustering), has been proposed in [35, Chapter 10] to cluster a city map into several zones of homogeneous morphology. Once this work has been done, statistical fitting procedures are run to identify each zone with a random tessellation model with determined characteristics $(\lambda_0, \lambda_1, \lambda_2, L_A)$. The city segmentation algorithm has been implemented in linear time and tested out against real city maps: most of the time, Poisson-Voronoi tessellations have proven to yield good fits of relatively dense urban street systems.

3.2.2 Doubly-stochastic percolation models for wireless networks

Considering the need for an appropriate modelling of street systems in mathematical models of wireless networks, it makes sense to think of *doubly-stochastic models*, in the following sense:

- A first layer of randomness could consist in a random tessellation or any appropriate random model for the support of the network.
- A second layer of randomness could consist in a point process modelling the nodes of the network, supported by the aforementioned random model for the street system.

If we consider a Poissonian model for the nodes of the network, very much as in the classical Gilbert graph [55] or the SINR graph [40], supported by a random tessellation, or, more generally, a random support, we arrive at models based on *Cox processes*.

Percolation of the Gilbert graph of Cox point processes has been studied very recently [70]. In this work, the authors considered the Gilbert graph in $\mathbb{R}^d$ of Cox point processes driven by a random intensity measure of the form $\lambda\Lambda$, where Λ is a random measure renormalised in such a way that $\mathbb{E}\Lambda\,[0,1]^d = 1$ and the scalar $\lambda > 0$ can be seen as an equivalent of the linear intensity of Cox points. The authors underlined that, without further assumptions on Λ, percolation of the Gilbert graph based on such point processes can fail. Furthermore, they proved that, under certain conditions of spatial decorrelation and of connectedness on the support of Λ, percolation of the Gilbert graph exhibits a non-trivial phase transition. Using the techniques developed in the previous work, SINR percolation for Cox point processes was studied in [128]. Assuming on the driving measure Λ the same conditions than those needed for percolation of the Gilbert graph, analogous results of SINR percolation available for Poisson point processes were derived for Cox point processes. In the line of thought of the previous works, additional extensions of classical models in the Cox case have also been considered very recently: in [74], SINR percolation with random powers has been considered, while the more general Boolean model with random communication radii has been investigated in [75].

On a more applied approach, Gilbert graphs based on Cox point processes supported by random tessellations are good models for networks of users communicating via multihop D2D mechanisms. In this fashion, a case of great interest considered in [70] is when the driving measure $\lambda\Lambda$ of the Cox point process is of the form $\lambda\Lambda(dx) = \lambda\nu_1(S \cap dx)$, where S denotes a stationary random tessellation in $\mathbb{R}^d$ and where ν_1 denotes the 1-dimensional Hausdorff measure. In such a case, the Cox points are supported by the random tessellation S and the scalar λ is equal to the mean number of Cox points per unit length of edge of S. Also, the number of Cox points on each edge e of S is a Poisson random variable with mean $\lambda\nu_1(e)$, with $\nu_1(e)$ denoting the length of e.
This case has been studied and investigated on a numerical perspective in [29]. The authors consider a street system given by a planar PVT or PDT, supporting a Cox process of network users X^λ driven by the random measure $\lambda\nu_1(S \cap dx)$, so that $\lambda > 0$ denotes the mean number of users per unit edge length. The multihop D2D network is then modelled by the Gilbert graph of this Cox point process, where two Cox points are connected by an edge if and only if they are at distance less than a communication radius $r > 0$. Finally, they address the following network engineering questions and answer them with mathematical formalism and numerical simulations:

- What is the minimal density of network users ensuring large scale connectivity of the network? This is answered by numerically estimating the critical intensity $\lambda_c = \lambda_c(r)$ for percolation of the Gilbert graph.

- What is the probability that a randomly selected user of the network is in the large connected component of the network? This is dealt with by estimating the percolation function $\theta(\lambda, r, \gamma) := \mathbb{P}^0(0 \leftrightsquigarrow \infty)$, where $\mathbb{P}^0$ denotes the Palm probability of the Cox process of network users.

- Given the distance between two points in the same connected component, how many hops are needed to connect them? Here, this question is answered by estimating

the shape factor $\mu(\lambda, r, \gamma)$ for the percolation of the Gilbert graph[6].

3.3 Our contribution: a new model for the study of connectivity of D2D networks in urban scenarios

3.3.1 Our approach

In the previous sections, we have reviewed various existing models of wireless networks available in the literature. As a first model, Gilbert's work started to make the case for a stochastic geometry and percolation approach of wireless networks modelling. However, we have noted that two main critics can be addressed to Gilbert's model:

1. The presence or absence of connection between two nodes only depends on their relative Euclidean distance, which is far from physical reality, e.g. because of interference.

2. The location of users is not constrained on any particular domain of $\mathbb{R}^d$, which is highly questionable in urban environments where users are located on streets.

While keeping up with a percolation approach, extensions have been proposed in the mathematical literature to address the previous points, by adopting one of the following strategies:

- Either put some emphasis on the modelling of more realistic connectivity conditions by using appropriate radio propagation models. This is for instance what has been done in classical SINR percolation. We call this approach a *telecommunications-oriented modelling*. Such models however fail to take into account a particular topology (e.g. a street system) for the support of the network.

- Or put some emphasis on a proper modelling of the support of the network and considering that users cannot be located everywhere in space. This is for instance what has been done by studying percolation of the Gilbert graph of Cox point processes. We call this approach a *geometry-oriented modelling*. However, such models fail to consider connectivity conditions which are more representative of physical reality.

Very few models did try to unite these both worlds by being telecommunications-oriented and geometry-oriented. Addressing the two previous points by considering more realistic connectivity conditions *and* a particular topology for the support of the network has only been considered very recently with the study of SINR percolation for Cox processes [128]. Even though this work considers interference, which is a good first step, some physical realities related to radio propagation such as fading or shadowing are still not taken into account. Moreover, in the engineering literature, the main available works treating the

[6]Contrary to the Poisson case, the theoretical existence of this shape factor has not been established for Cox processes. In the Poisson case, the existence of μ strongly relies on the subadditivity of the number of hops in the Gilbert graph. Since the number of hops remains subadditive in the Gilbert graph considered in [29], the authors assume that the existence of a similar shape factor in the Cox case is reasonable.

subject of D2D rather focus on possible applications of D2D communications, and most of the time take for granted that large-scale connectivity in the network is established.

In this thesis, we wanted to reconcile the two worlds of telecommunications-oriented and geometry-oriented approaches by introducing new models for D2D networks in urban areas where both refined connectivity conditions arising from physical reality and some topology of the support of the network are considered. In particular, the connectivity mechanism of our model considers the possibility of *shadowing* and of eventual *reflection* and *scattering* of radio signals, as we shall see in the next section. Finally, we especially wanted to address the question of *connectivity* of the modelled network by considering it under a percolation approach.

3.3.2 The model

We are now ready to introduce our model. We propose to see it as a *superposition* of four more elementary bricks, answering to the following four questions:

1. How is the network supported?
2. How are users of the network spread over the network?
3. Are there additional relays for extending D2D coverage or ensuring connectivity between adjacent streets in case of shadowing?
4. How are the nodes of the network interconnected?

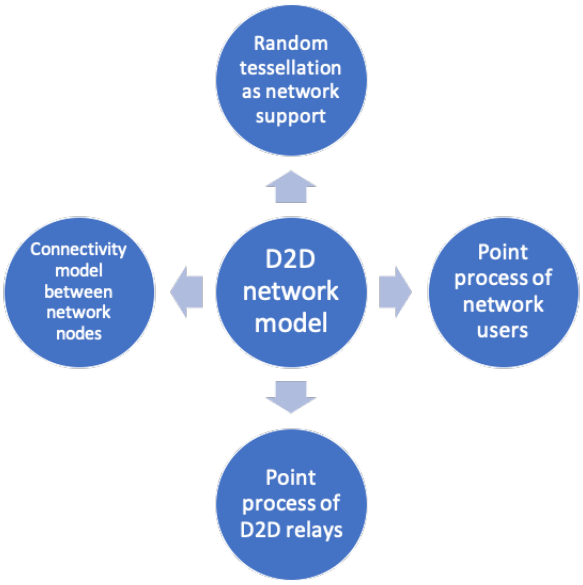


Figure 3.4 – Modelling approach. A random tessellation will serve as street system model and will prescribe a support for the network. Point processes for network users and relays, supported by the previous tessellation, will model how users and eventual D2D relays are spread over the network. Finally, an appropriate connectivity model describes the possible connections between the nodes of the network.

Each of these four questions is answered by an elementary brick of our modelling, see Figure 3.4. In more detail and given some probability space $(\Omega, \mathcal{A}, \mathbb{P})$ on which the considered random objects will be defined, our modelling is as follows.

First, the network will be supported by an urban *street system* given by a Poisson Voronoi tessellation (PVT) S generated by a planar homogeneous Poisson point process X_S with finite positive intensity $\lambda_S > 0$. We will denote by $L_A := 2\sqrt{\lambda_S}$ the length intensity of S, i.e. the mean edge length in a 1-area observation window. Seeing this PVT street system as a planar graph embedded in the plane $\mathbb{R}^2$ (as has been explained in Section 1.3), we respectively denote by E and V the sets of edges and vertices of S. Interpreting the PVT S as a modelling of a street system, an edge $e \in E$ is thought of as a *street* and a vertex $v \in V$ is thought of as a *crossroad* being the intersection of several streets.

We denote by $\Lambda(dx) := \nu_1(S \cap dx)$ the random stationary measure given by the restriction of the one-dimensional Hausdorff measure ν_1 to the PVT S. In other words, for any edge $e \in E$, $\Lambda(e) = \nu_1(e)$ denotes the length of e and, more generally, for any given observation window $W \subset \mathbb{R}^2$, $\Lambda(W)$ denotes the total edge length of S contained in W. Network users, equipped with mobile devices (which could be cellphones, tablets, smartwatches ...) will be modelled by a *Cox process* X^λ supported by E and driven by the intensity measure $\lambda\Lambda(dx) := \lambda\nu_1(S \cap dx)$, where $\lambda > 0$ is the *user linear intensity*, i.e. the mean number of network users per unit of edge length of S.

Since the conditional distribution of X^λ given S is the one of a Poisson point process supported by the edges of S, the probability of finding a Cox point at any given crossroad $v \in V$ is exactly 0. Therefore, an additional model for network nodes located at these crossroads has to be considered. Such nodes located at crossroads will play a critical role by relaying radio signals and ensuring connectivity between adjacent streets, which is why we call the network nodes at crossroads *relays* or *fixed relays*. These relays, supported by V, will be modelled by a *doubly stochastic Bernoulli point process* Y with parameter $p \in [0, 1]$. Equivalently, this means that a relay is present at a crossroad $v \in V$ with probability p, independently from all other crossroads $V \ni v' \neq v$. On a more applied approach, we can interpret p as the *proportion of crossroads equipped with a relay*. Note that such relays can be of various natures, for instance:

- A network user serving as D2D relay
- A fixed antenna
- A base station
- A connected device serving as D2D relay

We assume that the two point processes of network users X^λ and of relays Y are conditionally independent given the street system S, and we denote by $Z := X^\lambda \cup Y$ the superposition of the former two point processes. The points of Z thus represent the nodes of the modelled D2D network. Now, we need to specify how D2D connections between network nodes are modelled. We will consider two possible types of D2D connections

between two network nodes $Z_i \neq Z_j$.

On the one hand, we consider the *line-of-sight* (LOS) connections. Recall, as explained in Section 2.1.3, that such connections are possible when there is no physical obstacle preventing direct transmission between a transmitter and a receiver. In our model, we will consider that two network nodes are in LOS whenever they belong to the *same street*, i.e. the *same edge* of S. We denote by $r \in \mathbb{R}_+$ the D2D range in LOS, so that two points $Z_i \neq Z_j$ in LOS are connected whenever their mutual Euclidean distance $\|Z_i - Z_j\|_2$ is not larger than r. In other words, Z_i and Z_j are connected by a LOS D2D connection if and only if:

$$\begin{cases} \exists e \in E \, : Z_i \in e \, \text{and} \, Z_j \in e \\ \|Z_i - Z_j\|_2 \leq r \end{cases} \tag{3.3.1}$$

On the other hand, we consider connections due to *scattering* or *reflection* of radio signals by a physical obstacle (e.g. a building) located at the crossroad of two adjacent streets, that is to say streets that are incident to a common crossroad. Such connections make it possible to connect two network users (i.e. points of $X^\lambda = Z \setminus Y$) that are located on two adjacent streets if the total distance run by the signal is not too large. Denote by $r' \in \mathbb{R}_+$ the threshold for the distance run by the scattered signal and label e_{Z_i} the unique edge[7] on a which network user $Z_i \in X^\lambda$ is located on. Then two network users $Z_i \neq Z_j$ are connected by a D2D connection due to scattering or reflection of the signal if and only if:

$$\begin{cases} Z_i \in X^\lambda, Z_j \in X^\lambda \\ e_{Z_i} \neq e_{Z_j} \\ \exists v \in V, : e_{Z_i} \cap e_{Z_j} =: \{v\} \\ \|Z_i - v\|_2 + \|Z_j - v\|_2 \leq r' \end{cases} \tag{3.3.2}$$

In other words, D2D connections due to scattering or reflection of the signal can connect any two network users $Z_i, Z_j \in X^\lambda$ that are located on adjacent streets when the total distance run by the signal from Z_i to Z_j (or the other way around) along the street system is not larger than r'.

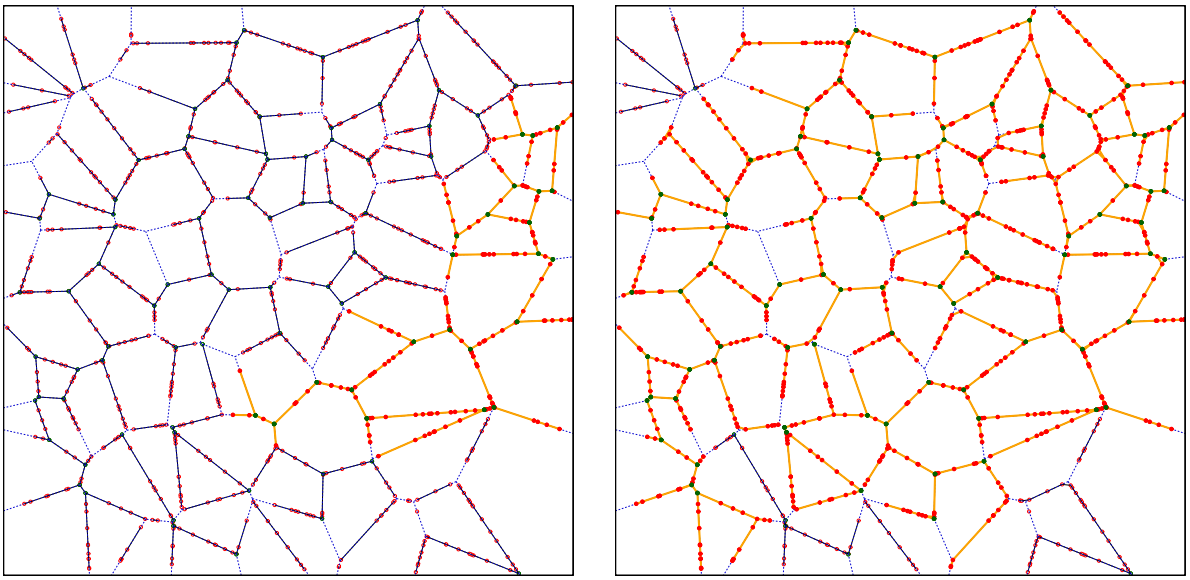
Finally, the network is modelled by its (random) *connectivity graph* $\mathcal{G} := \mathcal{G}_{p,\lambda,r,r'}$ with vertices being the points of Z, i.e. the network nodes, and where an undirected edge $Z_i \rightsquigarrow Z_j$ is drawn between two network nodes $Z_i \neq Z_j$ if and only if they are connected by a D2D LOS connection or by a D2D connection due to scattering or reflection, that is to say the pair Z_i, Z_j satisfies either (3.3.1) or (3.3.2).
Interpreting the edges connectivity graph as the possible D2D communications between nodes of the network, two network nodes Z_i and Z_j belonging to the same connected component of $\mathcal{G}$ can thus communicate by a multihop D2D communication in a finite number of hops. The network is considered to have good connectivity if it allows for connections at large-scale, that is to say between nodes separated by an arbitrarily large distance. In mathematical terms, good connectivity of the network is thus interpreted as

[7] Since the network users are given by the Cox process X^λ, which is simple and supported by the edges of S, such an edge is necessarily almost surely unique.

percolation of its (random) connectivity graph, in other words:

$$\mathbb{P}(\mathcal{G} \text{ has an unbounded connected component}) > 0$$

Examples of simulated networks are illustrated by Figure 3.5. A case of particular interest, studied in Chapter 4 and illustrated by Figure 3.5a is the case where $r' = 0$, i.e. only LOS communications along the edges of S are possible. Such a situation is often referred to as *canyon shadowing* in the literature [71, 72]. Canyon shadowing corresponds to a physical situation where the boundaries of the physical obstacles of the environment do not allow for penetration of the radio waves through them. From a mathematical perspective, we model this by only allowing for communications between a transmitter and a receiver that are in LOS. As mentioned earlier, and even more particularly in a canyon shadowing situation, the presence of additional Bernoulli relays located at crossroads is absolutely essential: they will allow to ensure connectivity between adjacent streets. Finally, Figure 3.5b shows a simulated network in the general case where both $r > 0$ and $r' > 0$.



(a) An example of a canyon shawdowing situation, where $r' = 0$. Connections between network nodes are only possible along the edges of the PVT street system S.

(b) General case where $r > 0$ and $r' > 0$. The positivity of r' allows for supplementary connections between adjacent edges of the PVT street system S.

Figure 3.5 – Examples of simulated networks. The blue dashed lines represent the PVT street system S, the red points represent network users (Cox process X^λ), the green points represent relays (Bernoulli process Y). The largest connected component in the simulation window is emphasised as follows: points of X^λ and Y are filled with their respective colours and connections are highlighted in orange. Edges of the smaller connected components are highlighted in black. In both cases, the realisations of S, X^λ and Y are the same.

3.3.3 Main system assumptions

We now present the crucial system assumptions made in the proposed model.

The model is two-dimensional

For the sake of simplicity, the problem is considered to be two-dimensional. This assumption can mostly be justified by the fact that a D2D communication is more stable when the path run by the radio wave is as direct as possible and does not involve a lot of diffractions and reflections in the third dimension of space.

The area covered by the network is morphologically homogeneous

In our model, we consider that the street system of the area covered by the network is morphologically homogeneous and can hence be modelled by a single random tessellation. Based on earlier results and fitting procedures (e.g. see [57, 58]) showing that PVT street systems provide the best fits for street systems in urban areas, we chose to model the street system in our model by a PVT.

Signal transmission through physical obstacles is neglected

We assume that two network nodes located on two different streets will be separated by at least one, if not many, physical obstacles. Moreover, we assume that any physical obstacle encountered will be sufficiently absorbing at the considered radio frequencies and cannot transmit any signal. As a matter of fact, only LOS communications or communications due to scattering and/or reflections at the crossroads of two adjacent streets are possible. Communications in other directions, for instance between two different streets that are not adjacent, are not possible, due to absorption of the radio waves by physical obstacles present in the environment of the network.

Scattering and reflection at crossroads comes with a loss of power greater than path loss for LOS communications

As has been noted in signal propagation models and by empirical measurements [126, Chapter 7], scattering or reflection of the signal by a physical obstacle located at the crossroad of two adjacent streets induces a loss of power in the received signal. In our model, we assume that this loss is considerably larger than the one due to path loss in LOS communications. As a matter of fact, the connectivity threshold r' for connections due to scattering is smaller than the one for LOS connections: $r' \leq r$. In practice, we may even expect $r' \ll r$, typically $r' = r/10$.

Constant communication radii

For the sake of simplicity, user mobility is neglected, so that our model can be seen as a static picture of the real network at a given time instant.
Regarding propagation effects, we assume that the only fading effect at stake is shadowing, which is already taken into account in our model. We thus do not need to consider fading coefficients for the wireless links. Moreover, for the sake of simplicity, we assume the following:

- Interference is neglected
- Noise power is a global deterministic constant $N > 0$

- The powers respectively emitted by users' mobile devices and network relays are equal to a common global, deterministic constant $P > 0$

As has been noted in [56], all of the previous assumptions imply that we can consider the communication radii r and r' to be global deterministic constants, which makes the model more tractable on an analytical perspective.

Assumptions on the D2D mode used

We assume that the D2D scenario used for D2D communications in our model is outband autonomous D2D. In particular, this means that BSs are not essential for establishing D2D communications in the network.

Chapter 4

The case of line-of-sight propagation only: canyon shadowing

The only thing worse than being blind is having sight but no vision.

Helen Keller

In this chapter, we study the case where $r' = 0$ in the model presented in Section 3.3. In other words, only line-of-sight (LOS) communications *along the edges* of the random tessellation S supporting the network are possible. Such a situation is referred to as *canyon shadowing*.

As explained earlier, the chosen Cox process modelling for network users makes the presence of Bernoulli relays at vertices of S essential for ensuring connectivity between adjacent edges of S. We may thus expect that some minimal proportion p^* of relays is a necessary condition for percolation of the network graph, no matter what. We shall see this is indeed the case. Furthermore, we have identified several *connectivity regimes* of very different natures for the percolation of our model, making its phase transition more elaborated than a simple dichotomy between a subcritical and a supercritical phase. More precisely, we exhibit a *relay-limited* regime where percolation of the connectivity graph can solely be ensured by the points of the Bernoulli process Y, even in the absence of Cox points (i.e. setting the user density $\lambda = 0$). Conversely, there also exists a *relay-and-user-limited* regime, where the users (Cox points) are absolutely essential for percolation of the connectivity graph. We prove the existence of such regimes with complete mathematical rigour.

On a more applied approach, resorting to numerical simulations conducted via original algorithms, we estimate the *frontiers* between the different connectivity regimes. Relating these numerical estimations to network parameters, we emphasise the importance of knowing those frontiers for *economic planning* of large-scale connectivity of the D2D network.

This chapter is based on the publication [87] and on the submission [89].

4.1 Introduction

4.1.1 Some other related works

The model introduced in Section 3.3 is of doubly stochastic nature and, by studying the percolation of the network connectivity graph, we will be faced with a percolation process living in a random environment. As mentioned earlier, (SINR) percolation of Cox processes, closely related to our work, has been considered only very recently [70, 128]. More generally, the study of percolation processes living in random environments had been considered a few years before the previous works and outlined that many standard techniques from Bernoulli or continuum percolation cannot be applied. As a matter of fact, new tools and techniques had to be introduced. In this regard, the paper from Bollobás and Riordan [25] on the threshold of Voronoi tiling percolation in the plane is pioneering. Later on, [4, 124] brought additional results concerning this model. Other percolation models [129], tessellations [27] and other random graphs [16, 18, 20] have also been considered. A more general study of Bernoulli and first-passage percolation on random tessellations has been conducted in [136, 137].

A repeatedly used technique for dealing with continuum percolation models (and more generally for percolation models featuring spatial dependencies) is the so-called *renormalization argument* or *coarse-graining argument*. The main idea is to discretize the original percolation process into a site percolation process on a discrete lattice (usually the Euclidean lattice $\mathbb{L}^d$), constructed in such a way that if the discrete process percolates, then so does the original one (or the converse, depending on whether one wants to prove percolation or the absence of percolation of the original model). The discrete percolation process is then stochastically dominated by an independent Bernoulli percolation process, seen as a product measure. This stochastic domination is most of the time obtained via resorting to a theorem from Liggett, Schonmann and Stacey [91]. More precisely, letting $d, k \in \mathbb{N}^*$ and $p \in (0,1)$, the authors consider a family $(X_s)_{s\in\mathbb{Z}^d}$ of $\{0,1\}$-valued random variables satisfying the following property:

$$\forall s \in \mathbb{Z}^d,\ \mathbb{P}(X_s = 1 \,|\, \sigma(\{X_u, u \in \mathbb{Z}^d : \|s-u\|_\infty \geq k\})) \geq p.$$

The class of random fields $(X_s)_{s\in\mathbb{Z}^d}$ satisfying the above property is denoted by $\mathscr{C}(d,k,p)$. The result is as follows:

Theorem 4.1.1 ([91, Theorem 0.0]). *For each d and k, when p is large enough, the random fields in $\mathscr{C}(d,k,p)$ are stochastically dominated from below by the product random field with density ρ, where ρ is a constant depending on d, k and p. One can make the density ρ of the product random field become arbitrarily close to 1 by taking p large enough.*

In particular, an important subclass of $\mathscr{C}(d,k,p)$ is the one of random fields which are k-dependent (as defined in Definition 1.4.11) for the supremum metric $\|\cdot\|_\infty$ and have marginals greater than or equal to p.

Regarding the canyon shadowing assumption, so-called *line-of-sight percolation* models for networks in environments with regular obstructions have already found some interest among the mathematical community. In [48], asymptotically tight results on k-connectivity of the connectivity graphs arising from such models are studied. Bollobás,

Janson and Riordan [24] extended these results by introducing a line-of-sight site percolation model on the discrete square lattice $\mathbb{Z}^2$ and the two-dimensional n-torus $[0 \times n]^2$. Asymptotical results for the critical probability were derived as well and interesting connections to Gilbert's continuum percolation model were also investigated.

However, the study of line-of-sight percolation in a continuum setting with a random environment has not, as far as we know, been studied yet.

4.1.2 Context

Network model

Recalling the notations introduced in Section 3.3, elementary bricks of the network model are as follows:

- A probability space $(\Omega, \mathcal{A}, \mathbb{P})$, on which the random bricks of our model are defined, is given.
- The street system supporting the network is still a random Poisson-Voronoi tessellation (PVT) S generated by a homogeneous Poisson point process X_S of intensity $\lambda_S > 0$ in $\mathbb{R}^2$. The line density is given by $L_A = 2\sqrt{\lambda_S}$ (see Table 1.1) and S is motion-invariant, i.e. its distribution is invariant under translations and rotations of $\mathbb{R}^2$.
- Network users are given by a Cox point process X^λ driven by the intensity measure $\Lambda(dx) = \lambda\nu_1(S \cap dx)$, where $\lambda > 0$.
- Relays are given by a (doubly stochastic) Bernoulli process Y on V, with parameter $p \in [0, 1]$.
- X^λ and Y are conditionally independent given their random support S, which we denote by $X^\lambda \perp\!\!\!\perp Y \,|\, S$. Note that since the random measure Λ is a deterministic function of the PVT S, it is equivalent to say that $X^\lambda \perp\!\!\!\perp Y \,|\, \Lambda$.

Some notation and terminology

The set of edges of S is denoted by E, and, furthering the identification of S with a street system, elements of E will be called *streets* from now onwards. Likewise, the set of vertices of S is denoted by V and elements of V are called *crossroads*. The length of a street $e \in E$ will be denoted $|e|$.
The chosen models for network users and relays imply that the number of users on a given street $e \in E$ is a Poisson random variable with mean $\lambda|e|$ and the number of users on two disjoint subsets of E are random variables being conditionally independent given the realisation of S. Moreover, given S, every crossroad $v \in V$ is occupied by a relay with probability p and the state of v (i.e. occupied or not) is independent of the state of any other crossroad $v' \neq v$.

Connectivity graph

Recall that we denote by $Z := X^\lambda \cup Y$ the superposition of the point processes of network users and relays. The points of Z will be called *network nodes*. With the canyon shadowing assumption, $r' = 0$ and so the network is modelled by the random connectivity graph $\mathcal{G}_{p,\lambda,r,0}$, where $r > 0$ designates the D2D range for LOS communications. Recall from Section 3.3 that the vertices of this graph are the network nodes and that an undirected edge $Z_i \leftrightsquigarrow Z_j$ between two network nodes Z_i and Z_j is drawn whenever they are in LOS and within range of communication. In other words, as formalised by (3.3.1), Z_i and Z_j belong to the same street $e \in E$ and we have $\|Z_i - Z_j\|_2 \leq r$. We are interested in the possibility of having arbitrarily long-range multihop D2D communications over the network. Hence, we interpret *percolation* of the connectivity graph $\mathcal{G}_{p,\lambda,r,0}$ as good connectivity of the network.

From now onwards, recalling that $r' = 0$ in the rest of this chapter, we will use the abuse of notation $\mathcal{G}_{p,\lambda,r}$ to designate the general network graph $\mathcal{G}_{p,\lambda,r,0}$ in the canyon-shadowing case. When the context is clear, we shall abbreviate $\mathcal{G}_{p,\lambda,r}$ by $\mathcal{G}$.

4.1.3 Dimensionless scale-invariant model parameters

Relevant model parameters are:

- The line density of the PVT street system $L_A > 0$, which we also refer to as *street intensity*. Note that larger values of L_A imply that the street system gets denser, while smaller values of L_A are representative of sparser streets systems. By definition, L_A is the average street length in any 1-area observation window (by stationarity of S, L_A does not depend on the chosen observation window). Hence, the dimension of L_A is the inverse of a length.

- The *users' linear intensity* $\lambda > 0$, which is the mean number of users per unit street length, so that the dimension of λ is the inverse of a length.

- The relay proportion $p \in [0, 1]$, representing the probability of having a relay installed at any crossroad $v \in V$. In frequentist terms, we can also interpret p as the proportion of crossroads of the street system that are equipped with a relay by the operator. p is dimensionless.

- The D2D communication range $r > 0$, having the dimension of a length.

From a mathematical point of view, note that, by motion-invariance of S, we can set $L_A = 1$ without loss of generality for the study of the percolation of the connectivity graph representing the network. From an engineering point of view, all the parameters (L_A, λ, p, r) are meaningful. From a numerical perspective, note that an infinite graph cannot be simulated, and so we will have to simulate the network in a given observation window. It is customary to consider squared simulation windows of the form $[0, \texttt{win}]^2$ (e.g. see [17, 106]), but the size `win` of the simulation window may come as an additional parameter in the study of the model. In our case, this can be avoided by noting that our model features *scale invariances*. Indeed, S, X^λ and Y are motion invariant. Furthermore, as has been noted in [29], changing L_A to aL_A for some scaling factor $a > 0$ is

equivalent to considering a rescaled network where the users' linear intensity λ has been replaced by $a\lambda$ and where the connectivity range r has been replaced by r/a. As a matter of fact, the two dimensionless parameters λ/L_A and rL_A are scale-invariant.

Actually, we may even go further by considering refined versions of the previous two scale-invariant parameters and relating them to the length of a *typical street*, or *typical edge* of the PVT S. Recall, as has been seen in Section 1.3.5, that such a typical edge is the typical point of the process of edges of S. Following [31, Section 9.4], the length of this typical edge of S is equal to L_A/λ_1, where λ_1 denotes the intensity of the process of edges of S, see Section 1.3.3. Since S is a planar PVT having been generated by a homogeneous Poisson point process X_S of intensity λ_S, we have $L_A = 2\sqrt{\lambda_S}$ and $\lambda_1 = 3\lambda_S$, see Table 1.1. Thus, $\lambda_S = L_A^2/4$ and so the length of the typical edge (or typical street) of S is simply given by:

$$\frac{L_A}{\lambda_1} = \frac{L_A}{3\lambda_S} = \frac{L_A}{3\frac{L_A^2}{4}} = \frac{4}{3L_A}.$$

It thus makes sense to consider the respective dimensionless and scale-invariant parameters U and H, defined as follows:

$$U := \frac{4}{3}\frac{\lambda}{L_A} \tag{4.1.1}$$

$$H := \frac{4}{3}\frac{1}{rL_A}. \tag{4.1.2}$$

The letter U stands for "users" while the letter H stands for "hops". Indeed, since $4/(3L_A)$ is the length of a street of S and $\lambda > 0$ represents the mean number of users per unit edge length, U is equal to the *mean number of users per typical street*. In the same way, since a single-hop D2D communication covers a length r of a street, $1/r$ can be interpreted as the "number" of hops covering a street of unit length and so H represents the *mean number of D2D hops* needed to ensure connectivity on a typical street of S. Note that H represents the interplay between the street system and the transmission range related to D2D technology: larger values of H mean that more hops will be needed to traverse streets while smaller values of H mean that fewer hops will be needed to traverse streets. In other words, larger values of H corresponds to cases where the streets are somehow long compared to the D2D range r, while smaller values of H corresponds to cases where the streets are somehow short compared to the D2D range.
Introducing the previous parameters allows us to study the model in terms of the parameters (p, U, H): this will be particularly convenient for numerical purposes, as this allows to perform simulations for a given value of L_A and deduce other results for different values of L_A by scale invariance, so that fixing a particular value of L_A in numerical experiments does not matter.

As a matter of fact, from now onwards and in the rest of this chapter, we shall denote the network connectivity graph in the canyon shadowing situation indifferently by $\mathcal{G}_{p,\lambda,r}$, $\mathcal{G}_{p,U,H}$ or simply $\mathcal{G}$ when the context is clear, so as to avoid heavy notation.

4.2 Phase transitions between different connectivity regimes

Using the dimensionless scale-invariant parameters introduced in Section 4.1.3, we consider the network connectvity graph $\mathcal{G} := \mathcal{G}_{p,U,H}$ parametrized by (p, U, H) and denote the *percolation probability* by:

$$P(p, U, H) := \mathbb{P}(\mathcal{G}_{p,U,H} \text{ has an unbounded connected component}). \tag{4.2.1}$$

As usual, we say that $\mathcal{G}$ percolates if $P(p, U, H) > 0$ and does not percolate otherwise. A first observation is the following monotonicity of the model: P is increasing in p and U and decreasing in H. Indeed, larger values of p and U respectively mean more relays and more users and the network, making more possible connections and percolation easier to occur. On the contrary, larger values of H mean that the number of hops needed to traverse a typical street is larger, making percolation more difficult to occur.

For given $p \geq 0, H \geq 0$ consider the following critical value of the mean number of users per street U:

$$U_c(p, H) := \inf\{U \geq 0 : P(p, U, H) > 0\}, \tag{4.2.2}$$

with $U_c(p, H) := \infty$ if $P(p, U, H) = 0$ for all $U \geq 0$. We aim at showing that there is a region (i.e. a connected subset) of parameters (p, H) such that $0 < U_c(p, H) < \infty$. Such a region is where the percolation of $\mathcal{G}$ exhibits non-trivial phase transition in the density of users. Before stating our results, we discuss special cases of the model.

4.2.1 Special cases of the model

PVT site percolation

For $H = 0$, or equivalently $r = \infty$ (recall that $L_A > 0$ and $H = 4/(3rL_A)$), $P(p, U, 0) =: P_{PVT}(p)$ does not depend on U and corresponds to the percolation probability of the *Bernoulli site percolation model on the planar PVT*[1]. Denote the critical parameter of this model by

$$p^* := \inf\{p \in [0, 1] : P_{PVT}(p) > 0\}. \tag{4.2.3}$$

Clearly, by the monotonicity of the model, $\mathcal{G}_{p,U,H}$ does not percolate for $p < p^*$, whatever $U \geq 0$, $H \geq 0$.

PVT hard-geometric bond percolation

For $U = 0$ (no mobile users) and $p = 1$ (a relay is present at each crossroad), $\mathcal{G}_{1,0,H}$ corresponds to a non-standard inhomogeneous bond percolation model on the PVT, in which the edges of the PVT are open or closed depending whether their length is smaller or larger than the threshold $r = 4/(3HL_A)$. We call this model *PVT hard-geometric bond*

[1]The PVT site percolation model should not be confused with the *Voronoi tiling percolation* model, which consists in coloring each *cell* of a PVT in black independently from all other cells with some fixed probability p and investigating the random tiling of black cells. The critical probability for this latter model in the plane has been proven to be $1/2$ in [25].

percolation. It seems that this model has not been studied in the literature. We define the critical bond parameter of this model by:

$$H_c \coloneqq \sup\{H \geq 0 : P(1, 0, H) > 0\}. \tag{4.2.4}$$

PVT soft-geometric bond percolation

Keeping $p = 1$ and $H > 0$, considering $U > 0$ introduces to our model the possibility of opening some long edges, which are not open in the PVT hard-geometric bond percolation. Note that this is equivalent to yet another bond percolation, in which the edges of the PVT are open independently with probabilities depending on their lengths. We call it *PVT soft-geometric bond percolation*. It seems that such a model has not been studied in the literature either.

$\mathcal{G}$ as a superposition of three percolation models

Using the aforementioned particular cases of our model, the general connectivity graph $\mathcal{G}_{p,U,H}$ for non-trivial values of the parameters (p, U, H) can be seen as a superposition of the three previous percolation models: site model, hard-geometric bond model and soft-geometric bond model.

4.2.2 Analytical results

We now state our analytical results, which ensure the non-triviality of the previously defined critical thresholds p^*, H_c and $U_c(p, H)$.
First, the existence of a region where the percolation of the network connectivity graph $\mathcal{G}_{p,U,H}$ has a non-trivial phase transition in U is ensured by the following two results:

Theorem 4.2.1 (Existence of subcritical intensities of users)**.** *For large enough $H \in [0, \infty)$ and small enough $U > 0$ we have $P(1, U, H) = 0$ and, consequently, by monotonicity of the model, $P(p, U, H) = 0$ for any $p \in [0, 1]$.*

Theorem 4.2.2 (Existence of supercritical intensities of users)**.** *For large enough $p \in (0, 1)$ we have $P(p, U, H) > 0$ for any $H \in [0, \infty)$ and large enough $U < \infty$ (depending on H).*

A direct consequence of Theorem 4.2.2 and of standard percolation arguments is the non-triviality of the PVT site percolation threshold:

Corollary 4.2.2.1. *The critical parameter for PVT site percolation p^* defined in (4.2.3) is non-trivial, i.e. $p^* \in (0, 1)$.*

Furthermore, we obtain the existence of a regime where $\mathcal{G}$ percolates, whatever small the density of users. In other words, there exists a region of parameters (p, H) for which $U_c(p, H) = 0$. We call such a region the *permanently supercritical range* and its existence is ensured by the following result:

Theorem 4.2.3 (Existence of the permanently supercritical range)**.** *For large enough $p < 1$ and small enough $H > 0$, we have that $P(p, 0, H) > 0$.*

As a corollary of Theorems 4.2.1 and 4.2.3, we obtain the non-triviality of the critical threshold for PVT hard-geometric bond percolation:

Corollary 4.2.3.1. *The critical parameter for PVT hard-geometric bond percolation H_c defined by (4.2.4) is non-trivial, i.e. $0 < H_c < \infty$.*

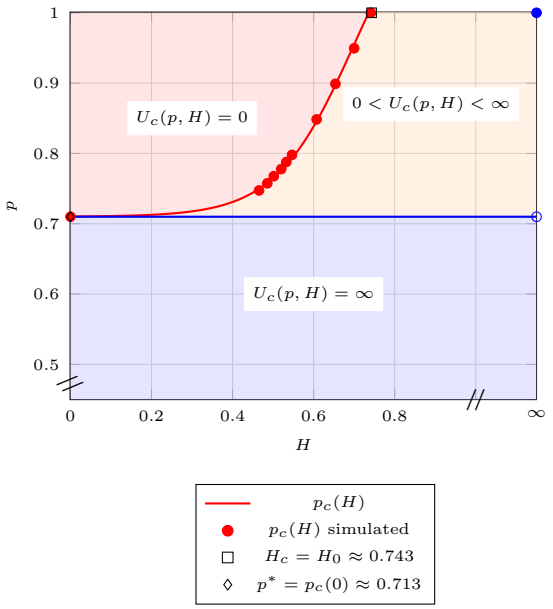
The proofs of all the above results are postponed to Section 4.4.

4.2.3 Connectivity regimes

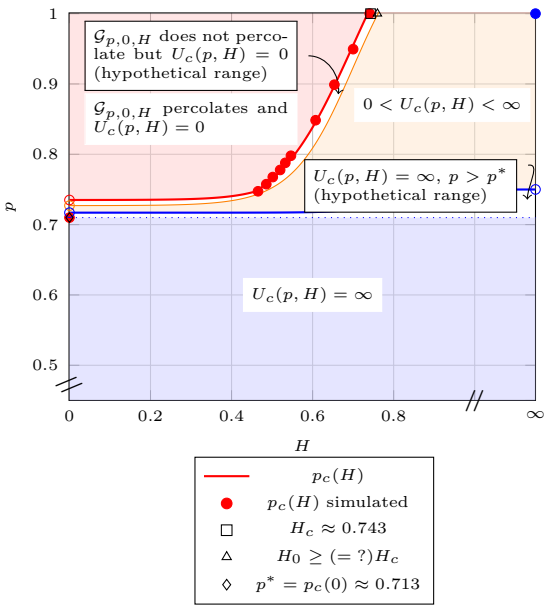
In light of the previous analytical results, there are three and may be up to five different ranges of parameters (p, H) of interest in our model, giving rise to different *connectivity regimes* of the network.

The range of parameters (p, H) where $0 < U_c(p, H) < \infty$, schematically presented in orange on Figure 4.1a, can be seen as the *critical range of* (p, H) in the sense that it separates the following two ranges of (p, H):

- The *permanently subcritical range* (schematically presented in blue on Figure 4.1a), where $\mathcal{G}$ does not percolate whatever large the density of users ($U_c(p, H) = \infty$).
- The *permanently supercritical range* (schematically presented in red on Figure 4.1a), where $\mathcal{G}$ percolates with positive probability, whatever small the density of users ($U_c(p, H) = 0$).



(a) Conjectured phase transition diagram of $\mathcal{G}$. The existence of the orange region where $0 < U_c(p, H) < \infty$ follows from Theorems 4.2.1 and 4.2.2.



(b) Phase transition diagram of $\mathcal{G}$ with hypothetical ranges of (p, H).

Figure 4.1 – Possible phase transition diagrams of the network connectivity graph $\mathcal{G}$.

From a theoretical perspective, we cannot exclude that the permanently supercritical range contains a non-empty subset of (p, H) such that $\mathcal{G}$ does not percolate without users $(U = 0)$ but percolates with positive probability for arbitrarily small densities of users, as depicted on Figure 4.1b. Moreover, we do not know whether the permanently subcritical range contains some $p > p^*$, as also depicted on Figure 4.1b. Note that we do not know the exact shapes of the curves separating these ranges except that they are monotonic. Even continuity is not known.

By monotonicity of the model, $\mathcal{G}$ does not percolate for all $p < p^*$, whatever $U \geq 0$ and $H \geq 0$. Thus, the permanently subcritical range of parameters is delimited by the the PVT site percolation threshold p^* given by the blue horizontal line on Figure 4.1a. In other words, large-scale connectivity of the network is delimited by a minimal proportion of crossroads p^* equipped with relays. Below that proportion, large-scale communications are not possible, regardless of all other network parameters.
For (p, H) in the (strictly) permanently supercritical range, $\mathcal{G}$ percolates whatever small the density of users. This means that long-range communications over the network are possible whatever small the number of D2D users in the network. As a matter of fact, large-scale connectivity of the network can solely be ensured by the relays: this is what we call *relay-limited connectivity*. In such a situation, given the interplay between the geometry of the streets and the D2D technology represented by H, the operator can ensure good connectivity of the network if it accepts to deploy a sufficiently high amount of relays over the crossroads of the street system.
Finally, in the critical range, where $0 < U_c(p, H) < \infty$, the percolation of $\mathcal{G}$ exhibits a non-trivial phase transition in the density of users. This means that large-scale connectivity of the network can be ensured by a finite positive density of users supplementing the relays. In other words, both relays and users are essential to allow for long-range communications: this is what we call *relay-and-user-limited connectivity*.

Still examining the frontiers between the different ranges of parameters, let us also define the following quantities:

$$H_0 \coloneqq \sup\{H \geq 0 : U_c(1, H) = 0\} \tag{4.2.5}$$

$$\forall\, H < H_c,\ p_c(H) \coloneqq \inf\{p > 0 \ :\ P(p, 0, H) > 0\}. \tag{4.2.6}$$

The first quantity is well-defined. Indeed, it is easy to see that $H_0 \geq H_c$, and, by Corollary 4.2.3.1, $H_c > 0$. Moreover, by Theorem 4.2.1, $H_0 < \infty$. In the definition of H_0, the condition $U_c(1, H) = 0$ means that we are looking at all these values of H for which percolation of the network with all relays deployed $(p = 1)$ remains possible whatever small the density of users $(U_c = 0)$. We conjecture that $H_c = H_0$, as illustrated by Figure 4.1a. However, if the hypothetical range where $\mathcal{G}$ does not percolate without users but percolates for arbitrarily small densities of users exists, then we would have $H_0 > H_c$, see Figure 4.1b.

Regarding the well-definedness of $p_c(H)$, recall that $0 < H_c < \infty$ by Corollary 4.2.3.1. For $H < H_c$, necessarily $P(1, 0, H) > 0$ by definition of H_c and monotonicity of the model. Hence the set $\{p > 0 \ :\ P(p, 0, H)\}$ is non-empty and lower-bounded by 0, so the infimum $p_c(H)$ is well-defined. The function $H \mapsto p_c(H)$ represents the minimal proportion of

relays p needed to ensure percolation of the network in the relay-limited regime. In other words, this is the boundary of the permanently supercritical range for (p, H), given by the red curve on Figures 4.1a and 4.1b.

Numerical results with estimations of the thresholds p^*, H_c as well as the functions $H \mapsto p_c(H)$ and $(p, H) \mapsto U_c(p, H)$ are detailed in Section 4.2.4.

4.2.4 Numerical results: estimating the frontiers between the different connectivity regimes

We performed numerical simulations to estimate the frontiers between the different connectivity regimes presented in Figure 4.1 and in the previous section. All of our simulations have been performed using the statistical software `R`. We now present the results of these numerical simulations and their consequences in terms of network connectivity. Details on the methodology of the simulations and the algorithms used will be provided in Section 4.3.

Estimation of the PVT site percolation threshold p^*

We obtained the following estimate for the PVT site percolation threshold p^*:

$$p^* \approx 0.713 \tag{4.2.7}$$

Our value only slightly differs from the most recent estimation available in the literature [19]: $p^* \approx 0.71410 \pm 0.00002$. While the authors providing this estimate proceeded with Monte-Carlo simulations with periodic boundary conditions and investigated the growth of the largest cluster, we chose a window-crossing method to obtain our estimate of p^*. More details on our window-crossing method will be provided in Section 4.3. Note that p^* is the upper limit of the permanently subcritical range for the percolation of $\mathcal{G}$. In terms of network design considerations, the estimate $p^* \approx 0,713$ means that the D2D network cannot allow for arbitrarily long-range multihop D2D communications if less than 71.3% of crossroads are equipped with a relay, whatever the values of the user density and of the D2D range. Such a proportion of 71.3% is considerably high and is due to the fact that with the canyon shadowing assumption, only LOS communications are possible. We shall however see in Chapter 5 that much less relays are actually needed at large scale.

Estimation of the PVT hard-geometric bond percolation threshold H_c

We found the following estimate for H_c:

$$H_c \approx 0.743, \tag{4.2.8}$$

meaning that the critical threshold for PVT hard-geometric bond percolation corresponds to the case where the length of the typical edge of the PVT is equal to 74.3% of the D2D range. Note that H_c is the maximal compromise that can be made on H in the permanently supercritical range of (p, H), see Figures 4.1a and 4.1b. Within this range, $U_C(p, H) = 0$ and so $\mathcal{G}$ percolates whatever small the density of users. On a more applied perspective, this means that when the typical street is not longer than 74.3% of the

D2D range, the connectivity of the network is relay-limited and long-range multihop D2D communications are possible in the absence of users.

Estimations of $p_c(H)$

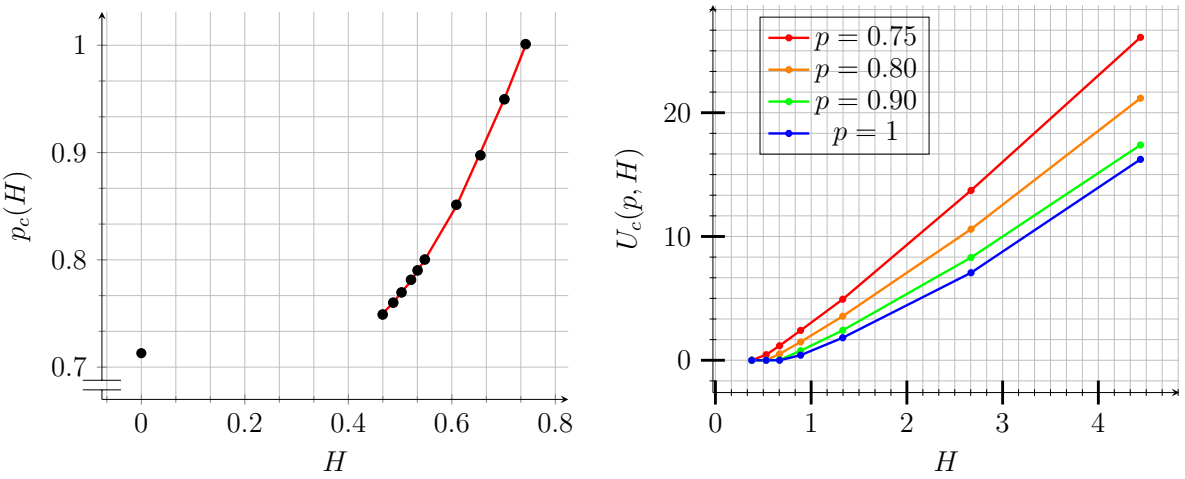
Recall that $p_c(H)$ is the boundary of the permanently supercritical range for the percolation of $\mathcal{G}$. In engineering terms, this corresponds to the relay-limited regime, where large-scale connectivity of the D2D network can be solely ensured by relays, even in the absence of users. Thus, $p_c(H)$ represents the minimal proportion of crossroads of the street system that have to be equipped with a relay to allow for long-range multihop D2D communications in the network. Since installing these relays comes with a cost, the mapping $H \mapsto p_c(H)$ can be seen as the investment in relays that the operator must make, given the interplay H between the geometry of the street system and the D2D range.

Table 4.1 – Estimated values of $p_c(H)$.

H	$p_c(H)$
0.467	0.75
0.487	0.76
0.503	0.77
0.521	0.78
0.534	0.79
0.548	0.80
0.609	0.85
0.655	0.90
0.702	0.95
$H_c \approx 0.743$	1

Estimated values of $p_c(H)$ for several values of H are given in Table 4.1 and illustrated by the black points on Figure 4.2a. We were only able to consider values $H > 0.46$, as the percolation regime of $\mathcal{G}$ for smaller values of H started to have an erratic behaviour. This can be explained by two facts. On the one hand, the simulated graphs cannot be infinite and have to be simulated in bounded simulation windows, which can induce boundary effects. On the other hand, when H gets smaller, one approaches the permanently supercritical regime and it is thus much trickier to perform numerical simulations. However, we do know that $p_c(0) = p^*$: this corresponds to the isolated black point on Figure 4.2a. Finally, in order to get a continuous approximation of $p_c(H)$ for $0.46 < H < H_c$, we interpolated the estimated values and found out that a quadratic fit:

$$\begin{cases} p_c(H) = aH^2 + bH + c \\ a \approx 1.45, b \approx -0.84, c \approx 0.83 \end{cases} \tag{4.2.9}$$



(a) $p_c(H)$ as a function of H. The black points are the discrete values from Table 4.1 obtained by simulation and the red curve is the estimated quadratic fit. The isolated point $p_c(0) =: p^*$ is the estimate of the PVT site percolation given by (4.2.7).

(b) Plotting of $U_c(p, H)$ as a function of H for certain values of p. Red curve: $p = 0.75$. Orange curve: $p = 0.80$. Green curve: $p = 0.90$. Blue curve: $p = 1$.

Figure 4.2 – Plotting of $p_c(H)$ as a function of H (left) and of $U_c(p, H)$ as a function of H for certain values of p (right).

is able to explain 99% of the variance. The values of the coefficients a, b and c are estimated by a generalised linear regression. This quadratic fit is illustrated by the red curves on Figure 4.2a and on Figure 4.1. On this latter figure, the continuous fit has been extended to $p_c(0) = p^*$ by smoothing the curve, for clarity of the phase transition diagram.

Estimations of $U_c(p, H)$

Figure 4.2b shows estimations of the critical user density $U_c(p, H)$ for several values of the parameters (p, H) mainly in the critical range where, recall, $U_c(p, H) \in (0, \infty)$. Exact values are given in Table 4.2. In this table, a comparison with the model studied in [29] has been conducted in the column entitled NoSha. This stands for "No Shadowing", as the model previously studied in [29] is the Gilbert graph of the same Cox process X^λ: any two Cox points are connected whenever they are separated by less than the connectivity threshold $r = 4/(3HL_A)$. Relays are thus not needed and shadowing is not taken into account since connections are possible in all directions.
We see that the mapping $(p, H) \mapsto U_c(p, H)$ is decreasing in p and increasing in H, which was expected. Indeed, for given H, larger values of p imply more relays in the network and hence less users are needed for percolation. Likewise, for given p, larger values of H mean that more hops are needed to traverse streets and hence more users are needed.
Interestingly, compared to the previous estimates provided in [29], taking the canyon shadowing into account does not necessarily mean that more users are needed for percolation of $\mathcal{G}$ and thus large-scale connectivity of the D2D network. The relay proportion p also plays its own role and there is a compromise to make between the relay propor-

Table 4.2 – Estimated values of the critical user density $U_c(p, H)$.

$U_c(p, H)$					
H	$p = 1$	$p = 0.9$	$p = 0.8$	$p = 0.75$	NoSha [29]
4.44	16.23	17.39	21.17	26.09	15.87
2.67	7.07	8.30	10.59	13.72	7.44
1.33	1.82	2.42	3.56	4.93	–
0.89	0.41	0.77	1.48	2.41	1
0.67	0	0.03	0.51	1.17	–
0.53	0	0	0	0.45	0.32
0.38	0	0	0	0	0.16

tion p and the user density U. For an operator, equipping crossroads with relays comes with considerable costs, while increasing the user density U by gaining new customers is feasible, but hard to predict and master.

4.3 Methodology of the numerical simulations

We now turn on to presenting the methodology used in the numerical simulations performed to obtain the above estimations of critical network parameters. All of our simulations have been performed using the statistical software `R`.

Each simulation of a realisation of the network connectivity graph $\mathcal{G}$ corresponds to the following process:

1. For given network parameters, simulate the PVT street system S, the process of network users X^λ and the process of relays Y.
2. Use appropriate algorithms to generate the connected components of the network connectivity graph $\mathcal{G}$.
3. Assess whether the simulated connectivity graph percolates or not.

Once we know whether a simulated graph percolates or not, appropriate statistical methods are used to estimate critical network parameters such as p^*, H_c or $U_c(p, H)$.

4.3.1 Simulating the network

First and foremost, note that infinite networks cannot be simulated and so we have to choose a simulation window in order to perform our simulations. Such a simulation window is chosen to be a square of side `win`, expressed in kilometres (km). Most of

the time and when possible (especially in terms of running time of the simulations), `win` $= 30\,\mathrm{km}$, a value chosen sufficiently large so as to avoid any boundary effects due to the finiteness of the simulation window.
For practical reasons, we perform our simulations using the original network parameters (L_A, p, λ, r) and then compute the associated dimensionless parameters U and H. Given (L_A, p, λ, r), one first needs to simulate the elementary bricks of the network, i.e. the PVT street system S, the Cox process of network users X^λ and the Bernoulli process of relays Y. This is easily done by using the packages[2] `spatstat` [14] and `deldir` [127] from `R`. In more detail, we proceeed as follows:

1. Generate a planar homogeneous Poisson point process X_S with intensity $\lambda_S = L_A^2/4$.

2. Generate the Voronoi tessellation S associated with X_S. The associated line density is indeed equal to $2\sqrt{\lambda_S} = L_A$.

3. Generate the Cox process X^λ.

4. Generate the Bernoulli process Y.

Generating the Poisson point process X_S is straightforwardly done using the command `rpoispp` from `spatstat`. In the same way, the command `deldir` from the package with the same name generates the associated PVT and PDT in a list `L`, from which the PVT is easily accessed. The Cox process X^λ is also easy to simulate with the command `rpoisppOnLines`, taking as arguments the users' linear intensity λ and a line segment pattern (type `psp` in `R`) which will support the simulated Cox process. Converting the edges of S into such a line segment pattern is also straightforward, using the command `psp`. Generating the Bernoulli process Y requires a bit more work, as no direct command is available in `spatstat`. We generate a realisation of Y by applying Algorithm 4.1.

Algorithm 4.1: Simulating a realisation of a doubly stochastic Bernoulli process Y with parameter p, supported by the crossroads of a PVT street system S.

Input: Realisation of a PVT S, $p \in [0,1]$
Output: Realisation of Y
$V \leftarrow$ list of crossroads of S. v_i denotes the i-th crossroad.
$n \leftarrow \#(V)$
$Y \leftarrow \emptyset$
Generate a vector $U = (U_1, \ldots, U_n)$ of n i.i.d. random variables $U_i \sim \mathcal{U}([0,1])$
$x \leftarrow \{i \,:\, U_i \leq p\}$
for $i \in x$ **do**
  $Y \leftarrow Y \cup \{v_i\}$
end

[2] On a more general perspective, `spatstat` is a very complete package overwhelmingly used for spatial statistics. It is regularly updated and documentation of the package is now 1500 pages long.

4.3.2 Computing the connections of $\mathcal{G}$ and generating the connected components

While generating the support of the network, it is fairly easy to assign a unique street number denoted `street_label` (an integer number between 1 and the number of streets $\#(E)$) to each street of the PVT street system S. Then, when we generate the nodes of the network, we assign an additional attribute `node_street_label` (a list of integers) to each node of the network: this vector `node_street_label` contains the numbers of the streets on which this particular node is located . For instance, a network node with `node_street_label` $= \{4, 9, 12\}$ is located at the crossroad of streets numbered 4, 9 and 12. Given the structure of the network, note that we are in either one of the following cases:

- Either a node is a network user (Cox point), in which case it is located on a unique edge of S and so `node_street_label` contains only one element.
- Or the node is a relay (Bernoulli point), in which case it is located on a crossroad of S. Since any crossroad of S has a degree almost surely equal to 3[3], `node_street_label` almost surely contains 3 elements.

Obviously, two nodes of the network (Cox or Bernoulli points) are in LOS if and only if their respective `node_street_label` share a common element.

Once the network is generated, given the D2D range r, we need to compute possible connections in the network. We use Algorithm 4.2. In more detail, we proceed street by street. For a given street $e \in E$, we obtain the users and relays located on that specific street. In practice, they are found by checking whether the attributes `node_street_label` contain the `street_label` of the considered street $e \in E$. Once we have obtained these users and relays, we collapse them into a simple list `nodes` containing the coordinates of all the network nodes located on e, and sort it by ascending x-coordinate. By doing so, we now have all the network nodes located on e in consecutive order of appearance from one endpoint of e to the other. If the length $|e|$ of e is smaller than the D2D range r, all successive gaps will be smaller than r. Hence, the first point will be connected to the second, the second to the third, and so on. If the length of e is larger than the D2D range r, we compute the successive distances d_i between consecutive nodes and all d_i smaller than r correspond to a connection in the network graph $\mathcal{G}$.
Note that the previous algorithm does not compute *all* connections of $\mathcal{G}$ exhaustively. Indeed, for a given street $e \in E$, we only compute the connections between successive network nodes along the street and ignore the connections between any two network nodes that are not consecutive on the street. This does not matter, as we will then be interested in the percolation of $\mathcal{G}$ and hence in the connected components of $\mathcal{G}$. In particular, as long as two network nodes are in the same connected component of $\mathcal{G}$, it won't matter whether they are connected in 1 or more hops when studying percolation of $\mathcal{G}$.

To generate and keep track of the connected components of $\mathcal{G}$, we use a *union-find algorithm* [50], as has been suggested in [106]. The principle of this algorithm is the following one:

[3]The degree of any vertex in a planar PVT is almost surely equal to 3, see [108].

Algorithm 4.2: Computing the D2D connections between the nodes of the network

Input: Realisations of the PVT street system S, the processes of network users X^λ and relays Y, the D2D range r

Output: A set `D` of pairs of network nodes representing (some of the) direct connections in $\mathcal{G}$. Each element of `D` is thus a pair (Z_i, Z_j) indicating that the two network nodes Z_i and Z_j are connected by an edge connectivity graph $\mathcal{G}$, i.e. $Z_i \leftrightsquigarrow Z_j$.

$E \leftarrow$ list of streets of S
`D` $\leftarrow \emptyset$
for $e \in E$ **do**
  `nodes` $\leftarrow \left(X^\lambda \cap e\right) \cup (Y \cap e)$
  `nodes` $\leftarrow$ Order `nodes` by ascending x-coordinate $=: \{Z'_1, \ldots, Z'_k\}$
  $k \leftarrow \#$(`nodes`)
  if length of $e =: |e| \leq r$ **then**
    $D \leftarrow D \cup \{(Z'_1, Z'_2), (Z'_2, Z'_3), \ldots, (Z'_{k-1}, Z'_k)\}$
  else
    for $i \in \{1, 2, \ldots, k-1\}$ **do**
      $d_i \leftarrow \|Z'_i - Z'_{i+1}\|_2$
      if $d_i \leq r$ **then**
        $D \leftarrow D \cup (Z'_i, Z'_{i+1})$
      end
    end
  end
  Proceed to next street
end

1. Each connected component of $\mathcal{G}$ is represented by one of its members called *root*. Within a connected component of $\mathcal{G}$, each network node points towards another node called its *parent*. The root representing a component is its own parent.

2. At the beginning, each network node is its own parent. This initialisation step is done by a function `Initialise`.

3. Using the output of Algorithm 4.2, we know when two network nodes are connected in $\mathcal{G}$. For each connection $Z_i \leftrightsquigarrow Z_j$ computed, we look at their roots using a function `Find`. If their roots are the same, they are already in the same connected component and we do nothing. If not, then we merge their connected components using a function `Union`.

The previous approach is given by Algorithm 4.3. In computer science, union-find algorithms are often implemented using forest and tree structures. This asserts the above denomination of *root* for the representative member of a connected component. More precisely, $\mathcal{G}$ is represented by a forest data structure of trees and each connected component of $\mathcal{G}$ is represented by a tree of the forest. Within a tree, i.e. within a connected component of $\mathcal{G}$, each network node points towards some parent node. The whole tree is

Algorithm 4.3: Union-find algorithm for keeping track of the connected components of $\mathcal{G}$.

Input: The `D` of connections computed by Algorithm 4.2.
Output: An identification, for each network node $Z_i \in Z$, of the connected component of $\mathcal{G}$ to which Z_i belongs.

```
for (x, y) ∈ D do
    root_1 ← Find(x)
    root_2 ← Find(y)
    if root_1 = root_2 then
        Do nothing
    else
        Union(x, y)
    end
end
```

then represented by its root, which points towards itself. In this perspective, applying the `Find` function to some network node Z_i consists in finding the root of the tree representing the connected component to which Z_i belongs. In the same way, if two nodes $Z_i \neq Z_j$ are connected in $\mathcal{G}$, the `Union` function checks whether they are already in the same tree and, if not, merges the corresponding trees together. Several implementations for the `Find` and `Union` functions have been proposed in the literature. Depending on the implementation chosen, the time complexity of the Union-Find Algorithm 4.3 can hugely vary. So far, one of the most efficient approaches available consists in the following:

- *Path compression* for the `Find` function. This means that while looking at the root of a given node of the network, all parent nodes visited in the tree of that node, up to the root, are also updated so as to directly point to the root, thus flattening the structure of the tree. Pseudo-code for this approach is given by Algorithm 4.5.

- *Weighted union* for the `Union` function. This means that each tree of the forest is given a certain *weight* $w \in \mathbb{N}$ and that, when merging two trees, the tree with smaller weight will be attached to the tree with larger weight. In other words, root of the tree with smaller weight will point towards the root of the tree with greater weight. Pseudo-code for this approach is given by Algorithm 4.6.

As mentioned above, before performing the union-find algorithm, an initialisation step is required. This initialisation consists in ensuring that each node of the network is its own parent at the beginning. Moreover, with the weighted union approach for the `Union` function, one also needs to initialise the weight of any node to 1. Pseudo-code for the `Initialise` function is given by Algorithm 4.4.

Algorithm 4.4: Pseudo-code for the `Initialise` function.

```
Function Initialise(x):
    x.parent ← x
    x.weight ← 1
```

Algorithm 4.5: Pseudo-code for the `Find` function with path compression.

```
Function Find(x):
    if x.parent ≠ x then
        x.parent ← Find(x.parent)
    return x.parent
```

Algorithm 4.6: Pseudo-code for the for the `Union` function with weighted union.

```
Function Union(x, y):
    root_x ← Find(x)
    root_y ← Find(y)
    if root_x = root_y then
        Do nothing
    else
        if (root_x).weight < (root_y).weight then
            (root_x).parent ← (root_y).parent
            (root_y).weight ← (root_y).weight + (root_x).weight
        else
            (root_y).parent ← (root_x).parent
            (root_x).weight ← (root_x).weight + (root_y).weight
```

When both `Find` with path compression and `Union` with weighted union are used, the time complexity of the union-find algorithm presented in Algorithm 4.3 scales as $O(m\alpha(n))$ [123], where:

- m is the number of union and find operations performed.
- n is the number of network nodes
- α denotes the inverse Ackermann function defined in the following way. First, the Ackermann function $A : \mathbb{N}^2 \to \mathbb{N}$ is defined by :

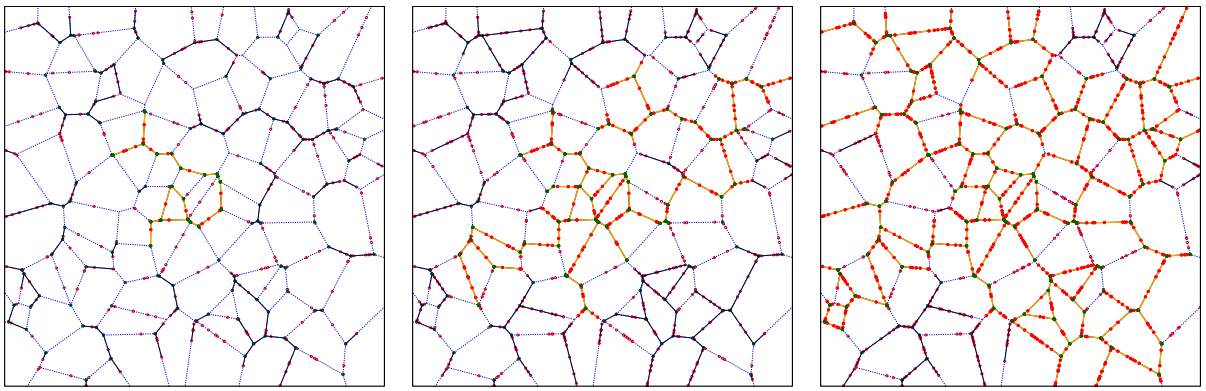
$$\begin{cases} A(0,n) & = n+1 \\ A(m+1,0) & = A(m,1) \\ A(m+1,n+1) & = A(m,A(m+1,n)) \end{cases} \tag{4.3.1}$$

  The inverse Ackermann function α is then defined as the inverse of the increasing mapping $n \mapsto A(n,n)$.

Since the mapping $n \mapsto A(n,n)$ increases incredibly fast, the inverse Ackermann function α increases very slowly. It is often found in the literature that $\alpha(n) \leq 4$ for any practical applications: this is due to the fact that $A(4,4) = 2 \uparrow\uparrow 7 - 3 =: 2^{2^{2^{2^{2^{2^{2}}}}}} - 3 = 2^{2^{2^{65536}}} - 3$. A direct consequence is that the amortized time complexity (i.e. the time complexity per network node) of Algorithm 4.3 is almost linear in the number m of union and find operations performed.

4.3.3 Assessing whether the simulated graph percolates

Once the union-find algorithm has been performed, we have access to all the connected components of the network connectivity graph $\mathcal{G}$. To assess the percolation of the simulated graph, we use a *window-crossing method*: we declare that the simulated graph percolates if there exists a left-right or a top-bottom crossing of the simulation window $W = [0, \texttt{win}]^2$ by a connected component of $\mathcal{G}$. Examples are provided in Figure 4.3.



(a) In this simulation, the connectivity graph does not percolate: the largest connected component of $\mathcal{G}$ does not cross the simulation window.

(b) In this simulation, the connectivity graph does not percolate: the largest connected component of $\mathcal{G}$ does not cross the simulation window. However, it reaches the right side of the simulation window.

(c) In this simulation, the connectivity graph percolates: the largest connected component of $\mathcal{G}$ crosses the simulation window from left to right and from bottom to top.

Figure 4.3 – Assessing whether a simulated connectivity graph percolates or not. As before, the blue dashed lines represent the PVT street system S, the red points represent network users (Cox process X^λ) and the green points represent relays (Bernoulli process Y). The largest connected component in the simulation window is emphasised as before: points of X^λ and Y are filled with their respective colours and connections are highlighted in orange. Edges of the smaller connected component are highlighted in black.

Note that we cannot require an *exact* crossing of the simulation window, i.e. a crossing between two network nodes $Z_i \neq Z_j$ located exactly on the boundary ∂W of the simulation window. Indeed, the chosen models for S, X^λ and Y imply the probability $\mathbb{P}(Z(\partial W) > 0)$ that some network node is located on the boundary of the simulation window is exactly 0. Thus, we look at crossings of the simulation window within a margin $\epsilon \ll \texttt{win}$. In other words, a left-right crossing of the simulation window happens when a connected component of $\mathcal{G}$ connects a network node located in the left margin $[0, \epsilon] \times [0, \texttt{win}]$ to a network node located in the right margin $[\texttt{win} - \epsilon, \texttt{win}] \times [0, \texttt{win}]$. In the same way, a top-bottom crossing of the simulation window happens when when a connected component of $\mathcal{G}$ connects a network node located in the top margin $[0, \texttt{win}] \times [\texttt{win} - \epsilon, \texttt{win}]$ to a network node located in the bottom margin $[0, \texttt{win}] \times [0, \epsilon]$. In practice, the margin for window-crossing simulations is most of the time chosen to be $\epsilon = r/2$. This is due to the following analogy: in networks modelled by Gilbert graphs, the coverage zone of a network node is a circle centered at that node and with radius $r/2$, as has been seen in Figure 3.1. Note

also that $\epsilon = r/2$ is sufficiently small compared to the size `win` of the simulation window.

Using the previous `Find` function (see Algorithm 4.5), checking whether a crossing of the simulation window exists is done straightforwardly. For instance, checking that a left-right crossing exists consists in checking whether a network node in the left margin of the simulation window W has the same root as a network node located in the right margin W. The same approach is done for checking the existence of a top-bottom crossing of the simulation window. Putting this altogether, checking whether the simulated graph $\mathcal{G}$ percolates is simply done by Algorithm 4.7.

Algorithm 4.7: Checking whether the simulated connectivity graph $\mathcal{G}$ percolates.

```
Output← "G does not percolate"
left_nodes ← {Z_i : Z_i ∈ Z ∩ [0, ε] × [0, win]}
right_nodes ← {Z_j : Z_j ∈ Z ∩ [win − ε, win] × [0, win]}
top_nodes ← {Z_k : Z_k ∈ Z ∩ [0, win] × [win − ε, win]}
bottom_nodes ← {Z_l : Z_l ∈ Z ∩ [0, win] × [0, ε]}
for Z_i ∈ left_nodes do
    root_left ← Find(Z_i)
    for Z_j ∈ right_nodes do
        root_right ← Find(Z_j)
        if root_left = root_right then
            Output ← "G percolates"
            break
for Z_k ∈ top_nodes do
    root_top ← Find(Z_k)
    for Z_l ∈ bottom_nodes do
        root_bottom ← Find(Z_l)
        if root_top = root_bottom then
            Output ← "G percolates"
            break
return Output
```

4.3.4 Estimating critical network parameters by appropriate statistical methods

We now have the tools to simulate a connectivity graph $\mathcal{G}$ for a given set of network parameters and check whether the simulated graph percolates. The question is now on how to estimate *critical network quantities* such as p^*, H_c, $p_c(H)$ or $U_c(p, H)$ in a relevant way. First and foremost, note that the definitions of all the above quantities have common points:

- They are all defined as a supremum or an infimum.
- In this supremum or infimum, two of the three network parameters (p, U, H) are fixed and the third one is varying.

- The condition defining the supremum or the infimum is about when the varying parameter reaches some critical value where the percolation probability $P = P(p, U, H)$ becomes positive.

As a matter of fact, the methods for estimating any one of the aforementioned critical network quantities will be quite similar:

- For a given set of parameters (p, U, H), we estimate $P(p, U, H)$ by a Monte-Carlo method:
  - We simulate N realisations (typically, $N = 100$ is sufficient for most practical applications) of the connectivity graph $\mathcal{G} = \mathcal{G}_{p,U,H}$.
  - To each realisation $1 \leq i \leq N$ of $\mathcal{G}_{p,U,H}$, we associate a random variable:

$$y_i = \begin{cases} 1 & \text{if the simulated graph percolates} \\ 0 & \text{if the simulated graph does not percolate} \end{cases}$$

    where, recall, in our simulations, percolation means existence of a left-right or of a top-bottom crossing of the simulation window $[0, \texttt{win}]^2$.
  - We estimate the percolation probability $P(p, U, H)$ by the empirical mean:

$$\hat{y} := \frac{1}{N} \sum_{i=1}^{N} y_i. \tag{4.3.2}$$

    In other words, the percolation probability $P(p, U, H)$ is estimated by the *proportion of simulations* where the simulated graph percolates. Since the $(y_i)_{i \leq i \leq N}$ are i.i.d. distributed like a Bernoulli random variable with parameter $P(p, U, H)$, the empirical mean $\hat{y}$ computed over N simulations is an unbiased estimator of $P(p, U, H)$.
- For a grid of values for the varying parameter (the two other being fixed), we estimate $P(p, U, H)$.
- Plotting $P(p, U, H)$ against the varying parameter, we use an appropriate statistical regression to assess where the critical point is.

We now give a bit more detail for each of the critical network quantities p^*, H_c, $p_c(H)$ and $U_c(p, H)$.

Estimating p^*

Recall that p^* denotes the critical threshold of the PVT site percolation model:

$$p^* := \inf\{p \in [0, 1] : P_{PVT}(p) > 0\}, \tag{4.2.3}$$

where $P_{PVT}(p) =: P(p, U, 0)$ does not depend on U. Moreover, p^* does not depend on L_A. Thus, when estimating p^*, we fix $L_A > 0$, $U = 0$, $H = 0$ (or, equivalently, $\lambda = 0$ and $r = \infty$) and estimate $P(p, 0, 0) = P_{PVT}(p)$ for a grid of values of p using the estimator defined in (4.3.2). The result of this procedure is illustrated by Figure 4.4a. Since for

each value of p, the percolation probability $P_{PVT}(p)$ is estimated by the empirical mean of N Bernoulli random variables (see (4.3.2)), we propose to interpolate the discrete values of the percolation probability obtained by simulation with a logistic model given by the blue curve on Figure 4.4a:

$$\text{logit}(P_{PVT}(p)) := \log\left(\frac{P_{PVT}(p)}{1-P_{PVT}(p)}\right) = ap + b, \tag{4.3.3}$$

where the parameters a and b are estimated by logistic regression. By analogy with discrete percolation theory[4], we estimate the threshold p^* by the inflection point of the logistic curve. Since $\text{logit}''(x) = 0 \Leftrightarrow x = 1/2$ and that $\text{logit}(1/2) = 0$, the logistic model (4.3.3) yields:

$$p^* \approx -\frac{b}{a}.$$

Estimating H_c

H_c denotes the critical threshold of the PVT hard-geometric bond percolation model:

$$H_c := \sup\{H \geq 0 \,:\, P(1,0,H) > 0\}. \tag{4.2.4}$$

Thus, we fix $p = 1$, $U = 0$ and, for a grid of values of H, we estimate $P(1,0,H)$ by the proportion of simulations where the simulated network graph $\mathcal{G}_{1,0,H}$ percolates. Results are illustrated by Figure 4.4b.

Since $P(1,0,H)$ is decreasing as a function of H, we interpolate the discrete values by a "reverse" logistic model given by the blue curve on Figure 4.4b:

$$\text{logit}(1-P(1,0,H)) =: \log\left(\frac{1-P(1,0,H)}{P(1,0,H)}\right) = aH + b, \tag{4.3.4}$$

and we estimate H_c by the inflection point of the "reverse" logistic curve. In the same way as before, this yields:

$$H_c \approx -\frac{b}{a}.$$

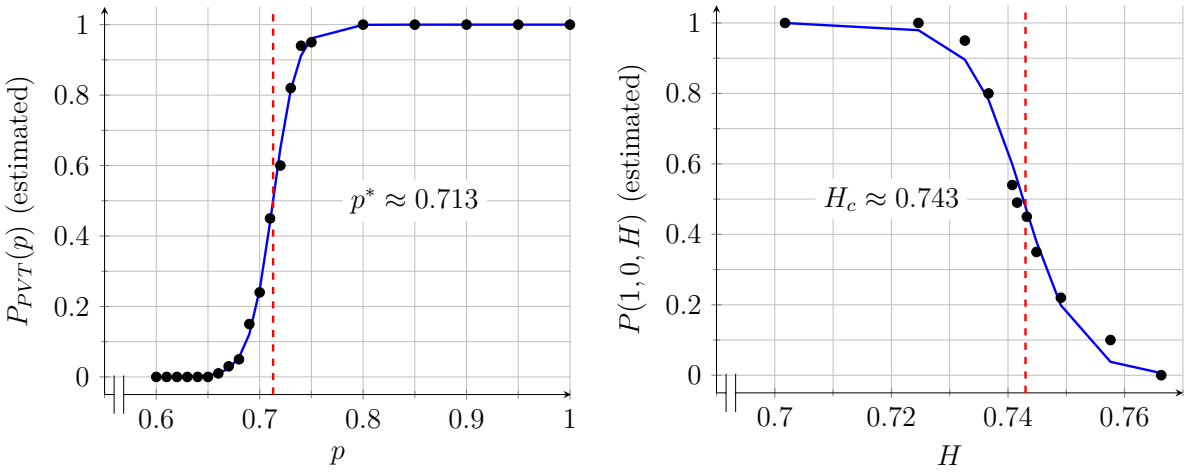
Estimating $p_c(H)$

For $H < H_c$, $p_c(H)$ denotes the boundary of the permanently supercritical range of percolation of $\mathcal{G}$, i.e. the boundary of the relay-limited connectivity regime:

$$p_c(H) := \inf\{p > 0 \,:\, P(p,0,H) > 0\}. \tag{4.2.6}$$

Therefore, for estimating $p_c(H)$, we fix $H < H_c$ and $U = 0$ and, for a grid of values for p, we again estimate $P(p,0,H)$ by the proportion of simulations where the simulated

[4]Consider for instance independent bond percolation on the square lattice $\mathbb{L}^2$ with parameter p and denote $p_c := p_c(2)$ the critical parameter. The event {there exists *some* infinite connected component} (not necessarily containing the origin!) is a tail event. Denoting $\Theta(p)$ the probability of this event, it is easy to show, thanks to Kolmogorov's 0-1 law, that $\Theta(p) = 0$ for $p < p_c$ and $\Theta(p) = 1$ for $p > p_c$. As a matter of fact, the plot of the mapping $p \mapsto \Theta(p)$ is a 0-1 step function with step occurring at p_c, so that we can somehow see p_c as the inflection point of this curve.



(a) Estimation of the PVT site percolation threshold p^*.

(b) Estimation of the PVT hard-geometric bond percolation threshold H_c.

Figure 4.4 – Left: Estimation of the PVT site percolation threshold p^*. Right: Estimation of the PVT hard-geometric bond percolation threshold H_c. The black points represent the estimations of the percolation probability obtained by simulations, the blue curves represent the logistic (left figure) and "reverse" logistic (right figure) models. The red vertical line corresponds to the inflection point of the logistic curve and is supposed to intercept the percolation threshold. The simulations have been done with a square simulation window of size $\texttt{win} = 30\,\text{km}$ and a PVT street density $L_A = 20\,\text{km}^{-1}$.

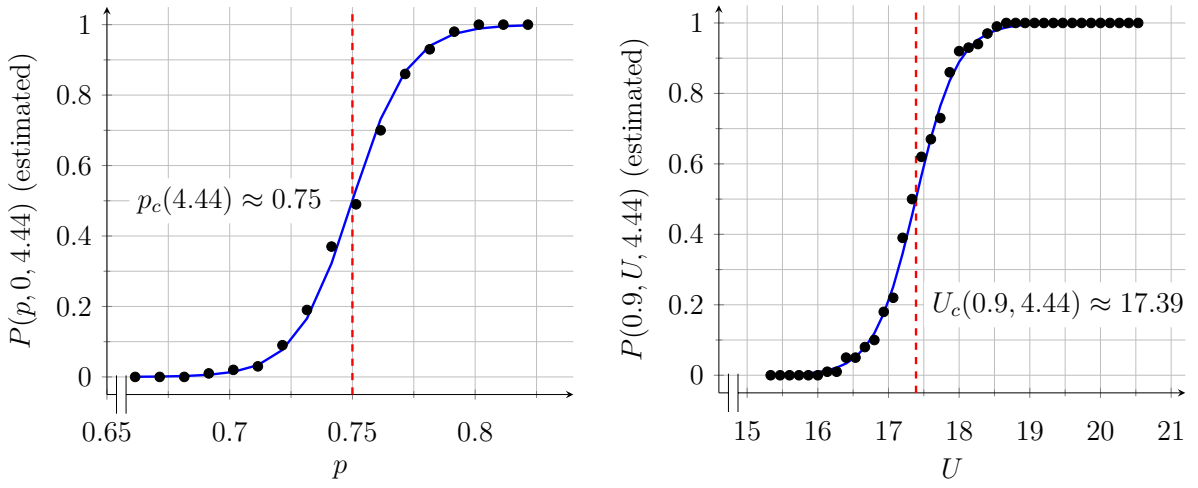
graph $\mathcal{G}_{p,0,H}$ percolates. $P(p, 0, H)$ being increasing as a function of H, we can again resort to logistic regression to interpolate the discrete values of the percolation probability $P(p, 0, H)$ obtained by discrete simulations. Once again, the threshold $p_c(H)$ is estimated by the inflection point of the logistic curve, see Figure 4.5a.

Estimating $U_c(p, H)$

$U_c(p, H)$ denotes the critical user density in the critical range of percolation of $\mathcal{G}$ (relay-and-user-limited connectivity regime):

$$U_c(p, H) := \inf\{U \geq 0 : P(p, U, H) > 0\} \tag{4.2.2}$$

In the same fashion as before, for fixed (p, H) and for a grid of values for U, we estimate $P(p, U, H)$ using the proportion of simulations where the simulated graph $\mathcal{G}_{p,U,H}$ percolates. $P(p, U, H)$ being increasing as a function of U, we can again interpolate the discrete values of the percolation probability $P(p, U, H)$ by a logistic model and the percolation threshold $U_c(p, H)$ is estimated by the inflection point of the logistic curve, see Figure 4.5b.



(a) Estimation of $p_c(H = 4.44)$.

(b) Estimation of $U_c(p = 0.9, H = 4.44)$.

Figure 4.5 – Left: Example of estimation of $p_c(H)$ in the permanently subcritical range of percolation of $\mathcal{G}$ (relay-limited connectivity). Right: Example of estimation of the critical user density $U_c(p, H)$ in the critical range of percolation of $\mathcal{G}$ (relay-and-user-limited connectivity). The black points represent the estimations of the percolation probability obtained by simulations, the blue curves represent the logistic models. The red vertical line corresponds to the inflection point of the logistic curve and is supposed to intercept the percolation threshold. The simulations have been done with a square simulation window of size $\texttt{win} = 15\,\text{km}$ and a PVT street density $L_A = 20\,\text{km}^{-1}$.

4.4 Proofs of analytical results

4.4.1 General approach

Theorems 4.2.1 to 4.2.3 and their corollaries have been stated in terms of the dimensionless parameters (p, U, H) introduced in Section 4.1.3. This allowed us to introduce the different connectivity regimes and their frontiers in a more applied fashion, as illustrated by Figure 4.1.

However, when working on the proofs of the aforementioned results, it will be much more convenient to come back to the original parameters (L_A, p, λ, r). This is mostly due to the fact that H, being inversely proportional to r, is less easy to work with when considering particular events related to connectivity in the network graph. Moreover, as mentioned earlier, S is motion-invariant and percolation of the connectivity graph does not depend on the street intensity L_A, so that we can set $L_A = 1$ without loss of generality in what follows. Switching back to the original network parameters $(L_A = 1, p, \lambda, r)$, we will therefore now refer to the connectivity graph as $\mathcal{G} = \mathcal{G}_{p,\lambda,r}$ and prove the following equivalent formulations of Theorems 4.2.1 to 4.2.3:

Reformulation 1 (Reformulation of Theorem 4.2.1)**.** *For small enough $r \in (0, \infty]$ and small enough $\lambda > 0$, $\mathcal{G}_{1,\lambda,r}$ does not percolate and, consequently, by monotonicity of the model, $\mathcal{G}_{p,\lambda,r}$ does not percolate either.*

Reformulation 2 (Reformulation of Theorem 4.2.2)**.** *For large enough $p \in (0, 1)$, $\mathcal{G}_{p,\lambda,r}$ percolates for any $r \in (0, \infty]$ and large enough $\lambda < \infty$ (depending on r).*

Reformulation 3 (Reformulation of Theorem 4.2.3)**.** *For large enough $p < 1$ and large enough $r < \infty$, $\mathcal{G}_{p,0,r}$ percolates.*

The general idea of the proofs of the above reformulations will consist in using a renormalization argument. More precisely, we will most of the time proceed as follows:

1. Map the percolation of $\mathcal{G}$ to a discretized percolation process on a rescaled Euclidean lattice
2. Depending on the needs, relate the percolation (or the absence of percolation) of $\mathcal{G}$ to the percolation (or absence of percolation) of the discretized process
3. Prove that the discretized process is k-dependent for some $k \geq 1$.
4. Apply the result from Liggett, Schonmann and Stacey (Theorem 4.1.1) to conclude that the discretized process is dominated by a Bernoulli percolation process and thereby draw conclusions on $\mathcal{G}$.

4.4.2 Preparation and notations

Before proceeding with the proofs of the analytical results stated in Section 4.2.2, we begin with introducing a few notations and definitions that will be useful for the purposes of our developments.

Recall that V and E respectively denote the set of vertices and edges of the PVT street system S and that, furthering the analogy between S and a street system, elements of V (respectively of E) are called crossroads (respectively streets). A topologically connected subset $s \subset e$ of some street $e \in E$ will be called *street segment*. The length of a street $e \in E$ is noted by $|e|$ and the length of a street segment s will likewise be denoted by $|s|$.

For $A \subset \mathbb{R}^2$ and $B \subset \mathbb{R}^2$, we denote the Euclidean distance between A and B by:

$$\mathrm{dist}_2(A, B) \coloneqq \inf\{\|x - y\|_2,\, x \in A,\, y \in B\}.$$

For $x \in \mathbb{R}^2$ and $a > 0$ we denote by $Q_a(x) \coloneqq x + [-a/2, a/2]^2$ the square of side a centered at x. We note that this is exactly the definition of the closed ball $\mathscr{B}(x, a/2)$ with center x and radius $a/2$ for the supremum norm of $\mathbb{R}^2$:

$$Q_a(x) = \{y : \|y - x\|_\infty \leq a/2\} = \mathscr{B}(x, a/2).$$

For simplicity, whenever $a = n \in \mathbb{N}$, we will write Q_n to mean $Q_n(0)$.

We denote by $\mathbf{M}$ the space of Borel measures on $\mathbb{R}^2$, equipped with the evaluation σ-algebra[5]. For a (possibly random) Borel measure μ on $\mathbb{R}^2$ and $A \subset \mathbb{R}^2$, we denote the restriction of μ to A by $\mu_A(\cdot) \coloneqq \mu(A \cap \cdot)$. We also adapt the definition of the support of μ as follows: the *support* of μ is the following set:

$$\mathrm{supp}(\mu) \coloneqq \{x \in \mathbb{R}^d : \forall \varepsilon > 0,\, \mu(Q_\varepsilon(x)) > 0\}$$

[5]Recall from Section 1.1.2, that the evaluation σ-algebra is the smallest σ-algebra making the mappings $\mathbf{M} \ni \Xi \mapsto \Xi(B)$ measurable for all Borel sets B of $\mathbb{R}^2$.

The following concepts of *stabilization* and *asymptotic essential connectedness*, both introduced in [70] will also be needed for investigating spatial dependencies of random measures.

Definition 4.4.1. [70, Definition 2.3] A random measure Ξ on $\mathbb{R}^2$ is called *stabilizing* if there exists a random field of stabilization radii $R = \{R_x\}_{x\in\mathbb{R}^2}$ defined on the same probability space as Ξ and Ξ-measurable, such that:

(1) (Ξ, R) are jointly stationary,

(2) $\lim\limits_{n\uparrow\infty} \mathbb{P}\left(\sup\limits_{y\in Q_n\cap\mathbb{Q}^2} R_y < n\right) = 1,$

(3) for all $n \geq 1$, the random variables

$$\left\{ f\left(\Xi_{Q_n(x)}\right) \mathbb{1}\left\{ \sup_{y\in Q_n(x)\cap\mathbb{Q}^2} R_y < n \right\} \right\}_{x\in\varphi}$$

are independent for all bounded measurable functions

$$f : \mathbf{M} \to [0, +\infty)$$

and finite $\varphi \subset \mathbb{R}^2$ such that $\forall\, x \in \varphi$, $\mathrm{dist}_2(x, \varphi \setminus \{x\}) > 3n$.

We slightly modify the definition of asymptotic essential connectedness given in [70] for the sake of simplicity and use the following definition:

Definition 4.4.2. Let Ξ be a random measure on $\mathbb{R}^2$. Then Ξ is *asymptotically essentially connected* if there exists a random field $R = \{R_x\}_{x\in\mathbb{R}^d}$ such that Ξ is stabilizing with R as in Definition 4.4.1 and if for all $n \geq 1$, whenever $\sup\limits_{y\in Q_{2n}\cap\mathbb{Q}^2} R_y < n/2$, the following assertions are satisfied:

(1) $\mathrm{supp}(\Xi_{Q_n}) \neq \emptyset$

(2) $\mathrm{supp}(\Xi_{Q_n})$ is contained in a connected component of $\mathrm{supp}(\Xi_{Q_{2n}})$

The following result is stated in [70, Example 3.1] for a slightly modified version of Definition 4.4.2. It is easy to check that it adapts in our case as follows:

Proposition 4.4.3. *Let Λ denote the random measure on $\mathbb{R}^2$ defined by $\Lambda(dx) = \nu_1(S \cap dx)$, where S is the PVT generated by a planar homogeneous stationary Poisson point process X_S. Then Λ is stabilizing and asymptotically essentially connected with the following stabilization field:*

$$\forall x \in \mathbb{R}^2,\ R_x := \inf\{\|x - X_{S,i}\|,\ X_{S,i} \in X_S\},$$

For simplicity, whenever $x \in \mathbb{R}^2$ and $n \in \mathbb{N}$ we denote:

$$R(Q_n(x)) := \sup_{y\in Q_n(x)\cap\mathbb{Q}^2} R_y.$$

Finally, we will use coarse-graining arguments to map our continuum models to discretized percolation processes. Hence, we will need to define notions of openness and closedness of crossroads and street segments (possibly the whole streets themselves). This is done as follows:

Definition 4.4.4 (Open/Closed crossroad)**.** We say a crossroad $v \in V$ is *open* if it is an atom of the point process of relays Y, i.e. $Y(\{v\}) = 1$. In other words, v is open if there is a relay at v. We say v is *closed* if it is not open.

Definition 4.4.5 (Open/Closed street segment)**.** Let $e \in E$ be a street and let $s \subseteq e$ be a non-empty street segment. We say s is *open* if either of the two following set of conditions are satisfied:

1. $|s| \leq r$

   OR

2. $\begin{cases} |s| > r \\ \forall c \subset s, \ (|c| = r \ , c \text{ connected and } c \text{ topologically closed}) \Rightarrow X^{\lambda}(c) \geq 1 \end{cases}$

In other words, an open street segment allows D2D signals to be transmitted over its range, either by being shorter than the D2D range r or by not having a disconnection due to a gap. We say that s is *closed* if s is not open.

4.4.3 Proof of Theorem 4.2.1

To prove Theorem 4.2.1, we switch back to the original network parameters and prove the equivalent Reformulation 1. In other words, we aim at showing that $\mathcal{G}$ does not percolate when $p = 1$ and λ, r are sufficiently small but positive. We will introduce a discrete site percolation model on the square lattice constructed in such a way that if it does not percolate, then neither does $\mathcal{G}$. Proving the absence of percolation of the square lattice model will then be done via appealing to its local dependence.

To this end, for $n \geq 1$, say a site $z \in \mathbb{Z}^2$ is *n-good* if the following conditions are satisfied:

(1) $R(Q_n(nz)) < n$

(2) $\forall e \in E$ if $s_{z,e} := e \cap Q_n(nz) \neq \emptyset$, then $s_{z,e}$ is closed.

Say a site $z \in \mathbb{Z}^2$ is *n-bad* if it is not n-good.

Our first claim is the following:

Lemma 4.4.6. *Percolation of $\mathcal{G}$ implies percolation of the process of n-bad sites.*

Proof. Assume $\mathcal{G}$ percolates and denote by $\mathcal{C}$ an unbounded (connected) component of $\mathcal{G}$. Denote $\mathcal{Z} = \mathcal{Z}_n := \{z \in \mathbb{Z}^2 : \mathcal{C} \cap Q_n(nz) \neq \emptyset\}$. Since $\mathcal{C}$ is unbounded we have $\#(\mathcal{Z}) = \infty$. Observe for all $z \in \mathcal{Z}$, z is n-bad since condition (2) of n-goodness is not satisfied (there exists an open street segment intersecting $Q_n(nz)$). Also, $\mathcal{Z}$ is almost surely connected [6] in $\mathbb{Z}^2$. This follows from the fact that the probability that some edge $e \in E$ of the PVT intersects $\mathbb{Z}^2$ is equal to zero [7]. Hence, the process of n-bad sites percolates. □

[6] $z, z' \in \mathbb{Z}^2$ with $z \neq z'$ are said to be connected in $\mathbb{Z}^2$ whenever $\|z - z'\|_1 = 1$.

[7] This is true for the Voronoi tessellation generated by any stationary point process as a consequence of the fact that such a process does not have points which are equidistant to a given, fixed location ; see e.g. [12, Lemma 6.2.6].

By Lemma 4.4.6, it suffices to prove that the process of n-bad sites does not percolate (for some n) when λ and r are sufficiently small but positive. This will be done using the fact that it is a 3-dependent percolation model on the square lattice.

Lemma 4.4.7. *For $z \in \mathbb{Z}^2$, set $\zeta_z := \mathbb{1}\{z \text{ is } n\text{-bad}\}$. Then $(\zeta_z)_{z\in\mathbb{Z}^2}$ is a 3-dependent random field for the supremum metric $\|\cdot\|_\infty$.*

Proof. As a starting point, note that $\forall z \in \mathbb{Z}^2, \zeta_z = 1 - \mathbb{1}\{z \text{ is } n\text{-good}\}$. It is therefore equivalent to prove that the process of n-good sites is 3-dependent.
For $z \in \mathbb{Z}^2$, set $\xi_z = \mathbb{1}\{z \text{ is } n\text{-good}\}$. Let $\{z_1, \dots z_q\} \subset \mathbb{Z}^2$ be such that $\forall i \neq j, \|z_i - z_j\|_\infty > 3$. We want to show that the random variables $(\xi_{z_i})_{1\leq i\leq q}$ are independent. Since we are dealing with indicator functions, this is equivalent to showing that:

$$\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) = \prod_{i=1}^{q} \mathbb{E}(\xi_{z_i}).$$

Now, we have:

$$\begin{aligned}
\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) &= \mathbb{E}\left[\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\middle|\Lambda\right)\right] \\
&= \mathbb{E}\left[\mathbb{E}\left(\prod_{i=1}^{q} \mathbb{1}\{R(Q_n(nz_i)) < n\} \prod_{i=1}^{q} \mathbb{1}\{\forall\, e \in E,\, s_{z_i,e} \text{ is closed}\}\middle|\Lambda\right)\right] \qquad (4.4.1) \\
&= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{R(Q_n(nz_i) < n\}\mathbb{E}\left(\prod_{i=1}^{q} \mathbb{1}\{\forall\, e \in E,\, s_{z_i,e} \text{ is closed}\}\middle|\Lambda\right)\right], \qquad (4.4.2)
\end{aligned}$$

where we have used Λ-measurability of the random variables $\{R_x\}_{x\in\mathbb{R}^2}$ in (4.4.2).

For $1 \leq i \leq q$, set $A_{z_i} := \{\forall\, e \in E,\, s_{z_i,e} \text{ is closed}\}$. According to Definition 4.4.5, for a given $1 \leq i \leq q$, the event A_{z_i} only depends on the configuration of the random measure Λ and of the Cox point process X^λ inside the square $Q_n(nz_i)$. Therefore, given Λ, the events $\{A_{z_i} : 1 \leq i \leq q\}$ only depend on $X^\lambda \cap Q_n(nz_i)$, $1 \leq i \leq q$. Since we have $\forall i \neq j, \|z_i - z_j\|_\infty > 3$, then the squares $Q_n(nz_i)$ are disjoint. Moreover, given Λ, X^λ has the distribution of a Poisson Point Process. Thus, by Poisson independence property, the events $(A_{z_i})_{1\leq i\leq q}$ are conditionally independent given Λ. Hence (4.4.2) yields:

$$\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) = \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{R(Q_n(nz_i) < n\} \prod_{i=1}^{q} \mathbb{E}\left(\mathbb{1}\{\forall\, e \in E,\, s_{z_i,e} \text{ is closed}\}\middle|\Lambda\right)\right]. \qquad (4.4.3)$$

Set $f(\Lambda_{Q_n(x)}) := \mathbb{E}\left(\mathbb{1}\{\forall\, e \in E,\, s_{x,e} \text{ is closed}\}\middle|\Lambda\right)$. Then f is a deterministic, bounded and measurable function of $\Lambda_{Q_n(x)}$. Moreover, the set $\varphi := \{nz_1, \dots, nz_q\} \subset \mathbb{R}^2$ is a finite subset of $\mathbb{R}^2$ satisfying:

$$\forall i \neq j, \|nz_i - nz_j\|_\infty > 3n.$$

Since the infinite norm is always upper bounded by the Euclidean norm, we have $\forall i \neq j, \|nz_i - nz_j\|_2 > 3n$, and so φ satisfies:

$$\forall x \in \varphi, \operatorname{dist}_2(x, \varphi \setminus \{x\}) > 3n.$$

Hence, by condition (3) in the definition of stabilization (Definition 4.4.1), the random variables appearing in the right-hand side of (4.4.3) are independent. This yields:

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\right) &= \prod_{i=1}^{q}\mathbb{E}\left(\mathbb{1}\{R(Q_n(nz_i)<n\}\prod_{i=1}^{q}\mathbb{E}\left(\mathbb{1}\{\forall\, e\in E,\, s_{z_i,e}\text{ is closed}\}\middle|\Lambda\right)\right)\\ &= \prod_{i=1}^{q}\mathbb{E}\left(\xi_{z_i}\right),\end{aligned}$$

thus concluding the proof of the lemma. □

Now we prove that the probability for an arbitrary site, which by stationarity can be chosen to be the origin $\mathbf{0}\in\mathbb{Z}^2$, to be n-bad can be made arbitrarily small when first taking some large enough finite n and then positive small enough λ, r, as stated in the following Lemma.

Lemma 4.4.8.

$$\lim_{n\uparrow\infty}\lim_{\lambda,r\downarrow 0}\mathbb{P}(\mathbf{0}\ \textit{is } n\textit{-bad}) = 0$$

Proof. Note that we have:

$$\begin{aligned}\mathbb{P}(0\,\text{is } n\text{-bad}) &= \mathbb{P}\Big(\{R(Q_n)\geq n\}\cup\{\exists\, e\in E:\, e\cap Q_n\neq\emptyset\text{ and open}\}\Big)\\ &\leq \mathbb{P}\left(R(Q_n)\geq n\right)+\mathbb{P}\left(\exists\, e\in E:\, e\cap Q_n\neq\emptyset\text{ and open}\right)\\ &\leq \mathbb{P}\left(R(Q_n)\geq n\right) &\text{(a)}\\ &\quad+\mathbb{P}\left(\exists\, e\in E:\ 0<|e\cap Q_n|\leq r\right) &\text{(b)}\\ &\quad+\mathbb{P}\left(\exists\, e\in E:\ e\cap Q_n\text{ satisfies condition (2) in Definition 4.4.5}\right). &\text{(c)}\end{aligned}$$

Take any $\epsilon>0$. By the stabilization property of the PVT (Proposition 4.4.3) we have $\lim_{n\uparrow\infty}\mathbb{P}(R(Q_n)\geq n)=0$ and so we can fix n large enough to make the probability in (a) smaller than $\epsilon/3$. Then, Q_n intersects almost surely zero or a finite number of edges $e\in E$. Hence the probability in (b) converges to 0 when $r\to 0$ and, consequently, we can take r small enough to make the probability in (b) smaller than $\epsilon/3$. Finally, for given n (and independently of r), we can take λ small enough to make the probability in (c) smaller than $\epsilon/3$. Indeed, this latter probability is dominated by the probability that $X^{\lambda}(Q_n)\geq 1$ and thus converges to 0 when $\lambda\to 0$ for any finite n. This concludes the proof of Lemma 4.4.8. □

By Lemmas 4.4.7 and 4.4.8, using Theorem 4.1.1, for large enough $n<\infty$ and small enough $r>0, \lambda>0$, the process of n-bad sites is stochastically dominated from above by an independent site percolation model on the square lattice where the probability of having an open site is arbitrarily small. Hence this independent site percolation model is sub-critical. Consequently, we can make the process of n-bad sites non-percolating. By Lemma 4.4.6 the same is true for $\mathcal{G}$, thus concluding the proof of Reformulation 1 and the equivalent Theorem 4.2.1.

4.4.4 Proof of Theorem 4.2.2

As before, we shall prove the equivalent Reformulation 2. This resorts to showing that for large enough $p < 1$ the connectivity graph $\mathcal{G}$ percolates for all $r > 0$ and large enough $\lambda < \infty$ (depending on r).

We will again use a renormalisation argument. To this end, consider the following percolation model on the square lattice $\mathbb{Z}^2$. For $n \geq 1$, say a site $z \in \mathbb{Z}^2$ is n-good if the following conditions are satisfied:

(1) $R(Q_{6n}(nz)) < 6n$.

(2) $E \cap Q_n(nz) \neq \emptyset$, i.e. the square $Q_n(nz)$ contains a *full* street (not just a street segment).

(3) There exists $e \in E \cap Q_n(nz)$ such that e is open, in the sense of Definition 4.4.5. In other words, there exists an open street which is fully included in the square $Q_n(nz)$.

(4) All crossroads in $Q_{6n}(nz)$ are open, in the sense of Definition 4.4.4.

(5) Every two open edges $e, e' \in E \cap Q_{3n}(nz)$ are connected by a path in $\mathcal{G} \cap Q_{6n}(nz)$.

We say a site $z \in \mathbb{Z}^2$ is n-bad if it is not n-good.

The n-good sites have been defined so as to satisfy the following implication.

Lemma 4.4.9. *Percolation of the process of n-good sites implies percolation of the connectivity graph $\mathcal{G}$.*

Proof. Let $\mathcal{C}$ be an infinite connected component of n-good sites. Consider $z, z' \in \mathcal{C}$ such that $\|z - z'\|_1 = 1$. Without loss of generality, assume $z = (a, b)$ for some $a, b \in \mathbb{Z}$ and $z' = (a+1, b)$. By conditions (2) and (3) in the definition of n-goodness, we can find open edges $e \in E \cap Q_n(nz)$ and $e' \in E \cap Q_n(nz')$. Since

$$\begin{aligned} Q_n(nz) &= [na - n/2, na + n/2] \times [nb - n/2, nb + n/2]\,, \\ Q_n(nz') &= [na + n/2, na + 3n/2] \times [nb - n/2, nb + n/2]\,, \\ Q_{3n}(nz) &= [na - 3n/2, na + 3n/2] \times [nb - 3n/2, nb + 3n/2]\,, \\ Q_{6n}(nz) &= [na - 3n, na + 3n] \times [nb - 3n, nb + 3n]\,, \end{aligned}$$

therefore, we have $Q_n(nz') \subset Q_{3n}(nz)$ and so $e' \in E \cap Q_n(nz')$ implies $e' \in E \cap Q_{3n}(nz)$. Since we also have $e \in E \cap Q_n(nz) \subset E \cap Q_{3n}(nz)$ and e, e' are both open, by conditions (4) and (5) in the definition of n-goodness, e and e' are connected by a path $\mathcal{L}$ in $\mathcal{G} \cap Q_{6n}(nz)$. Therefore, the path $\mathcal{L}$ also connects e and e' in $\mathcal{G}$, thus giving rise to an infinite connected component in $\mathcal{G}$. This concludes the proof of Lemma 4.4.9. □

Lemma 4.4.10. *For $z \in \mathbb{Z}^2$, set $\xi_z := \mathbb{1}\{z \text{ is } n\text{-good}\}$. Then $(\xi_z)_{z \in \mathbb{Z}^2}$ is an 18-dependent random field for the supremum metric $\|\cdot\|_\infty$.*

Proof. In the same way as in the proof of Lemma 4.4.7, it suffices to prove that for all finite $\psi = \{z_1, \dots z_q\} \subset \mathbb{Z}^2$ such that $\forall i \neq j, \|z_i - z_j\|_\infty > 18$, we have:

$$\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) = \prod_{i=1}^{q} \mathbb{E}(\xi_{z_i}).$$

Denote respectively by A_z, B_z, C_z, D_z, F_z the events that the conditions (1), (2), (3), (4), (5) in the definition of n-goodness hold for $z \in \mathbb{Z}^2$. We thus have:

$$\forall z \in \mathbb{Z}^2, \xi_z = \mathbb{1}\{A_z\}\mathbb{1}\{B_z\}\mathbb{1}\{C_z\}\mathbb{1}\{D_z\}\mathbb{1}\{F_z\}.$$

Note first that whenever $z \in \mathbb{Z}^2$, the indicators $\mathbb{1}\{A_z\}$ and $\mathbb{1}\{B_z\}$ are Λ-measurable. Thus, we have :

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) &= \mathbb{E}\left[\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\middle|\Lambda\right)\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{E}\left(\prod_{i=1}^{q} \mathbb{1}\{C_{z_i} \cap D_{z_i} \cap F_{z_i}\}\middle|\Lambda\right)\right].\end{aligned}$$

Now, note that conditioned on Λ, for each $1 \leq i \leq q$, the event $C_{z_i} \cap D_{z_i} \cap F_{z_i}$ only depends on the configuration of X^λ and Y inside of the square $Q_{6n}(nz_i)$. Since ψ satisfies $\forall i \neq j, \|z_i - z_j\|_\infty > 18$, then we have $\forall i \neq j, \|nz_i - nz_j\|_\infty > 18n$. As a matter of fact, the squares $\{Q_{6n}(nz_i)) \,:\, 1 \leq i \leq q\}$ are disjoint, i.e.

$$\forall i \neq j, Q_{6n}(nz_i) \cap Q_{6n}(nz_j) = \emptyset.$$

By the complete independence of Poisson and Bernoulli processes (recall that, given Λ, X^λ has the distribution of a Poisson point process and Y the distribution of a Bernoulli point process), we have

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{E}\left(\prod_{i=1}^{q} \mathbb{1}\{C_{z_i} \cap D_{z_i} \cap F_{z_i}\}\middle|\Lambda\right)\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\prod_{i=1}^{q}\mathbb{E}\left(\mathbb{1}\{C_{z_i} \cap D_{z_i} \cap F_{z_i}\}\middle|\Lambda\right)\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{A_{z_i}\}\prod_{i=1}^{q}\mathbb{E}\left(\mathbb{1}\{B_{z_i} \cap C_{z_i} \cap D_{z_i} \cap F_{z_i}\}\middle|\Lambda\right)\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{R(Q_{6n}(nz_i) < 6n\} f(\Lambda_{Q_{6n}(nz_i)})\right], \qquad (4.4.4)\end{aligned}$$

where $f(\Lambda_{Q_{6n}(x)}) \coloneqq \mathbb{E}\left(\mathbb{1}\{B_x \cap C_x \cap D_x \cap F_x\} \,|\, \Lambda\right)$, a bounded measurable deterministic function of $\Lambda_{Q_{6n}(x)}$, and where by Λ-measurability of the events $\{B_{z_i} \,:\, 1 \leq i \leq q\}$, we put their indicators into the conditional expectation given Λ. Now, the set $\varphi \coloneqq \{nz_1, \dots, nz_q\} \subset \mathbb{R}^2$ is finite and satisfies:

$$\forall i \neq j, \|nz_i - nz_j\|_\infty > 18n.$$

Since the infinite norm is always upper bounded by the Euclidean norm, we have $\forall i \neq j, \|nz_i - nz_j\|_2 > 18n$, and so φ satisfies:

$$\forall x \in \varphi,\ \mathrm{dist}_2(x, \varphi \setminus \{x\}) > 18n = 3 \times 6n.$$

We can therefore apply condition (3) in Definition 4.4.1 (with n replaced by $6n$) to get that the random variables appearing in the right-hand side of (4.4.4) are independent. Hence:

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) &= \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{R(Q_{6n}(nz_i) < 6n\} f(\Lambda_{Q_{6n}(nz_i)})\right] \\ &= \prod_{i=1}^{q} \mathbb{E}(\xi_{z_i}),\end{aligned}$$

which concludes the proof of Lemma 4.4.10. □

Again, we prove that the probability for an arbitrary site to be n-good can be made arbitrarily large:

Lemma 4.4.11. *For any $r > 0$ we have*

$$\lim_{n\uparrow\infty} \lim_{p\uparrow 1,\, \lambda\uparrow\infty} \mathbb{P}(\mathbf{0} \textit{ is } n\textit{-good}) = 1$$

Proof. We shall prove that

$$\lim_{n\uparrow\infty} \lim_{p\uparrow 1,\, \lambda\uparrow\infty} \mathbb{P}(\mathbf{0} \text{ is } n\text{-bad}) = 0.$$

Take any $\epsilon > 0$. Denote respectively by A, B, C, D, F the events that the conditions (1), (2), (3), (4), (5) in the definition of n-goodness hold for $z = \mathbf{0}$. Denote also by $\tilde{A}$ the event that $R(Q_{6n}) < n/2$. Note that $\tilde{A} \subset A$ and thus we have:

$$\begin{aligned}\mathbb{P}(\mathbf{0} \text{ is } n\text{-bad}) &= \mathbb{P}(A^c \cup B^c \cup C^c \cup D^c \cup F^c) \\ &\leq \mathbb{P}(\tilde{A}^c \cup B^c \cup C^c \cup D^c \cup F^c) \\ &\leq \mathbb{P}(\tilde{A}^c) + \mathbb{P}(B^c) + \mathbb{P}(B \cap C^c) + \mathbb{P}(D^c) + \mathbb{P}(\tilde{A} \cap D \cap F^c).\end{aligned}$$

First, partitioning the square Q_{6n} into $12^2 = 144$ subsquares $(Q_i)_{1\leq i\leq 144}$ of side length $n/2$, we get:

$$\begin{aligned}\mathbb{P}(\tilde{A}^c) &= \mathbb{P}(R(Q_{6n}) \geq n/2) \\ &= \mathbb{P}\left(\bigcup_{i=1}^{144} \{R(Q_i) \geq n/2\}\right) \\ &\leq 144\, \mathbb{P}(R(Q_{n/2}) \geq n/2) \qquad \text{by stationarity of the } R\text{'s.}\end{aligned}$$

Therefore, by condition (2) of Definition 4.4.1, we get $\lim_{n\uparrow\infty} \mathbb{P}(\tilde{A}^c) = 0$. Also

$$\mathbb{P}(B^c) = \mathbb{P}(E \cap Q_n = \emptyset)$$

and thus $\lim_{n\uparrow\infty} \mathbb{P}(B^c) = 0$. Fix n large enough such that $\mathbb{P}(\tilde{A}^c) \leq \epsilon/5$ and $\mathbb{P}(B^c) \leq \epsilon/5$. For such n, Q_n, Q_{3n} and Q_{6n} intersect almost surely zero or a finite number of edges and vertices.

Let's now deal with the quantity $\mathbb{P}(B \cap C^c)$. We have:

$$\mathbb{P}(B \cap C^c) = \mathbb{P}(E \cap Q_n \neq \emptyset \,\text{and}\, \forall\, e \in E \cap Q_n \ : e \,\text{is closed}).$$

This latter probability converges to 0 when $\lambda \to \infty$ (for fixed n and $r > 0$). Hence, for large enough $\lambda < \infty$ (depending on n, r) we have $\mathbb{P}(B \cap C^c) \leq \epsilon/5$.

Similarly,

$$\mathbb{P}(D^c) = \mathbb{P}(\exists v \in V \cap Q_{6n} : v \text{ is closed})$$

converges to 0 when $p \uparrow 1$ (for fixed n and $r > 0$), hence, for large enough $p < 1$, we have $\mathbb{P}(D^c) \leq \epsilon/5$.

Regarding the event $\tilde{A} \cap D \cap F^c$, note that under the event $\tilde{A}$, we have

$$R(Q_{6n}) < n/2 < 3n/2.$$

Hence, by asymptotic essential connectedness (see Definition 4.4.2), we have that $\text{supp}(\Lambda_{Q_{3n}}) \neq \emptyset$ and, moreover, there exists a connected component Δ of $\text{supp}(\Lambda_{Q_{6n}})$ such that $\text{supp}(\Lambda_{Q_{3n}}) \subset \Delta \subset \text{supp}(\Lambda_{Q_{6n}})$. Therefore

$$\tilde{A} \cap D \cap F^c \subset \{\exists e \in E \cap Q_{6n} : e \text{ is closed}\}.$$

Clearly, for fixed n, r and independently of p,

$$\lim_{\lambda\to\infty} \mathbb{P}(\exists e \in E \cap Q_{6n} : e \text{ is closed}) = 0.$$

Hence, we can find $\lambda < \infty$ large enough (depending on n, r) such that $\mathbb{P}(\tilde{A} \cap D \cap F^c) \leq \epsilon/5$.

Since $\epsilon > 0$ was arbitrary, this concludes the proof of Lemma 4.4.11. □

By Lemmas 4.4.10 and 4.4.11, using Theorem 4.1.1, the process of n-good sites is stochastically dominated from below by a supercritical Bernoulli process for large enough $n < \infty, \lambda < \infty, p < 1$. Thus, we can make the process of n-good sites percolating. By Lemma 4.4.9, the connectivity graph $\mathcal{G}$ with these values of λ, p percolates, thus concluding the proof of Reformulation 4.2.2 and the equivalent Theorem 4.2.2.

4.4.5 Proof of Theorem 4.2.3

Using Reformulation 3, proving Theorem 4.2.3 amounts to showing that $\mathcal{G}$ percolates when $\lambda = 0$, $p < 1$ is sufficiently large and $r < \infty$ is sufficiently large. We thus assume throughout the rest of this subsection that $\lambda = 0$, r and p are the varying parameters and we still refer to $\mathcal{G}$ for the associated connectivity graph.

Say a street $e \in E$ is *hard-geometric-open* if its length is smaller than the connectivity threshold r: $|e| \leq r$. If not, say e is *hard-geometric-closed*.

Once again, we will use a renormalization argument. Since the development is very similar to the one exposed in the previous subsection, we only give details on which modifications should be brought to the proof of Theorem 4.2.2 to prove Theorem 4.2.3.

To this end, we consider now the following percolation model on the square lattice $\mathbb{Z}^2$. For $n \geq 1$, say a site $z \in \mathbb{Z}^2$ is n-good if it satisfies the following conditions:

(1) $R(Q_{6n}(nz)) < 6n$.

(2) $E \cap Q_n(nz) \neq \emptyset$, i.e. the square $Q_n(nz)$ contains a *full* street.

($\hat{3}$) $\exists\, e \in E \cap Q_n(nz)$ such that $|e| \leq r$. In other words, there exists a hard-geometric-open street that is fully included in the square $Q_n(nz)$.

(4) All crossroads in $Q_{6n}(nz)$ are open, in the sense of Definition 4.4.4.

($\hat{5}$) Every two hard-geometric-open streets $e, e' \in E \cap Q_{3n}(nz)$ (i.e. such that $|e| \leq r$ and $|e'| \leq r$) are connected by a path in $\mathcal{G} \cap Q_{6n}(nz)$.

We say a site $z \in \mathbb{Z}^2$ is n-bad if it is not n-good.

Note that this new definition of n-goodness is exactly the same as the one given in the proof of Theorem 4.2.2 but with conditions (3) and (5) being replaced by ($\hat{3}$) and ($\hat{5}$). In other words, openness is replaced by hard-geometric-openness.

Since we are now dealing with hard-geometric openness and all the other conditions are unchanged, the following is straightforward by adapting the proof of Lemma 4.4.9:

Lemma 4.4.12. *Percolation of the process of n-good sites implies percolation of the connectivity graph $\mathcal{G}$.*

In the same way, we get:

Lemma 4.4.13. *For $z \in \mathbb{Z}^2$, set $\xi_z := \mathbb{1}\{z \text{ is } n\text{-good}\}$. Then $(\xi_z)_{z\in\mathbb{Z}^2}$ is an 18-dependent random field for the supremum metric $\|\cdot\|_\infty$.*

Proof. It suffices to adapt the proof of Lemma 4.4.10 as follows.

Denote respectively by $A_z, B_z, \hat{C}_z, D_z, \hat{F}_z$ the events that the conditions (1), (2), ($\hat{3}$), (4) and ($\hat{5}$) in the definition of n-goodness hold for $z \in \mathbb{Z}^2$.

Note first that whenever $z \in \mathbb{Z}^2$, the indicators $\mathbb{1}\{A_z\}$, $\mathbb{1}\{B_z\}$, $\mathbb{1}\{\hat{C}_z\}$ are all Λ-measurable. Doing the exact same calculations as in the proof of Lemma 4.4.10, we arrive at dealing with the quantity

$$\mathbb{E}\left(\prod_{i=1}^{q} \xi'_{z_i}\right) = \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{1}\{\hat{C}_{z_i}\}\mathbb{E}\left(\prod_{i=1}^{q} \mathbb{1}\{D_{z_i} \cap \hat{F}_{z_i}\}\middle|\Lambda\right)\right].$$

Now, note that conditioned on Λ, for each $1 \leq i \leq q$, the event $D_{z_i} \cap \hat{F}_{z_i}$ only depends on the configuration of Y inside of the square $Q_{6n}(nz_i)$. We can thus proceed as in the aforementioned proof by using the complete independence of Y (recall that, given Λ, Y has the distribution of a Bernoulli point process). Finally, it is clear that $\mathbb{1}\{\hat{C}_z\}$ is a bounded deterministic function of $\Lambda_{Q_n(nz)}$ and that $\mathbb{E}\left(\mathbb{1}\{D_z \cap \hat{F}_z\}\middle|\Lambda\right)$ is a bounded deterministic function of $\Lambda_{Q_{6n}(nz)}$. This allows to proceed exactly as in the proof of Lemma 4.4.10 and conclude. □

Finally, for the hard-geometric model, we still have:

Lemma 4.4.14.

$$\lim_{n\uparrow\infty}\lim_{p\uparrow 1, r\uparrow\infty}\mathbb{P}(\mathbf{0}\ \textit{is}\ n\textit{-good}) = 1$$

Proof. Again, we shall prove that:

$$\lim_{n\uparrow\infty}\lim_{p\uparrow 1, r\uparrow\infty}\mathbb{P}(\mathbf{0}\,\text{is } n\text{-bad}) = 0.$$

Take any $\epsilon > 0$. We adapt the proof of Lemma 4.4.11 as follows. Denote respectively by $A, B, \hat{C}, D, \hat{F}$ the events that the conditions (1), (2), ($\hat{3}$), (4) and ($\hat{5}$) in the definition of n-goodness hold for $z = \mathbf{0}$. Denote also by $\tilde{A}$ the event that $R(Q_{6n}) < n/2$. As in the aforementioned proof, we have

$$\mathbb{P}(\mathbf{0}\,\text{is } n\text{-bad}) \leq \mathbb{P}(\tilde{A}^c) + \mathbb{P}(B^c) + \mathbb{P}(B \cap \hat{C}^c) + \mathbb{P}(D^c) + \mathbb{P}(\tilde{A} \cap D \cap \hat{F}^c).$$

In the above inequality, we deal with the first, second and fourth quantities as before and so we can fix n large enough such that $\mathbb{P}(\tilde{A}^c) \leq \epsilon/5$ and $\mathbb{P}(B^c) \leq \epsilon/5$. For such n, Q_n, Q_{3n} and Q_{6n} intersect almost surely zero or finitely many edges and vertices. We can then fix $p < 1$ large enough such that $\mathbb{P}(D^c) \leq \epsilon/5$.

Let's now deal with the quantity $\mathbb{P}(B \cap \hat{C}^c)$. We have:

$$\begin{aligned}\mathbb{P}(B \cap \hat{C}^c) &= \mathbb{P}(E \cap Q_n \neq \emptyset \text{ and } \forall\, e \in E \cap Q_n\ |e| > r) \\ &= \mathbb{E}\left(\mathbb{1}\{E \cap Q_n \neq \emptyset\} \prod_{e \in E \cap Q_n} \mathbb{1}\{|e| > r\}\right).\end{aligned}$$

Note first that on the event $\{E \cap Q_n \neq \emptyset\}$, the latter product is non-empty. Moreover, since $E \cap Q_n$ contains finitely many edges (recall that n is fixed) and since we have

$$\forall e \in E,\ \lim_{r\uparrow\infty} \mathbb{1}\{|e| > r\} = 0 \text{ a.s.},$$

by dominated convergence, we have that the latter expectation converges to 0 when $r \uparrow \infty$ (for fixed n). Therefore, $\lim_{r\uparrow\infty} \mathbb{P}(B \cap \hat{C}^c) = 0$ (for fixed n).

Regarding the event $\tilde{A} \cap D \cap \hat{F}^c$, we proceed as before and use asymptotic essential connectedness to get

$$\tilde{A} \cap D \cap \hat{F}^c \subset (\exists e \in E \cap Q_{6n} :\ e \text{ is hard-geometric-closed})$$

Clearly, for fixed n,

$$\begin{aligned}\lim_{r\uparrow\infty} \mathbb{P}(\exists e \in E \cap Q_{6n} :\ e \text{ is hard-geometric-closed}) &= \lim_{r\uparrow\infty} \mathbb{P}(\exists e \in E \cap Q_{6n} :\ |e| > r) \\ &= 0\end{aligned}$$

Hence, we can find $r < \infty$ large enough (depending on n) such that $\mathbb{P}(B \cap \hat{C}^c) \leq \epsilon/5$ and $\mathbb{P}(\tilde{A} \cap D \cap \hat{F}^c) \leq \epsilon/5$. Since $\epsilon > 0$ was arbitrary, this concludes the proof of Lemma 4.4.14.

□

By Lemmas 4.4.13 and 4.4.14, using Theorem 4.1.1, the process of n-good sites is stochastically dominated from below by a supercritical Bernoulli process for large enough $n, p < 1, r < \infty$. Thus, we can make the process of n-good sites percolating. By Lemma 4.4.12, the connectivity graph $\mathcal{G}$ with these values of p and r percolates, thus concluding the proof of Theorem 4.1.1.

4.4.6 Proof of Corollaries 4.2.2.1 and 4.2.3.1

We now prove the corollaries 4.2.2.1 and 4.2.3.1 as direct consequences of Theorems 4.2.2 and 4.2.3 and come back to the dimensionless parameters (p, U, H) in terms of which the Corollaries are stated.

Proof of Corollary 4.2.2.1

$\boxed{p^* < 1}$ By Theorem 4.2.2, we can find large enough $p \in (0,1)$ such that $P(p,U,H) > 0$ for any $H \in [0,\infty)$ and large enough $U < \infty$ depending on the chosen H. Choosing $H = 0$, we thus obtain the existence of some $p \in (0,1)$ such that $P(p,U,0) > 0$ for some $U < \infty$. Now, as noted in Section 4.2.1, the case $H = 0$ corresponds to PVT site percolation and the percolation probability $P(p,U,H) =: P_{PVT}(p)$ does not depend on U. Thus, $P_{PVT}(p) > 0$ for some $p < 1$, and so $p^* < 1$.

□

$\boxed{p^* > 0}$ It is known that the degree of all sites of a planar PVT is almost surely equal to 3, as a consequence of the fact that, almost surely, no 4 points of a homogeneous Poisson point process in $\mathbb{R}^2$ are cocyclic, see e.g. [12, Exercise 11.3.1]. Moreover, S, being a planar tessellation, is locally finite. This observation combined with the degree bound allows to use an adapted path-count argument, as follows.

First, denote by $\tilde{\mathcal{G}}_p =: \mathcal{G}_{p,0,0}$ the random graph arising from the PVT site percolation process with parameter p and note that percolation of $\tilde{\mathcal{G}}_p$ is equivalent to the percolation of $\mathcal{G}_{p,U,0}$ whenever $p \in [0,1]$ and $U \geq 0$ (indeed, the fact that $H = 0$ makes the percolation independent of the Cox points in our model). Hence:

$$P_{PVT}(p) = \mathbb{P}\left(\tilde{\mathcal{G}}_p \text{ has an infinite connected component}\right). \tag{4.4.5}$$

Denote by Φ the point process of crossroads (i.e. vertices of the PVT S) and denote by $\mathbb{P}^0$ its Palm probability. Since Y is a doubly stochastic Bernoulli point process supported by the crossroads, the conditional distribution of Y given S is the same under the stationary probability $\mathbb{P}$ and under the Palm probability $\mathbb{P}^0$. As a matter of fact, for every crossroad $v \in V$, we have: $\mathbb{P}^0(Y(\{v\}) > 0 \,|\, S) = \mathbb{P}(Y(\{v\}) > 0 \,|\, S) = p$. Moreover, conditionally to the realisation of the PVT S, the states of distinct crossroads (i.e. open or closed) remain independent.

Fix some $n \geq 1$. A *self-avoiding path* γ of length n starting from the typical crossroad $\mathbf{0}$ is a sequence of crossroads $\mathbf{0} = v_1, \ldots, v_n \in V$ with $v_i \neq v_j$ for $i \neq j$ and such that v_i and v_{i+1} are adjacent in S whenever $1 \leq i \leq n-1$. If the typical crossroad $\mathbf{0}$ belongs to an infinite connected component in $\tilde{\mathcal{G}}_p$ (which we denote by $\mathbf{0} \rightsquigarrow \infty$), there must exist such a path with a Bernoulli point present at all crossroads of the path. Denote this event by A_n.
Then we have:

$$\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) \leq \mathbb{P}^0(A_n).$$

Let SAP_n denote the set of self-avoiding paths of length n starting from the typical crossroad $\mathbf{0}$. By the union bound, we have:

$$\begin{aligned}\mathbb{P}^0(A_n) &\leq \mathbb{E}^0\left[\sum_{(v_1,\dots,v_n)=\gamma\in SAP_n} \mathbb{P}^0\left(\bigcap_{i=1}^n \{Y(\{v_i\})>0\}\,\middle|\, S\right)\right]\\ &= \mathbb{E}^0\left[\sum_{(v_1,\dots,v_n)=\gamma\in SAP_n} p^n\right]\\ &= \mathbb{E}^0\left[\#(SAP_n)p^n\right],\end{aligned}$$

where $\#(SAP_n)$ denotes the cardinal of SAP_n and where we have used the conditional independence of the states of the vertices as well as the conditional distribution of Y given S to get the first equality. Now, using the fact that $\forall v \in V, \deg v = 3$ a.s., we get that $\#(SAP_n) \leq 3 \times 2^{n-1}$. Hence:

$$\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) \leq 3 \times 2^{n-1}p^n = \frac{3}{2}(2p)^{n-1}.$$

When $p < 1/2$, the quantity in the right-hand side converges to 0 as $n \uparrow \infty$. Hence, for $p < 1/2$, we have $\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) = 0$.
To conclude that $\tilde{\mathcal{G}}_p$ does not percolate, we proceed as follows. For a crossroad $v \in V$, denote by $\{v \rightsquigarrow \infty\}$ the event that v belongs to an infinite connected component of the PVT site percolation graph $\tilde{\mathcal{G}}_p$. By Markov's inequality, we have:

$$\begin{aligned}\mathbb{P}\left(\tilde{\mathcal{G}}_p \text{ has an infinite connected component}\right) &=: \mathbb{P}(\exists v \in V \,:\, v \rightsquigarrow \infty)\\ &\leq \mathbb{E}\left[\#\{v \in V \,:\, v \rightsquigarrow \infty\}\right],\end{aligned}$$

and so, by (4.4.5), we get:

$$P_{PVT}(p) \leq \mathbb{E}\left[\#\{v \in V \,:\, v \rightsquigarrow \infty\}\right]. \tag{4.4.6}$$

Denote by $\lambda_0 := 2\lambda_S$ the intensity of the point process Φ of crossroads of S and fix some $p < 1/2$. By the Campbell-Little-Mecke-Matthes theorem (Theorem 1.2.10), we have:

$$\begin{aligned}\mathbb{E}\left[\#\{v \in V \,:\, v \rightsquigarrow \infty\}\right] &= \mathbb{E}\left[\int_{\mathbb{R}^2} \mathbb{1}\{x \rightsquigarrow \infty\}\Phi(dx)\right]\\ &= \lambda_0 \int_{\mathbb{R}^2} \mathbb{E}^0\left[\mathbb{1}\{\mathbf{0} \rightsquigarrow \infty\}\right] dx\\ &= \lambda_0 \int_{\mathbb{R}^2} \mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) dx\\ &= 0,\end{aligned}$$

where we have used the fact that $p < 1/2 \Rightarrow \mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) = 0$ to get the last equality. By (4.4.6), we have $P_{PVT}(p) = 0$ whenever $p < 1/2$ and thus $p^* \geq 1/2 > 0$.

□

Proof of Corollary 4.2.3.1

$\boxed{H_c < \infty}$ By Theorem 4.2.1, we can fix some $H' < \infty$ and $U' < \infty$ such that $P(1, U', H') = 0$. Note that since $P(1, 0, 0) = 1$, we have $H' > 0$. By monotonicity of the model, P is increasing in U and thus $P(1, 0, H') = 0$. Moreover, P is also decreasing in H and so $H \geq H' \Rightarrow P(1, 0, H) = 0$. Hence, by contraposition, $P(1, 0, H) > 0 \Rightarrow H < H'$. This latter condition is possible since $H' > 0$. Thus, we have the following inclusion of non-trivial sets:

$$\{H \geq 0 \,:\, P(1, 0, H) > 0\} \subseteq [0, H'[$$

and so, taking the suprema, we obtain $H_c \leq H' < \infty$.

□

$\boxed{H_c > 0}$ By Theorem 4.2.3, we can fix large enough $p < 1$ and small enough $H > 0$ such that $P(p, 0, H) > 0$. By monotonicity of the model, P is increasing in p and so $P(1, 0, H) > 0$. Hence $H_c \geq H > 0$.

□

4.5 Conclusion

In this chapter, we have studied the case of canyon shadowing, where only LOS communications along the edges of the PVT street system S supporting the network are possible. In such a case, the presence of relays at crossroads (given by the doubly stochastic Bernoulli process with parameter p) are essential to ensure connectivity between adjacent streets of S. Such a doubly stochastic model with restrictive connectivity conditions between network nodes has not been considered in the literature before.

Introducing relevant dimensionless scale-invariant parameters, namely the mean number of users U and the mean number of D2D hops H per typical street of S, we have identified several particular cases of our network percolation model, some of which have not been considered in the literature earlier. Furthermore, the percolation of the network connectivity graph $\mathcal{G} = \mathcal{G}_{p,U,H}$ exhibits phase transitions in U between several connectivity regimes corresponding to different ranges of network parameters (p, H):

- The permanently subcritical range, where percolation of $\mathcal{G}$ is possible no matter how small the density of users. This means that large-scale connectivity of the network is relay-limited, in so far as it can be solely ensured by relays present at crossroads.

- The critical range, where percolation of $\mathcal{G}$ is possible with a positive and finite density of users. Large-scale connectivity of the network is thus called relay-and-user-limited, as it crucially depends on both relays and network users in sufficient amounts.

- The permanently supercritical range, where $\mathcal{G}$ does not percolate, whatever large the density of users. Large-scale connectivity of the network is thus not possible with a finite density of users, regardless of all other network parameters as long as they are not trivial.

Estimating the frontiers between these different connectivity regimes is of critical importance for network design and economic planning. Indeed, these frontiers indicate to operators how large-scale connectivity of the network can be ensured (either solely by relays or by both users and relays) and what are the minimal network parameters allowing for long-range multihop D2D communications. We thus performed simulations of our model and, using efficient algorithms adapted to our problem, we proposed estimations of the critical network parameters and quantities delimiting the different connectivity regimes of the network.

Our results bring both qualitative (existence of the different connectivity regimes) and quantitative arguments (estimations of the frontiers between these connectivity regimes) to possible deployment scenarios of multihop D2D networks.

Chapter 5

Extending the canyon shadowing assumption: two variants

All models are approximations.
Essentially, all models are wrong, but some are useful.

George E.P. Box

In Chapter 4, we considered the case of canyon shadowing in our model, meaning that only line-of-sight (LOS) communications between network nodes located on the same street are possible. Such highly restrictive connectivity conditions have heavy consequences: relays must be installed by operators at crossroads of the streets to ensure connectivity between adjacent streets. What is more, we demonstrated that if less than $p^* \approx 71.3\%$ of crossroads are equipped with a relay, then good connectivity of the network is not possible, regardless of all the other network parameters. In urban scenarios, where street systems are particularly dense, deploying so many relays is not feasible, as it would induce considerable costs for operators.

Furthermore, the canyon shadowing assumption is too restrictive for modelling real-world networks, where connections between two nodes are possible even if they are not in LOS. In particular, scattering and reflections of the signal may allow for non-line-of-sight communications (NLOS) and, more importantly, communications between network nodes that are not located on the same street of the support of the network.

In this chapter, we thus propose two complementary approaches for relaxing the canyon shadowing assumption:

- On a theoretical perspective, we study the general case where $r' > 0$ in the D2D network model introduced in Section 3.3. Considering $r' > 0$ will allow to connect adjacent streets without needing a relay.

- On a more applied perspective, since relays are located at crossroads, we propose to see each crossroad as some geometric figure of non-null area, where relaying of the signal can be done either by a relay (Bernoulli point) or by a network user (Cox point).

We will see how both approaches complete each other and are related to a crucial network design consideration for an operator: relay deployment.

This chapter is partly based on the publication [88].

5.1 Introduction

5.1.1 The bottleneck of signal relaying at crossroads

In our modelling, we see that crossroads play a critical role for large-scale connectivity of the network. Indeed, the possibility of having long-range communications in the network via successive D2D hops crucially relies on signal relaying at crossroads and on establishing connectivity between adjacent streets of S. Connecting adjacent streets by crossing their intersections (crossroads) thus constitutes a *bottleneck* for large-scale connectivity of the network and one may even expect that if the probability of directly relaying the signal at a crossroad $v \in V$ is uniformly small for every $v \in V$, then the emergence of large connected components in the network graph (and so percolation) will be very unlikely.

In the canyon shadowing case studied in Chapter 4, signal relaying at crossroads could only be done by the presence of relays (Bernoulli points). We have proven that this induces a considerable proportion of relays needed to get good connectivity of the network (at least $p^* \approx 71.3\%$). However, deploying so many relays may not be feasible in practice, mostly for economic reasons. Moreover, the canyon shadowing assumption does not take into account additional connections between network nodes that are not located on the same street of the PVT street system S.

Noticing this previous bottleneck of signal relaying at crossroads and the restrictions arising from that the canyon shadowing assumption, the need for refined models appears clearly.

5.1.2 The need for refined models

In this chapter, we propose two approaches to relax the canyon shadowing assumption.

Our first approach is of theoretical nature. We study the general case of the model introduced in Section 3.3 by considering $r' > 0$, thus adding the possibility of direct connections between adjacent streets of the PVT S without requiring for a Bernoulli relay to be present at their crossroad. Our main interest here will be to see whether allowing for supplementary connections between adjacent streets will allow to explore the permanently subcritical range that existed in the canyon shadowing case (see the blue regions on Figures 4.1a and 4.1b). In particular, one may wonder whether percolation of $\mathcal{G}_{p,\lambda,r,r'}$ may happen under a non-trivial density of users $0 < \lambda < \infty$ when the relay proportion p is below the PVT site percolation threshold, i.e. $p < p^*$. We shall prove this is indeed the case.

In a more applied fashion, our second approach will consist in adopting new models for the geometry of crossroads. Indeed, the street system being a planar PVT, crossroads are punctual and hence have a null area. As mentioned in Section 3.3, the fact that crossroads are punctual, combined with the Cox modelling for network users, implies that the probability of finding a network user (i.e. a Cox point) at a crossroad is exactly 0.

In real-world street systems, however, crossroads have a non-null area and streets have a certain width. Taking into account a non-null area for crossroads in our model could allow for a positive probability of having network users located at crossroads and relaying the signal. This, in turn, would mean that less relays could be needed for ensuring good connectivity in the network. Keeping this in mind, we thus propose a more applied approach to consider crossroads with non-null area in our model. Resorting to numerical simulations, in the same fashion as in Section 4.2.4, we will see how this results in requiring much less relays for ensuring good connectivity of the network.

5.2 Percolation of the general network graph $\mathcal{G}_{p,\lambda,r,r'}$

5.2.1 Context

We now treat the general case where $r' > 0$ in the D2D network model presented in Section 3.3. The elementary bricks of the network model have been recalled in Section 4.1.2: the street system is given by a PVT S, the network users are given by the Cox process X^λ and the relays by the (doubly stochastic) Bernoulli process Y. Moreover, we adopt the same terminology as the one introduced in the aforementioned section, mainly:

- Edges of the PVT street system S are called streets. The set of streets of S is denoted by E.
- Vertices of the PVT street system S are called crossroads. The set of crossroads of S is denoted by V.
- A street segment $s \subset e$ is a topologically connected subset of a street $e \in E$.

The only difference with the global context introduced in Section 4.1.2 is that we now consider the case where the range for D2D connections due to scattering or reflections on crossroads is now positive: $r' > 0$. The network is thus modelled by the connectivity graph $\mathcal{G} = \mathcal{G}_{p,\lambda,r,r'}$, where, recall, two network nodes $Z_i, Z_j \in Z := X^\lambda \cup Y$ can now be connected by an edge if they satisfy either of the following sets of conditions:

$$\begin{cases} \exists e \in E \,:\, Z_i \in e \,\text{and}\, Z_j \in e \\ \|Z_i - Z_j\|_2 \leq r \end{cases} \tag{3.3.1}$$

OR

$$\begin{cases} Z_i \in X^\lambda, Z_j \in X^\lambda \\ e_{Z_i} \neq e_{Z_j} \\ \exists v \in V, : e_{Z_i} \cap e_{Z_j} =: \{v\} \\ \|Z_i - v\|_2 + \|Z_j - v\|_2 \leq r'. \end{cases} \tag{3.3.2}$$

Contrary to the canyon shadowing case ($r' = 0$) studied in Chapter 4, the positivity of r' may allow one to ensure connectivity between two adjacent streets without requiring a relay to be present at the crossroad of these streets.

5.2.2 Results

As has been done in Chapter 4, denote the percolation probability by

$$P(p, \lambda, r, r') := \mathbb{P}(\mathcal{G}_{p,\lambda,r,r'} \text{ has an unbounded connected component})$$

and observe P is increasing in p, in λ, in r and in r'. For given $p \geq 0$, $0 < r \leq \infty$ and $0 < r' \leq r$, consider the following critical value for the user intensity λ:

$$\lambda_c(p, r, r') := \inf\{\lambda \geq 0 \,:\, P(p, \lambda, r, r') > 0\},$$

with $\lambda_c(p, r, r') := \infty$ if $P(p, \lambda, r, r') = 0$ for all $\lambda \geq 0$.
As in Section 4.2, our goal is to show that there is a region (connected subset) of parameters (p, r, r') such that $0 < \lambda_c(p, r, r') < \infty$. Such a region is where the percolation of the connectivity graph $\mathcal{G}_{p,\lambda,r,r'}$ exhibits a non-trivial phase transition in the density of users. The following results ensure the existence of this region:

Theorem 5.2.1 (Existence of sub-critical intensities of users)**.** *Fix $r' > 0$. Then, for small enough value of $1 - (1-p)e^{-3\lambda r'} > 0$, we have $P(p, \lambda, r, r') = 0$ for all $r \geq r'$.*

In the above theorem, note that both p and λ need to be small enough to ensure that $P(p, \lambda, r, r') = 0$. The mutual variation of p and λ is captured by the constant deterministic expression $1 - (1-p)e^{-3\lambda r'}$.

Theorem 5.2.2 (Existence of super-critical intensities of users)**.** *For all $0 < r' \leq r \leq \infty$ and large enough $\lambda < \infty$ (depending on r and r'), we have that $P(0, \lambda, r, r') > 0$ and, consequently, by monotonicity of the model, $P(p, \lambda, r, r') > 0$ for all $p \in [0, 1]$.*

The proof of Theorem 5.2.1 can be done in a rather simple way by stochastically dominating $\mathcal{G}_{p,\lambda,r,r'}$ from above by a 1-dependent percolation model on S and using a path-count argument. The proof of Theorem 5.2.2 requires the use of renormalization techniques similar to the ones used in Section 4.4 and via appealing to the local dependence of S.

5.2.3 Proof of Theorem 5.2.1

Fixing $r' > 0$, proving Theorem 5.2.1 is equivalent to proving that $\mathcal{G} = \mathcal{G}_{p,\lambda,r,r'}$ does not percolate when p, λ are sufficiently small but positive for all $r \geq r'$. We will introduce a site percolation model on the PVT street system S defined in such a way that if it does percolate, $\mathcal{G}$ does not. Proving the percolation of the site percolation model will be done using the fact that it features a conditional 1-dependence, given the realisation of the street system S.

We first define our site percolation model on the PVT S by defining closed and open crossroads in the following way:

Definition 5.2.3 (Open/Closed crossroad)**.** A crossroad $v \in V$ is said to be *closed* if the following conditions are satisfied:

(1) $Y(\{v\}) = 0$, i.e. there isn't a fixed relay at v.

(2) $X^\lambda(\mathcal{B}(v,r') \cap E_v) = 0$ where $\mathcal{B}(v,r') := \{x \in \mathbb{R}^2 : \|x - v\|_2 \leq r'\}$ is the closed Euclidean ball with centre v and radius r' and $E_v := \{e \in E : v \in e\}$ is the set of streets of S that are incident to v. In other words, there are no Cox points on incident streets to v that are within distance r' of v.

We say that v is *open* if it is not closed.

Note that the definition is independent of the choice of r.
Closed crossroads somehow block the routing of the signal. Indeed, if v is closed, the connection cannot be ensured between two streets incident to v: LOS connections are blocked by the absence of a fixed relay and connections due to reflections or scattering are blocked because no Cox points are within a distance r' from v. This is formalised as follows:

Lemma 5.2.4. *Fix a set of parameters (λ, p, r', r) with $\lambda > 0$, $p > 0$, $0 < r' \leq r < \infty$. If the process of open crossroads does not percolate, then neither does $\mathcal{G}$.*

Proof. It is equivalent to show that if $\mathcal{G}$ percolates, then so does the process of open crossroads. Therefore, assume $\mathcal{G}$ percolates and denote by $\mathcal{C}$ an unbounded connected component of $\mathcal{G}$. We claim that, without loss of generality, we can think of $\mathcal{C}$ as being supported by S. Indeed, for any edge $Z_i \rightsquigarrow Z_j$ in $\mathcal{C}$, we are in one of the two following cases:

- Either Z_i and Z_j belong to the same edge of S, in that case the edge $Z_i \rightsquigarrow Z_j$ is indeed supported by S. For instance, this is the case for all the connections $Z_i \rightsquigarrow Z_{i+1}, 2 \leq i \leq 5$ on Figure 5.1.
- Or Z_i and Z_j belong to different edges of S. Then, by definition of the connection mechanism, e_{Z_i} and e_{Z_j} are incident to a common crossroad $v \in V$ and so we virtually divert the connection $Z_i \rightsquigarrow Z_j$ by considering the following path along S: $Z_i \rightsquigarrow v \rightsquigarrow Z_j$. An example is provided by the connection $Z_1 \rightsquigarrow Z_2$ on Figure 5.1: the direct orange dotted connection is diverted along S by the black thick lines.

Doing so is just a matter of perspective in considering the connections of $\mathcal{C}$ and does not change the nature of $\mathcal{C}$: in particular, this new version of $\mathcal{C}$ along the edges of S remains unbounded and connected.

Now, consider the set $\mathcal{C} \cap V =: \{v \in V : v \in \mathcal{C}\}$. Since $\mathcal{C}$ is unbounded, one must have $\#(\mathcal{C} \cap V) = \infty$. Now, we claim that each $v \in \mathcal{C} \cap V$ is an open crossroad, in the sense of Definition 5.2.3. Indeed, for each $v \in \mathcal{C} \cap V$, there is either a relay at v, so $Y(\{v\}) > 0$; or there is a connection in $\mathcal{C}$ between two edges of S that are incident to v and so $X^\lambda(\mathcal{B}(v,r')) \geq 2 > 0$. Moreover, any two consecutive crossroads in $\mathcal{C} \cap V$ are connected in S, since $\mathcal{C}$ is connected.
Hence $\mathcal{C}$ gives rise to an infinite connected component of open crossroads and so the process of open crossroads percolates. □

By Lemma 5.2.4, it suffices to prove that the process of open crossroads does not percolate for sufficiently small but positive $p > 0$ and $\lambda > 0$. This will be done by using the fact that the process of open crossroads features short-range dependencies only. More precisely, we have the following:

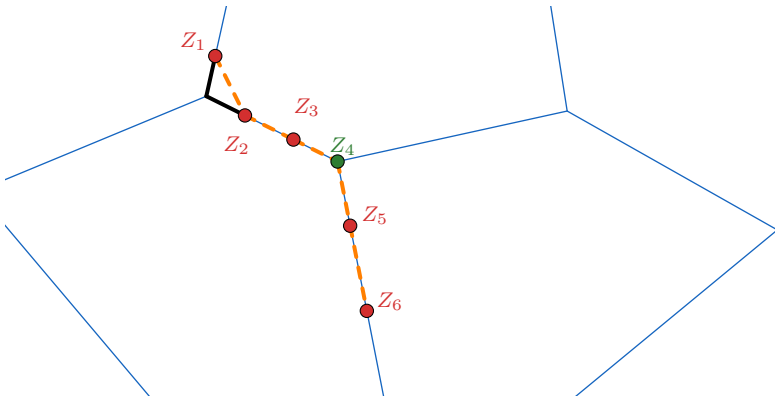


Figure 5.1 – Illustration of the model and of some part of the component $\mathcal{C}$. Blue lines represent the edges of S, red points represent Cox points, the green point represents a Bernoulli point and orange dotted segments represent connections in the connectivity graph $\mathcal{G}$. When considering $\mathcal{C}$, if a connection is already along S, it is left unchanged (e.g. $Z_2 \leftrightsquigarrow Z_3$). If a connection is between two points that are on different edges of S, we divert it by following the edges of S: for instance the connection $Z_1 \leftrightsquigarrow Z_2$ is diverted by the black thick lines.

Lemma 5.2.5. *For $v \in V$, denote $I_v := \mathbb{1}\{v \text{ is open}\}$. Denote by d the graph distance on the PVT street system S. Then, for every finite set of distinct crossroads $U = \{v_1, \dots, v_q\} \subset V$ with the property that $\forall i \neq j, d(v_i, v_j) > 1$, the random variables $(I_v)_{v \in U}$ are independent given Λ. In other words, conditioned on Λ (or, equivalently, on the street system S), the process of open crossroads is 1-dependent for the graph distance d of S.*

Proof. As a starting point, note that it is equivalent to show that the process of closed sites is 1-dependent. Therefore, set $\tilde{I}_v := \mathbb{1}\{v \text{ is closed}\}$ whenever $v \in V$ and let us show the property with I_v replaced by $\tilde{I}_v$. Let U be a finite subset of V as in Lemma 5.2.5. To prove the conditional independence of the random variables $(\tilde{I}_v)_{v \in U}$, it suffices to establish the following:

$$\mathbb{E}\left[\prod_{i=1}^{q} \tilde{I}_{v_i} \,\middle|\, \Lambda\right] = \prod_{i=1}^{q} \mathbb{E}\left[\tilde{I}_{v_i} \,|\, \Lambda\right]$$

We have:

$$\begin{aligned}\mathbb{E}\left[\prod_{i=1}^{q} \tilde{I}_{v_i} \,\middle|\, S\right] &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{Y(\{v_i\}) = 0\} \prod_{i=1}^{q} \mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \,\middle|\, \Lambda\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{Y(\{v_i\}) = 0\} \,\middle|\, \Lambda\right] \mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \,\middle|\, \Lambda\right],\end{aligned}$$

where we have used the fact that Y and X^λ are conditionally independent given Λ in the last line. Now, note that given Λ, $i \neq j \Rightarrow v_i \neq v_j$ and so the events $\mathbb{1}\{Y(\{v_i\}) = 0 \,:\, 1 \leq i \leq q\}$ are conditionally independent given Λ. Thus:

$$\mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{Y(\{v_i\}) = 0\} \,\middle|\, \Lambda\right] = \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{Y(\{v_i\}) = 0\} \,|\, \Lambda\right].$$

Moreover, given Λ, X^λ has the distribution of a Poisson point process with intensity measure $\lambda\Lambda$. Moreover, note that we have $d(v_i, v_j) > 1$ whenever $i \neq j$, so that every two distinct crossroads of U are at least at graph distance 2. Therefore, we have $E_{v_i} \cap E_{v_j} = \emptyset$ whenever $i \neq j$. As a matter of fact, given Λ, the sets $\mathcal{B}(v_i, r') \cap E_{v_i}, 1 \leq i \leq q$ are disjoint, and so the events $X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0$ are independent. Hence:

$$\mathbb{E}\left[\prod_{i=1}^{q} \mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \middle| \Lambda\right] = \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \,|\, \Lambda\right].$$

In all, we have:

$$\begin{aligned}
\mathbb{E}\left[\prod_{i=1}^{q} \tilde{I}_{v_i} \middle| \Lambda\right] &= \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{Y(\{v_i\}) = 0\} \,|\, \Lambda\right] \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \,|\, \Lambda\right] \\
&= \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{Y(\{v_i\}) = 0\} \,|\, \Lambda\right] \mathbb{E}\left[\mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \,|\, \Lambda\right] \\
&= \prod_{i=1}^{q} \mathbb{E}\left[\mathbb{1}\{Y(\{v_i\}) = 0\} \mathbb{1}\{X^\lambda(\mathcal{B}(v_i, r') \cap E_{v_i}) = 0\} \,|\, \Lambda\right] \\
&= \prod_{i=1}^{q} \mathbb{E}\left[\tilde{I}_{v_i} \,|\, \Lambda\right],
\end{aligned}$$

where we have again used the conditional independence of X^λ and Y given Λ to reunite both conditional expectations into a single one. This concludes the proof of Lemma 5.2.5. □

We finally prove that the process of good crossroads does not percolate does not percolate for small enough $p > 0$, $\lambda > 0$:

Lemma 5.2.6. *Fix $r' > 0$. For small small enough $p > 0$, $\lambda > 0$, the process of open crossroads does not percolate.*

Proof. Denote by Φ the point process of crossroads of S and denote by $\mathbb{P}^0$ its Palm probability. Note first that Lemma 5.2.5 and its proof remain true if we replace the stationary probability $\mathbb{P}$ by $\mathbb{P}^0$. Indeed, this is due to the fact that the joint distribution of X^λ and Y given the random support Λ remains the same under $\mathbb{P}$ and $\mathbb{P}^0$ (see [38, Example 13.1(a)] or [31, Section 5.2.3] for more details).
Next, we claim that given Λ, the Palm probability that any crossroad $v \in V$ is open is upper-bounded by $q = q_{p,\lambda,r'} := 1 - (1-p)e^{-3\lambda r'}$.
Indeed, we have:

$$\begin{aligned}
\mathbb{P}^0(v \text{ is open} \,|\, \Lambda) &= 1 - \mathbb{P}^0(v \text{ is closed} \,|\, \Lambda) \\
&= 1 - \mathbb{P}^0\left[Y(\{v\}) = 0, X^\lambda(\mathcal{B}(v, r') \cap E_v) = 0 \,|\, \Lambda\right] \\
&= 1 - \mathbb{E}^0\left[\mathbb{1}\{Y(\{v\}) = 0\} \mathbb{1}\{X^\lambda(\mathcal{B}(v, r') \cap E_v) = 0\} \,|\, \Lambda\right] \\
&= 1 - \mathbb{E}^0\left[\mathbb{1}\{Y(\{v\}) = 0\} \,|\, \Lambda\right] \mathbb{E}^0\left[\mathbb{1}\{\cap X^\lambda(\mathcal{B}(v, r') \cap E_v) = 0\} \,|\, \Lambda\right] \\
&= 1 - (1-p)\mathbb{E}^0(e^{-\lambda\Lambda(B(v,r') \cap E_v)} \,|\, \Lambda) \\
&= 1 - (1-p)e^{-\lambda\Lambda(B(v,r') \cap E_v)}, \qquad (5.2.1)
\end{aligned}$$

where we have used the conditional independence of X^λ and Y given Λ in the antepenultimate equality and the form of their conditional distributions in the penultimate equality. Now, since S is a PVT, it is known that for all $v \in V$, $\deg(v) = 3$ almost surely (e.g. see [108]). As a matter of fact, for each $v \in V$, E_v almost surely consists of three different edges of E incident to v. Denote them by $e_v^{(i)}, i = 1, 2, 3$ and recall their respective lengths are denoted by $|e_v^{(i)}|$. From this and (5.2.1), we thus get:

$$\begin{aligned}\mathbb{P}^0(v \text{ is open} \,|\, \Lambda) &= 1 - (1-p)e^{-\lambda\Lambda(B(v,r')\cap E_v)} \\ &= 1 - (1-p)\exp\left(-\lambda \sum_{i=1}^{3} \min(|e_v^{(i)}|, r')\right).\end{aligned}$$

Since we have $\min(|e_v^{(i)}|, r') \leq r'$ whenever $1 \leq i \leq 3$, we indeed have:

$$\mathbb{P}^0(v \text{ is open} \,|\, \Lambda) \leq 1 - (1-p)e^{-3\lambda r'} =: q,$$

as claimed. Moreover, it is clear that $\lim_{p\downarrow 0} \lim_{\lambda\downarrow 0} q_{p,\lambda,r'} = 0$ whenever $0 < r' < \infty$.

We now use a path-count argument for the (conditionally) 1-dependent process of open crossroads, very much as in the proof of Theorem 1.4.13, and proceed as follows. Fix some $n \geq 1$. A *self-avoiding path* γ of length n starting from the typical crossroad $\mathbf{0}$ is a sequence of crossroads $\mathbf{0} = v_1, \ldots, v_n \in V$ with $v_i \neq v_j$ for $i \neq j$ and such that $d(v_i, v_{i+1}) = 1$ for every $1 \leq i \leq n-1$, where, recall, d denotes the graph distance on the PVT street system S.
If the typical crossroad $\mathbf{0}$ belongs to an infinite connected component of open crossroads (which we denote by $\mathbf{0} \rightsquigarrow \infty$) , there must exist such a path where all crossroads are open. Denote this event by A_n.

Then we have

$$\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) \leq \mathbb{P}^0(A_n).$$

Let SAP_n denote the set of self-avoiding paths of length n starting from the typical crossroad $\mathbf{0}$. By the union bound, we have:

$$\mathbb{P}^0(A_n) \leq \mathbb{E}^0\left[\sum_{(v_1,\ldots,v_n)=\gamma\in SAP_n} \mathbb{P}^0\left(\bigcap_{i=1}^{n}\{v_i \text{ is open}\} \,\middle|\, \Lambda\right)\right].$$

Now, we can partition the set $\{v_1, \ldots, v_n\}$ into two sets U and W, such that any two distinct vertices in U, say $v_i \neq v_j$ are such that $d(v_i, v_j) > 1$. Since we have $\deg(v) = 3$ almost surely for every $v \in V$, U contains at least $\lfloor n/3 \rfloor$ crossroads of $\{v_1, \ldots, v_n\}$. Thus, we have:

$$\mathbb{P}^0\left(\bigcap_{i=1}^{n}\{v_i \text{ is open}\} \,\middle|\, \Lambda\right) \leq \mathbb{P}^0\left(\bigcap_{v\in U}\{v \text{ is open}\} \,\middle|\, \Lambda\right) = \prod_{v\in U} \mathbb{P}^0(v \text{ is open} \,|\, \Lambda) \leq q^{\lfloor n/3\rfloor},$$

where we have used Lemma 5.2.5 to get the equality and the previously found bound to get the last inequality. In all, we get:

$$\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) \leq \mathbb{E}^0\left[\sum_{(v_1,\ldots,v_n)=\gamma\in SAP_n} q^{\lfloor n/3\rfloor}\right] = \mathbb{E}^0\left[\#(SAP_n)q^{\lfloor n/3\rfloor}\right],$$

where $\#(SAP_n)$ denotes the cardinal of SAP_n. Again, since $\deg(v) = 3$ almost surely whenever $v \in V$, a crude bound for $\#(SAP_n)$ is $\#(SAP_n) \leq 3 \times 2^{n-1}$. Moreover, $\lfloor n/3 \rfloor > n/3 - 1$ implies $q^{\lfloor n/3 \rfloor} \leq q^{n/3-1}$. Hence:

$$\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) \leq 3 \times 2^{n-1} q^{n/3-1} = \frac{3}{2q}(2q^{1/3})^n$$

As $n \uparrow \infty$, the right-hand side converges to 0 when $q < 1/8$. Since we had $q \downarrow 0$ as $\lambda, p \downarrow 0$, we can find small enough $p > 0$ and $\lambda > 0$ such that $q < 1/8$ for all $0 < r' < \infty$ and independently of $r \in [r', \infty[$ (due to the fact that $q = q_{p,\lambda,r'}$ does not depend on r).
For such p, λ, we thus have $\mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) = 0$.

To finally conclude that the process of open crossroads does not percolate, we proceed as follows. For every crossroad $v \in V$, denote by $\{v \rightsquigarrow \infty\}$ the event that v belongs to an infinite connected component of open crossroads. Obviously, by Markov's inequality, we have:

$$\mathbb{P}(\text{open crossroads percolate}) =: \mathbb{P}\left(\exists v \in V \,:\, v \rightsquigarrow \infty\right) \leq \mathbb{E}\left[\#\{v \in V \,:\, v \rightsquigarrow \infty\}\right]. \quad (5.2.2)$$

Denote by $\lambda_0 := 2\lambda_S$ the intensity of the point process Φ of crossroads (i.e. vertices) of S and recall that $\mathbb{P}^0$ denotes the Palm probability of Φ. By the Campbell-Little-Mecke-Matthes theorem (Theorem 1.2.10), we have:

$$\begin{aligned}
\mathbb{E}\left[\#\{v \in V \,:\, v \rightsquigarrow \infty\}\right] &= \mathbb{E}\left[\int_{\mathbb{R}^2} \mathbb{1}\{x \rightsquigarrow \infty\}\Phi(dx)\right] \\
&= \lambda_0 \int_{\mathbb{R}^2} \mathbb{E}^0\left[\mathbb{1}\{\mathbf{0} \rightsquigarrow \infty\}\right] dx \\
&= \lambda_0 \int_{\mathbb{R}^2} \mathbb{P}^0(\mathbf{0} \rightsquigarrow \infty) dx \\
&= 0,
\end{aligned}$$

which, by (5.2.2), concludes the proof of Lemma 5.2.6. □

To finish the proof of Theorem 5.2.1, note that by Lemma 5.2.6, for fixed r', we can find small enough $\lambda, p > 0$ such that the process of open crossroads does not percolate for all $r \geq r'$. Hence, by Lemma 5.2.4, neither does $\mathcal{G}$. This concludes the proof of Theorem 5.2.1.

5.2.4 Proof of Theorem 5.2.2

We shall prove that for all $0 < r' \leq r \leq \infty$, in the absence of relays ($p = 0$), the model $\mathcal{G} = \mathcal{G}_{0,\lambda,r,r'}$ percolates with positive probability for large enough $\lambda < \infty$ depending on r and r'.
Contrary to the proof of Theorem 5.2.1, we will use a renormalization argument. The goal is to introduce a discrete percolation site model on the square lattice $\mathbb{Z}^2$, constructed in such a way that if the discrete model percolates, then so does $\mathcal{G}$. Then, proving that the discrete model percolates will be done via appealing to its local dependence.

To introduce the discrete model, we first introduce convenient notation and terminology.

First, we define concepts of openness and closedness for street segments (or possibly the whole streets themselves) exactly as in Section 4.4:

Definition 5.2.7 (Open/Closed street segment)**.** Let $e \in E$ be a street and let $\emptyset \neq s \subseteq e$ be a non-empty street segment of e or the whole street e itself.
Say s is *open* if either of the two following set of conditions are satisfied:

1. $|s| \leq r$

   OR

2. $\begin{cases} |s| > r \\ \forall c \subset s,\ (|c| = r\ , c \text{ connected and } c \text{ topologically closed}) \Rightarrow X^{\lambda}(c) \geq 1 \end{cases}$

Say s is *closed* if s is not open, i.e.:
$\begin{cases} |s| > r \\ \exists\, c \subset s, \text{ such that } |c| = r\ , c \text{ connected, } c \text{ topologically closed and } X^{\lambda}(c) = 0 \end{cases}$

Recall once again (see e.g. [108]) that if $v \in V$ is a crossroad of the PVT street system S, $\deg(v) = 3$ almost surely, i.e. v is almost surely the intersection of 3 streets. As a matter of fact, whenever $v \in V$, there almost surely exists a unique triple $(e_v^{(1)}; e_v^{(2)}; e_v^{(3)})$ of streets such that:

$$\{v\} = e_v^{(1)} \cap e_v^{(2)} \cap e_v^{(3)}$$

Then, for an integer $n \geq 1$ and a site $z \in \mathbb{Z}^2$, we introduce the following terminology, somehow analogous to the definition of openness in the proof of Theorem 5.2.1:

Definition 5.2.8 ((n,z)-open / (n,z)-closed crossroad)**.** Say a crossroad $v \in V$ is (n,z)-*open* if $v \in Q_{6n}(nz)$, $\deg(v) = 3$ and all of the three following conditions are satisfied:

- $X_v^{(1)} + X_v^{(2)} \leq r'$
- $X_v^{(2)} + X_v^{(3)} \leq r'$
- $X_v^{(3)} + X_v^{(1)} \leq r'$

where, for $i = 1, 2, 3$, $X_v^{(i)} := \inf\{\|x - v\|_2 \ : \ x \in X^{\lambda} \cap Q_{6n}(nz) \cap e_v^{(i)}\}$ denotes the (Euclidean) distance from v to its closest neighbour in the Cox process X^{λ}, in the square $Q_{6n}(nz)$ and along the edge $e_v^{(i)}$. We set $X_v^{(i)} := \infty$ if the (random) set $X^{\lambda} \cap Q_{6n}(nz) \cap e_v^{(i)}$ is empty.
As usual, we say that v is (n,z)-closed if it is not (n,z)-open.

We are now ready to introduce our discrete percolation model on the square lattice $z \in \mathbb{Z}^2$ as follows:

Definition 5.2.9 (n-good / n-bad site)**.** For $n \geq 1$, say a site $z \in \mathbb{Z}^2$ is n-good if the following conditions are satisfied:

(1) $R(Q_{6n}(nz)) < 6n$

(2) $E \cap Q_n(nz) \neq \emptyset$, i.e. the square $Q_n(nz)$ contains a *full* street (not just a street segment)

(3) There exists $e \in E \cap Q_n(nz)$ such that e is open, in the sense of Definition 5.2.7

(4) Every crossroad $v \in V \cap Q_{6n}(nz)$ is (n, z)-open, in the sense of Definition 5.2.8

(5) Every two open edges $e, e' \in E \cap Q_{3n}(nz)$ are connected by a path in $\mathcal{G} \cap Q_{6n}(nz)$

We say a site $z \in \mathbb{Z}^2$ is n-bad if it is not n-good.

The n-good sites have been defined so as to satisfy the following:

Lemma 5.2.10. *Percolation of the process of n-good sites implies percolation of the connectivity graph $\mathcal{G}_{0,\lambda,r,r'}$.*

Proof. Let $\mathcal{C}$ be an infinite connected component of n-good sites. Consider $z, z' \in \mathcal{C}$ such that $\|z - z'\|_1 = 1$. Without loss of generality, assume $z = (a, b)$ for some $a, b \in \mathbb{Z}$ and $z' = (a+1, b)$. By conditions (2) and (3) in the definition of n-goodness, we can find open streets $e \in E \cap Q_n(nz)$ and $e' \in E \cap Q_n(nz')$. Since

$$\begin{aligned} Q_n(nz) &= [na - n/2, na + n/2] \times [nb - n/2, nb + n/2]\,, \\ Q_n(nz') &= [na + n/2, na + 3n/2] \times [nb - n/2, nb + n/2]\,, \\ Q_{3n}(nz) &= [na - 3n/2, na + 3n/2] \times [nb - 3n/2, nb + 3n/2]\,, \end{aligned}$$

we have $Q_n(nz') \subset Q_{3n}(nz)$ and so $e' \in E \cap Q_n(nz')$ implies $e' \in E \cap Q_{3n}(nz)$. Since we also have $e \in E \cap Q_n(nz) \subset E \cap Q_{3n}(nz)$ and e, e' are both open, by conditions (4) and (5) in the definition of n-goodness, e and e' are connected by a path $\mathcal{L}$ in $\mathcal{G} \cap Q_{6n}(nz)$. Therefore, the path $\mathcal{L}$ also connects e and e' in $\mathcal{G}$, thus giving rise to an infinite connected component in $\mathcal{G}$. This concludes the proof of Lemma 5.2.10. □

By Lemma 5.2.10, proving Theorem 5.2.2 amounts to showing that, for all $0 < r' \leq r \leq \infty$, the process of n-good sites percolates (for some n) when $\lambda < \infty$ is sufficiently large but finite.

Proving the percolation of the process of n-good sites will be done via appealing to its local dependence. More precisely, the process of n-good sites is an 18-dependent percolation model on the square lattice:

Lemma 5.2.11. *For $z \in \mathbb{Z}^2$, set $\xi_z := \mathbb{1}\{z \text{ is } n\text{-good}\}$. Then $(\xi_z)_{z\in\mathbb{Z}^2}$ is an 18-dependent random field for the supremum metric $\|\cdot\|_\infty$.*

Proof. Consider a finite set of indices $\psi = \{z_1, \dots z_q\} \subset \mathbb{Z}^2$ such that $\forall i \neq j, \|z_i - z_j\|_\infty > 18$. We want to show that the random variables $(\xi_{z_i})_{1\leq i\leq q}$ are independent. Since the former random variables are indicator functions, it suffices to show that:

$$\mathbb{E}\left(\prod_{i=1}^{q} \xi_{z_i}\right) = \prod_{i=1}^{q} \mathbb{E}(\xi_{z_i})$$

Denote respectively by A_z, B_z, C_z, D_z, F_z the events that the conditions (1), (2), (3), (4), (5) in the definition of n-goodness hold for $z \in \mathbb{Z}^2$. We thus have:

$$\forall z \in \mathbb{Z}^2,\ \xi_z = \mathbb{1}\{A_z\}\mathbb{1}\{B_z\}\mathbb{1}\{C_z\}\mathbb{1}\{D_z\}\mathbb{1}\{F_z\}.$$

Note first that whenever $z \in \mathbb{Z}^2$, the indicators $\mathbb{1}\{A_z\}$ and $\mathbb{1}\{B_z\}$ are Λ-measurable. Thus, we have :

$$\begin{aligned}\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\right) &= \mathbb{E}\left[\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\middle|\Lambda\right)\right] \\ &= \mathbb{E}\left[\prod_{i=1}^{q}\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{E}\left(\prod_{i=1}^{q}\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\mathbb{1}\{D_{z_i}\}\middle|\Lambda\right)\right]. \qquad (5.2.3)\end{aligned}$$

Moreover, for all $1 \le i \le q$, we have:

$$\begin{aligned}\mathbb{1}\{D_{z_i}\} &= \prod_{\substack{v\in \\ V\cap Q_{6n}(nz_i)}} \mathbb{1}\{v \text{ is } (n,z_i)\text{-open}\} \\ &= \prod_{\substack{v\in \\ V\cap Q_{6n}(nz_i)}} \mathbb{1}\{\deg(v)=3\}\mathbb{1}\{v\in Q_{6n}(nz_i)\}\mathbb{1}\{G_v\}, \qquad (5.2.4)\end{aligned}$$

where, for $v \in V$, G_v denotes the following event:

$$G_v := \left(X_v^{(1)} + X_v^{(2)} \le r'\right) \cap \left(X_v^{(2)} + X_v^{(3)} \le r'\right) \cap \left(X_v^{(3)} + X_v^{(1)} \le r'\right).$$

The first two indicators appearing in the product in the right-hand side of (5.2.4) are Λ-measurable. Using this, the right-hand side of (5.2.3) is equal to:

$$\mathbb{E}\left[\prod_{i=1}^{q}\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\prod_{\substack{v\in \\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{\deg(v)=3\}\mathbb{E}\left(\prod_{i=1}^{q}\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\prod_{\substack{v\in \\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{G_v\}\middle|\Lambda\right)\right]$$

Consider some $1 \le i \le q$. It is clear from the definitions that, conditioned on Λ, the events C_{z_i} and D_{z_i} only depend on the configuration of X^λ and Y inside of the square $Q_{6n}(nz_i)$. The same is true for the event G_v whenever $v \in V \cap Q_{6n}(nz_i)$.
Moreover ψ satisfies $\forall i \ne j, \|z_i - z_j\|_\infty > 18$, then we have $\forall i \ne j, \|nz_i - nz_j\|_\infty > 18n$. As a matter of fact, the squares $\{Q_{6n}(nz_i)) \,:\, 1 \le i \le q\}$ are disjoint, i.e.

$$\forall i \ne j, Q_{6n}(nz_i) \cap Q_{6n}(nz_j) = \emptyset.$$

Since the respective conditional distribution of X^λ and Y given Λ are those of a Poisson point process and of a Bernoulli point process, by complete independence of Poisson and Bernoulli processes, we have :

$$\mathbb{E}\left(\prod_{i=1}^{q}\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{G_v\}\middle|\Lambda\right)=\prod_{i=1}^{q}\mathbb{E}\left[\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{G_v\}\middle|\Lambda\right]$$
$$=\prod_{i=1}^{q}f(\Lambda_{Q_{6n}(nz_i)}),$$

where $f(\Lambda_{Q_{6n}(x)}):=\mathbb{E}\left[\mathbb{1}\{C_x\cap F_x\}\prod_{\substack{v\in\\ V\cap Q_{6n}(x)}}\mathbb{1}\{G_v\}\middle|\Lambda\right]$ is a bounded measurable deterministic function of $\Lambda_{Q_{6n}(x)}$.

In all, we thus get:

$$\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\right)=\mathbb{E}\left[\prod_{i=1}^{q}\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}f(\Lambda_{Q_{6n}(nz_i)})\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{\deg(v)=3\}\right]. \tag{5.2.5}$$

For each $1\leq i\leq q$, we have that $\prod_{v\in V\cap Q_{6n}(nz_i)}\mathbb{1}\{\deg(v)=3\}(=1$ almost surely) and $\mathbb{1}\{B_{z_i}\}=:\mathbb{1}\{\exists e\in Q_{6n}(nz_i)\,:\,e$ is open$\}$ obviously are bounded measurable deterministic functions of $\Lambda_{Q_{6n}(nz_i)}$. Since f also is, we can write:

$$\mathbb{1}\{B_{z_i}\}\,f(\Lambda_{Q_{6n}(nz_i)})\prod_{v\in V\cap Q_{6n}(nz_i)}\mathbb{1}\{\deg(v)=3\}:=g(\Lambda_{Q_{6n}(nz_i)})$$

as one bounded measurable deterministic function g of the configuration of Λ inside the square $Q_{6n}(nz_i)$. Hence (5.2.5) yields:

$$\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\right)=\mathbb{E}\left[\prod_{i=1}^{q}\mathbb{1}\{A_{z_i}\}g(\Lambda_{Q_{6n}(nz_i)})\right]$$
$$=\mathbb{E}\left[\prod_{i=1}^{q}\mathbb{1}\{R(Q_{6n}(nz_i))<6n\}g(\Lambda_{Q_{6n}(nz_i)})\right]. \tag{5.2.6}$$

The set $\varphi:=\{nz_1,\ldots,nz_q\}\subset\mathbb{R}^2$ is a finite subset of $\mathbb{R}^2$ satisfying:

$$\forall i\neq j,\|nz_i-nz_j\|_\infty>18n$$

Since the infinite norm is always upper bounded by the Euclidean norm, we have $\forall i\neq j,\|nz_i-nz_j\|_2>18n$, and so φ satisfies:

$$\forall x\in\varphi,\,\mathrm{dist}_2(x,\varphi\setminus\{x\})>18n=3\times 6n.$$

Hence, we can apply condition (3) in the definition of stabilization (Definition 4.4.1) with n replaced by $6n$. As a consequence, we get that the random variables appearing in the right-hand side of (5.2.6) above are independent, and so:

$$\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\right)=\prod_{i=1}^{q}\mathbb{E}\left[\mathbb{1}\{R(Q_{6n}(nz_i))<6n\}g(\Lambda_{Q_{6n}(nz_i)})\right].$$

To finish the proof of Lemma 5.2.11, it only remains to show that for all $1\le i\le q$, we have

$$\mathbb{E}\left[\mathbb{1}\{R(Q_{6n}(nz_i))<6n\}g(\Lambda_{Q_{6n}(nz_i)})\right]=\mathbb{E}(\xi_{z_i}).$$

This is indeed the case:

$$\begin{aligned}
&\mathbb{E}\left[\mathbb{1}\{R(Q_{6n}(nz_i))<6n\}g(\Lambda_{Q_{6n}(nz_i)})\right]\\
&=\mathbb{E}\left[\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{E}\left[\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{G_v\}\middle|\Lambda\right]\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{\deg(v)=3\}\right]\\
&=\mathbb{E}\left[\mathbb{E}\left[\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{G_v\}\mathbb{1}\{\deg(v)=3\}\middle|\Lambda\right]\right]\\
&=\mathbb{E}\left[\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{1}\{C_{z_i}\cap F_{z_i}\}\prod_{\substack{v\in\\ V\cap Q_{6n}(nz_i)}}\mathbb{1}\{G_v\}\mathbb{1}\{\deg(v)=3\}\right]\\
&=\mathbb{E}\left[\mathbb{1}\{A_{z_i}\}\mathbb{1}\{B_{z_i}\}\mathbb{1}\{C_{z_i}\}\mathbb{1}\{F_{z_i}\}\mathbb{1}\{D_{z_i}\}\right]=:\mathbb{E}(\xi_{z_i}),
\end{aligned}$$

where we have used Λ-measurability of the indicators $\mathbb{1}\{A_{z_i}\}$, $\mathbb{1}\{B_{z_i}\}$ and $\mathbb{1}\{\deg(v)=3\}$ in the third line to put everything back into one single conditional expectation.

In all, we have $\mathbb{E}\left(\prod_{i=1}^{q}\xi_{z_i}\right)=\prod_{i=1}^{q}\mathbb{E}(\xi_{z_i})$ as needed. This concludes the proof of Lemma 5.2.11.

□

Now, for fixed $0<r'\le r\le\infty$, we prove that the probability for an arbitrary site (which by stationarity can be chosen to be the origin $\mathbf{0}\in\mathbb{Z}^2$), to be n-good can be made arbitrarily large when first taking some large enough finite n and then finite large enough λ, as stated in the following Lemma.

Lemma 5.2.12. *For all r,r' with $0<r'\le r\le\infty$, we have*

$$\lim_{n\uparrow\infty}\lim_{\lambda\uparrow\infty}\mathbb{P}(\mathbf{0}\textit{ is }n\textit{-good})=1$$

Proof. We shall rather prove that for all $0<r'\le r\le\infty$, we have

$$\lim_{n\uparrow\infty}\lim_{\lambda\uparrow\infty}\mathbb{P}(\mathbf{0}\text{ is }n\text{-bad})=0.$$

Fix such r and r'. Take any $\epsilon>0$. Denote respectively by A,B,C,D,F the events that the conditions (1), (2), (3), (4), (5) in the definition of n-goodness hold for $z=\mathbf{0}$.

Denote also by $\tilde{A}$ the event that $R(Q_{6n}) < n/2$. Note that $\tilde{A} \subset A$ and thus we have:

$$\begin{aligned}\mathbb{P}(\mathbf{0}\,\text{is } n\text{-bad}) &= \mathbb{P}(A^c \cup B^c \cup C^c \cup D^c \cup F^c)\\ &\leq \mathbb{P}(\tilde{A}^c \cup B^c \cup C^c \cup D^c \cup F^c)\\ &\leq \mathbb{P}(\tilde{A}^c) + \mathbb{P}(B^c) + \mathbb{P}(B \cap C^c) + \mathbb{P}(B \cap D^c) + \mathbb{P}(\tilde{A} \cap D \cap F^c).\end{aligned}$$

First, partitioning the square Q_{6n} into $12^2 = 144$ subsquares $(Q_i)_{1\leq i\leq 144}$ of side length $n/2$, we get:

$$\begin{aligned}\mathbb{P}(\tilde{A}^c) &= \mathbb{P}(R(Q_{6n}) \geq n/2)\\ &= \mathbb{P}\left(\bigcup_{i=1}^{144} \{R(Q_i) \geq n/2\}\right)\\ &\leq 144\,\mathbb{P}(R(Q_{n/2}) \geq n/2) \qquad \text{by stationarity of the } R\text{'s.}\end{aligned}$$

Therefore, by condition (2) of Definition 4.4.1, we get $\lim_{n\uparrow\infty} \mathbb{P}(\tilde{A}^c) = 0$. Also

$$\mathbb{P}(B^c) = \mathbb{P}(E \cap Q_n = \emptyset)$$

and thus $\lim_{n\uparrow\infty} \mathbb{P}(B^c) = 0$. Fix n large enough such that $\mathbb{P}(\tilde{A}^c) \leq \epsilon/5$ and $\mathbb{P}(B^c) \leq \epsilon/5$. For such n, Q_n, Q_{3n} and Q_{6n} intersect almost surely zero or a finite number of edges and vertices.

To deal with the quantity $\mathbb{P}(B \cap C^c)$, we simply write:

$$\mathbb{P}(B \cap C^c) = \mathbb{P}(E \cap Q_n \neq \emptyset \,\text{and}\, \forall\, e \in E \cap Q_n \ : e \,\text{is closed}).$$

For fixed n and $r > 0$, this latter probability now only depends on λ and converges to 0 when $\lambda \to \infty$. Hence, for large enough $\lambda < \infty$ (depending on n, r) we have $\mathbb{P}(B \cap C^c) \leq \epsilon/5$.

Let us now deal with the fourth quantity. We have:

$$\begin{aligned}\mathbb{P}(B \cap D^c) &= \mathbb{P}(E \cap Q_n \neq \emptyset \,\text{and}\, \exists\, v \in V \cap Q_{6n} \ : v \,\text{is } (n,\mathbf{0})\text{-closed})\\ &= \mathbb{E}\left[\mathbb{P}\left(\{E \cap Q_n \neq \emptyset\} \cap \bigcup_{v\in V\cap Q_{6n}} \{v \,\text{is } (n,\mathbf{0})\text{-closed}\} \,\middle|\, \Lambda\right)\right]\\ &= \mathbb{E}\left[\mathbb{1}\{E \cap Q_n \neq \emptyset\}\mathbb{P}\left(\bigcup_{v\in V\cap Q_{6n}} \{v \,\text{is } (n,\mathbf{0})\text{-closed}\} \,\middle|\, \Lambda\right)\right]\\ &\leq \mathbb{E}\left[\mathbb{1}\{E \cap Q_n \neq \emptyset\} \sum_{v\in V\cap Q_{6n}} \mathbb{P}\left(v \,\text{is } (n,\mathbf{0})\text{-closed} \,|\, \Lambda\right)\right],\end{aligned}$$

where we have used the Λ-measurability of the event $\{E \cap Q_n \neq \emptyset\}$ in the third equality and a union-bound on the conditional probability to get the last inequality. Going to the limit, we thus get:

$$\lim_{\lambda\uparrow\infty}\mathbb{P}(B\cap D^c)\leq\lim_{\lambda\uparrow\infty}\mathbb{E}\left[\mathbb{1}\{E\cap Q_n\neq\emptyset\}\sum_{v\in V\cap Q_{6n}}\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)\right]. \tag{5.2.7}$$

Now, note that we have

$$\mathbb{E}\left[\mathbb{1}\{E\cap Q_n\neq\emptyset\}\sum_{v\in V\cap Q_{6n}}\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)\right]\leq\mathbb{E}(\#(V\cap Q_{6n})).$$

Since S is a PVT generated by a homogeneous Poisson point process of intensity λ_S, the intensity of the point process of vertices of S is given by $\lambda_0=2\lambda_S$ (see [108]), so that $\mathbb{E}(\#(V\cap Q_{6n}))=e^{-2\lambda_S\times 36n^2}<\infty$ (recall that n is now fixed). Hence, by the dominated convergence theorem, (5.2.7) yields:

$$\begin{aligned}\lim_{\lambda\uparrow\infty}\mathbb{P}(B\cap D^c)&\leq\mathbb{E}\left[\lim_{\lambda\uparrow\infty}\mathbb{1}\{E\cap Q_n\neq\emptyset\}\sum_{v\in V\cap Q_{6n}}\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)\right]\\&=\mathbb{E}\left[\mathbb{1}\{E\cap Q_n\neq\emptyset\}\lim_{\lambda\uparrow\infty}\sum_{v\in V\cap Q_{6n}}\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)\right]\\&=\mathbb{E}\left[\mathbb{1}\{E\cap Q_n\neq\emptyset\}\sum_{v\in V\cap Q_{6n}}\lim_{\lambda\uparrow\infty}\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)\right],\end{aligned} \tag{5.2.8}$$

where we have used the facts that the event $\{E\cap Q_n\neq\emptyset\}$ is independent of λ and that, since n has been fixed, Q_{6n} almost surely intersects a finite number of vertices and so $\sum_{v\in V\cap Q_{6n}}$ is a finite sum almost surely.

Now, whenever $v\in V\cap Q_{6n}$, we have:

$$\begin{aligned}\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)&=1-\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-open}\,|\,\Lambda\right)\\&=1-\mathbb{P}\left[(v\in V\cap Q_{6n})\cap(\deg(v)=3)\cap G_v\,|\,\Lambda\right],\end{aligned}$$

where, recall, G_v, denotes the event:

$$G_v:=\left(X_v^{(1)}+X_v^{(2)}\leq r'\right)\cap\left(X_v^{(2)}+X_v^{(3)}\leq r'\right)\cap\left(X_v^{(3)}+X_v^{(1)}\leq r'\right).$$

Since the first two events appearing in the conditional probability above are Λ-measurable, we get:

$$\mathbb{P}\left(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda\right)=1-\mathbb{1}\{v\in V\cap Q_{6n}\}\mathbb{1}\{\deg(v)=3\}\mathbb{P}(G_v\,|\,\Lambda).$$

In the expression above, only $\mathbb{P}(G_v\,|\,\Lambda)$ depends on λ. Therefore, we have:

$$\lim_{\lambda\uparrow\infty}\mathbb{P}(v\text{ is }(n,\mathbf{0})\text{-closed}\,|\,\Lambda)=1-\mathbb{1}\{v\in V\cap Q_{6n}\}\mathbb{1}\{\deg(v)=3\}\lim_{\lambda\uparrow\infty}\mathbb{P}(G_v\,|\,\Lambda). \tag{5.2.9}$$

Now, from the definition of G_v, we have:

$$\mathbb{P}(G_v \,|\, \Lambda) \geq \mathbb{P}\left(X_v^{(1)} \leq r'/2,\, X_v^{(2)} \leq r'/2,\, X_v^{(3)} \leq r'/2 \,\middle|\, \Lambda\right).$$

Given Λ, X^λ has the distribution of a Poisson point process with mean measure $\lambda\Lambda$, and the three events $\{X_v^{(1)} \leq r'/2\}$, $\{X_v^{(1)} \leq r'/2\}$ and $\{X_v^{(3)} \leq r'/2\}$ depend on disjoint sets of edges of S, so they are conditionally independent given Λ. This yields:

$$\mathbb{P}(G_v \,|\, \Lambda) \geq \mathbb{P}\left(X_v^{(1)} \leq r'/2 \,\middle|\, \Lambda\right) \mathbb{P}\left(X_v^{(2)} \leq r'/2 \,\middle|\, \Lambda\right) \mathbb{P}\left(X_v^{(3)} \leq r'/2 \,\middle|\, \Lambda\right).$$

Obviously, as $\lambda \uparrow \infty$, the edges of S will eventually be densely filled with Cox points. Thus, the three conditional probabilities appearing in the right-hand side above all converge to 1 almost surely. Hence $\lim_{\lambda\uparrow\infty} \mathbb{P}(G_v \,|\, \Lambda) = 1$ almost surely. From (5.2.9), we thus get:

$$\lim_{\lambda\uparrow\infty} \mathbb{P}(v \text{ is } (n,\mathbf{0})\text{-closed} \,|\, \Lambda) = 1 - \mathbb{1}\{v \in V \cap Q_{6n}\}\mathbb{1}\{\deg(v) = 3\} \quad \text{a.s.}$$

Reporting this in (5.2.8), we obtain:

$$\lim_{\lambda\uparrow\infty} \mathbb{P}(B \cap D^c) \leq \mathbb{E}\left[\mathbb{1}\{E \cap Q_n \neq \emptyset\} \sum_{v \in V \cap Q_{6n}} (1 - \mathbb{1}\{v \in V \cap Q_{6n}\}\mathbb{1}\{\deg(v) = 3\})\right].$$

Now we claim that the sum in the right-hand side above is almost surely equal to $\mathbb{1}\{V \cap Q_{6n} = \emptyset\}$. Indeed, since n has been fixed, Q_{6n} contains almost surely zero or a finite number of vertices v. We are therefore in one of the two following cases:

- Either $V \cap Q_{6n} = \emptyset$, in which case the sum is empty and hence equal to $1 = \mathbb{1}\{V \cap Q_{6n} = \emptyset\}$.
- Or $V \cap Q_{6n} \neq \emptyset$, in which case the sum is not empty. But then, whenever $v \in V \cap Q_{6n}$, we obviously have $\mathbb{1}\{v \in V \cap Q_{6n}\} = 1$ and, since S is a PVT, $\deg(v) = 3$ almost surely. Hence the whole sum is equal to $0 = \mathbb{1}\{V \cap Q_{6n} = \emptyset\}$.

Thus, we finally get:

$$\lim_{\lambda\uparrow\infty} \mathbb{P}(B \cap D^c) \leq \mathbb{E}\left[\mathbb{1}\{E \cap Q_n \neq \emptyset\}\mathbb{1}\{V \cap Q_{6n} = \emptyset\}\right] = 0.$$

Therefore, for fixed n and $0 < r' \leq r \leq \infty$, $\lim_{\lambda\uparrow\infty} \mathbb{P}(B \cap D^c) = 0$ and so we can find $\lambda < \infty$ large enough (depending on n, r and r') such that $\mathbb{P}(B \cap D^c) \leq \epsilon/5$.

Regarding the event $\tilde{A} \cap D \cap F^c$, note that under the event $\tilde{A}$, we have

$$R(Q_{6n}) < n/2 < 3n/2.$$

Hence, by asymptotic essential connectedness (see Definition 4.4.2), we have that $\mathrm{supp}(\Lambda_{Q_{3n}}) \neq \emptyset$ and, moreover, there exists a connected component Δ of $\mathrm{supp}(\Lambda_{Q_{6n}})$ such that $\mathrm{supp}(\Lambda_{Q_{3n}}) \subset \Delta \subset \mathrm{supp}(\Lambda_{Q_{6n}})$. Therefore

$$\tilde{A} \cap D \cap F^c \subset \{\exists e \in E \cap Q_{6n} : e \text{ is closed}\}.$$

Clearly, for fixed n, r and independently of r',

$$\lim_{\lambda\uparrow\infty} \mathbb{P}(\exists e \in E \cap Q_{6n} : e \text{ is closed }) = 0.$$

Hence, we can find $\lambda < \infty$ large enough (depending on n, r) such that $\mathbb{P}(\tilde{A} \cap D \cap F^c) \leq \epsilon/5$.

As a matter of fact, we have that for sufficiently large $n < \infty$ and sufficiently large $\lambda < \infty$ (depending on n,r and r'), $\mathbb{P}(\mathbf{0}$ is n-bad$) \leq \epsilon$. Since $\epsilon > 0$ was arbitrary, we have indeed proven that

$$\lim_{n\uparrow\infty} \lim_{\lambda\uparrow\infty} \mathbb{P}(\mathbf{0} \text{ is } n\text{-bad}) = 0.$$

So, whenever $0 < r' \leq r \leq \infty$, $\lim_{n\uparrow\infty} \lim_{\lambda\uparrow\infty} \mathbb{P}(\mathbf{0}$ is n-good$) = 1$, as required. □

By Lemmas 5.2.11 and 5.2.12, using Theorem 4.1.1, for all $0 < r' \leq r \leq \infty$, the process of n-good sites is stochastically dominated from below by a supercritical Bernoulli process for large enough $n < \infty$ and, $\lambda < \infty$ (depending on r and r'). Thus, we can make the process of n-good sites percolating. By Lemma 5.2.10, the connectivity graph $\mathcal{G} = \mathcal{G}_{0,\lambda,r,r'}$ with these values of λ, r, r' percolates, thus concluding the proof of Theorem 5.2.2.

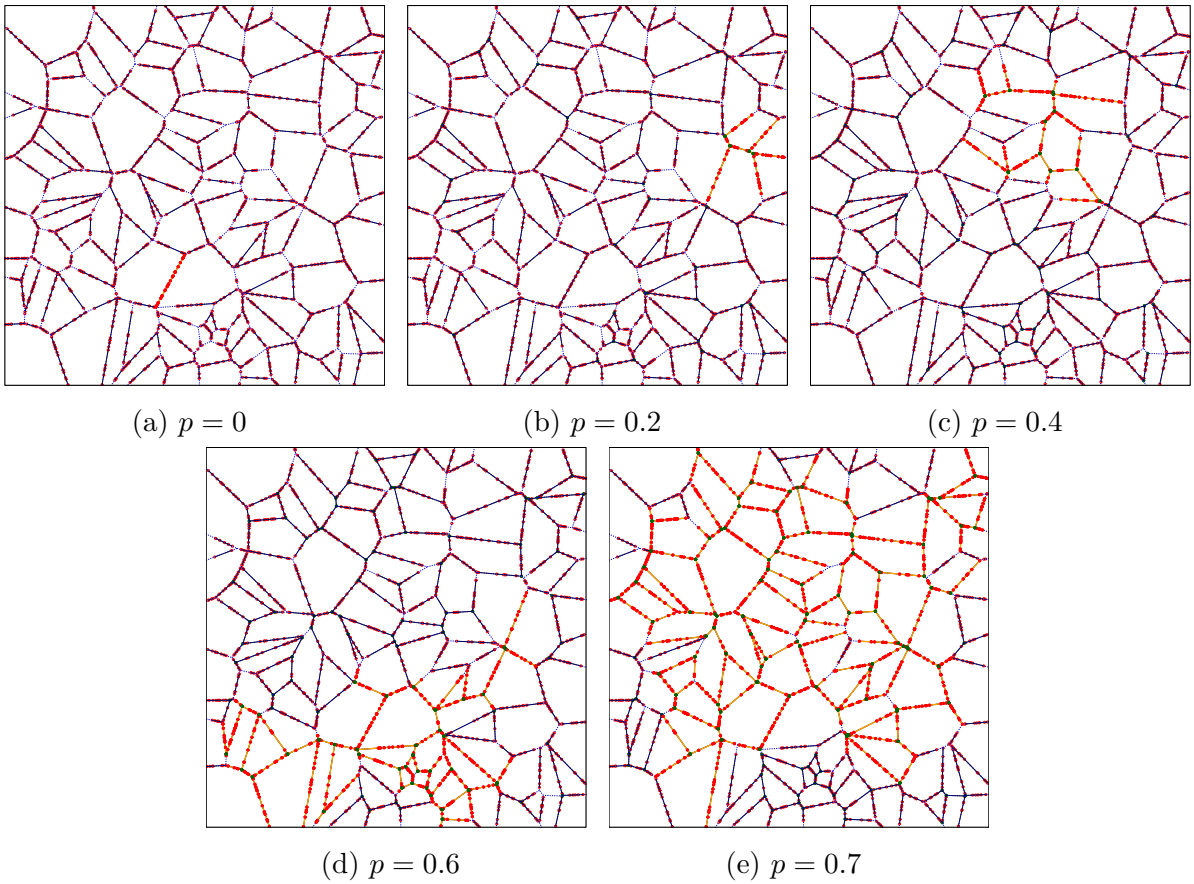
5.2.5 Discussion of the case of real-world networks: $r' \ll r$

As has been noted in Section 3.3.3, it makes sense to assume that $r' \leq r$, due to the fact that the loss of power due to scattering or reflection comes in addition to the path loss. Depending on the nature of physical obstacles encountered by the signal and of the environment of the network, one may even expect that r' is considerably smaller than r. Typically, in dense urban environments, one would expect $r' \approx r/10$.

On a theoretical perspective, Theorem 5.2.2 ensures that percolation of the general network graph $\mathcal{G}_{p,\lambda,r,r'}$ is theoretically possible under a finite density of users $\lambda_c(p, r) < \infty$ for all relay proportion $p \in [0, 1]$. This means that, theoretically, large-scale connectivity of the network can always be achieved under a finite density of D2D users for all values of p. In particular, taking $p = 0$, one can thus obtain the existence of a finite density of D2D users $\lambda_c(0, r) < \infty$ above which communications at large-scale are possible even if no relays are present in the network.

However, in practice, one may expect repercussions from the fact that $r' \ll r$. Indeed, r' being considerably smaller than r, most of the connections in the network may be due to LOS connections along the streets, rather than to scattering or reflection at the crossroads of adjacent streets. Thus, a very small proportion of relays p may need to be compensated by a very large density of users λ to ensure percolation of the network connectivity graph. Or, the other way around, for a fixed D2D user density λ, the fact that $r' \ll r$ may result in a large critical relay proportion p for ensuring good connectivity of the network. What if this critical relay proportion p cannot be achieved in practice?

Let us try to get an idea of the situation from a few pictures. Consider for instance the case of an urban scenario with a PVT street density $L_A = 20\,\mathrm{km}^{-1}$, corresponding to a dense urban street system. Also, consider a D2D LOS range corresponding to a 4G context $r = 100\,\mathrm{m}$ and take $r' = r/10 = 10\,\mathrm{m}$ for the range of D2D connections due to scattering or reflections at crossroads. Finally, consider a D2D user density $\lambda = 100\,\mathrm{km}^{-1}$, meaning that there is 1 D2D user every 10 metres in average, which corresponds to a



(a) $p = 0$ (b) $p = 0.2$ (c) $p = 0.4$

(d) $p = 0.6$ (e) $p = 0.7$

Figure 5.2 – Five simulations of the network connectivity graph $\mathcal{G}_{p,\lambda,r,r/10}$ for a LOS D2D range $r = 100\,\text{m}$, a D2D user density $\lambda = 100\,\text{km}^{-1}$, a PVT street density $L_A = 20\,\text{km}^{-1}$ and varying values of the relay proportion p. As before, blue dashed lines represent the PVT street, red points represent the D2D users (Cox points) and green points represent the relays (Bernoulli points). In the largest connected component of the window, network nodes are filled with their respective colours and edges are highlighted in orange. Edges of the smaller connected components are highlighted in black. In the five simulations, the realisations of the PVT street system S and of the Cox process of network users X^λ are the same.

rather dense population density.
Figure 5.2 shows examples of simulations of the connectivity graph $\mathcal{G}_{p,\lambda,r,r/10}$ for the previous values of the D2D range r and of the D2D user density λ, with a relay proportion $p \in \{0, 0.2, 0.4, 0.6, 0.7\}$. On this figure, we can see that the fact that $r' \ll r$, with a fixed D2D user density λ, forces the relay proportion p to be quite large for noticing the emergence of good connectivity of the network. Even for $p = 0.7$, a substantially large part of the simulation window remains out of the largest connected component of the network. This is due to the fact that since $r' \ll r$, connections due to scattering and/or reflections at crossroads only account for a small proportion of the total connections in the network.

From the previous figure and remarks, we note that while it is theoretically possible to have good connectivity of the network even in the absence of relays ($p = 0$), in practice

and for a fixed user density λ, the minimal relay proportion needed for large-scale connectivity of the network may still be quite large compared to the PVT site percolation threshold $p^* \approx 0.713$ estimated in Chapter 4. In other words, still for a fixed D2D user density λ, the fact that $r' \ll r$ may result in a very small influence of r' on the minimal relay proportion ensuring percolation of the network connectivity graph. Therefore, one may need to think of supplementary models that allow for more realistic estimations of this minimal relay proportion.

5.3 Revising the geometry of crossroads: supplementary geometric models

5.3.1 General methodology

As stated earlier, in the general D2D network model introduced in Section 3.3, the probability of finding a Cox point at a given crossroad of the street system is exactly 0. In practice, this means that network users cannot be on crossroads and play the role of relays for relaying the signal, thereby ensuring connectivity between adjacent streets. Moreover, with the chosen modelling, the only way to connect two adjacent streets $e \neq e'$ intersecting at the crossroad v without requiring the presence of a relay at v is to have a connection between two Cox points due to scattering or reflection of the signal at v, as described in the connectivity mechanism given by (3.3.2).
However, if we were to consider a refined geometry for the street system and especially a non-null area for the crossroads while keeping the same spreading of network users (Cox process X^λ) and relays (doubly stochastic Bernoulli process Y), the problem might change. In particular, there may be a positive probability to find a network user at a crossroad, able to relay the signal between streets incident to that crossroad. As a matter of fact, relays may be compensated by the presence of users at crossroads ensuring signal relaying, thus resulting in much less relays needed to ensure good connectivity of the network! Since relays represent a necessary investment for operators, the question of finding how many of them are needed to ensure connectivity of the network at large scale is of critical importance.
Therefore, we need supplementary geometric models to take into account the possibility that D2D users (Cox points) themselves can act as relays. Such models will in turn allow to estimate more precisely how much relays have to be deployed by an operator to ensure good connectivity of the network. Given this need of refined geometric models and of more precise estimations, we propose to proceed as follows:

- First, we "enlarge" the street system by giving a positive width $l > 0$, expressed in metres (m) to each street of the PVT S.
- Doing so, the intersections between adjacent streets are no longer punctual.
- We then imagine some geometric figure of non-null area (which, by abuse of terminology, will also be called crossroad) where a D2D user (Cox point) can serve as a relay between all the streets incident to that area.

- We estimate the probability that signal relaying between adjacent streets can effectively be ensured, by a relay *or* by a D2D user.
- We estimate the minimal proportion of relays needed to ensure good connectivity of the network.
- Relating the obtained estimates to our previous estimates in the canyon shadowing case, we show that large-scale connectivity of the D2D network in our model actually requires rather few relays in practice.

5.3.2 Geometric models for crossroads

Giving a non-null area to each crossroad

The street system S being a PVT, the degree of each crossroad is almost surely equal to 3 [108]. Consider the typical crossroad O of S, in other words the typical point of the point process of vertices of S, as defined in Section 1.3.3. This typical crossroad can be represented by the segments e_1, e_2 and e_3 of Figure 5.3, with (not necessarily equal) angles $(\alpha, \beta, \delta = 2\pi - \alpha - \beta)$ being random variables.

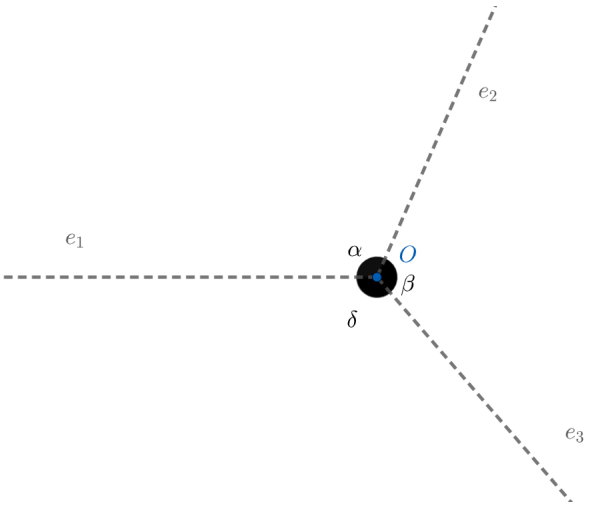


Figure 5.3 – Typical crossroad in the original street model. e_1, e_2, e_3 denote the streets of S incident to the typical crossroad O (there are almost surely 3 such streets).

To define a non-null surface for crossroads, we proceed as follows: we keep the PVT street system S but we think of "enlarging" S by giving a width $l > 0$, expressed in metres (m), to each street e of S. The typical crossroad (and, in fact, every crossroad of S) illustrated above now rather looks like in Figure 5.4. In this figure, the grey dashed segments e_1, e_2 and e_3 represent the original streets and the new enlarged streets are represented by the blue thick lines.

By considering the enlarged streets given by the blue lines, we imagine that the original punctual crossroad at O (given by the central blue point on Figures 5.3 and 5.4) becomes some geometric figure of non-null area. Such a geometric figure, by abuse of terminology, will also be called *crossroad* from now onwards and will be considered to be the region where either a relay (Bernoulli point) *or* a D2D user (Cox point) can also serve as a

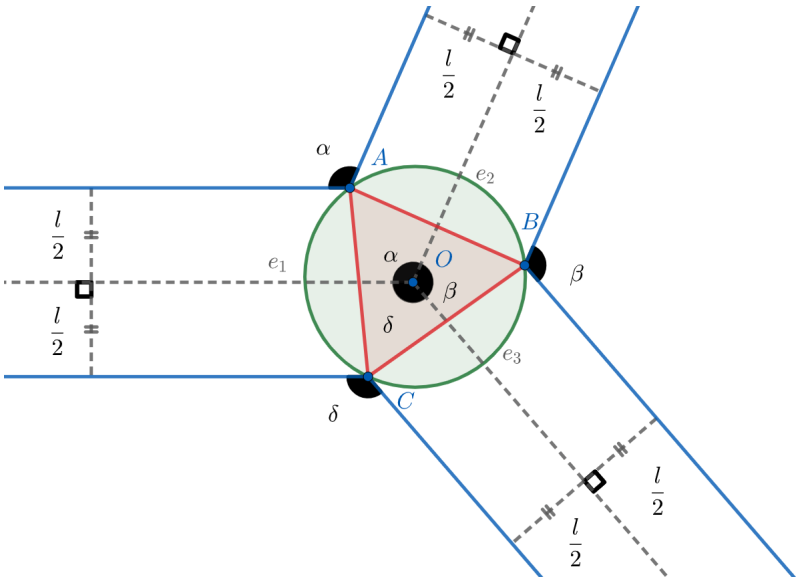


Figure 5.4 – Typical crossroad in the refined street model where all streets of S are prescribed with a width $l > 0$. The dashed segments e_1, e_2 and e_3 illustrate how the typical crossroad looked like in the original realisation of the PVT S (i.e. when streets have a null width $l = 0$). The blue lines represent the enlarged streets, of width $l > 0$. We consider two cases where the useful crossroad surface for relaying the signal between any two of the three enlarged streets respectively corresponds to the triangle filled with red or to the circle filled with green. Keep in mind that O designates the original (punctual) typical crossroad and is *not* the centre of the green circle (in fact, one can show that O is rather the centre of the innercircle of the red triangle ABC).

relay between two of the three streets. Indeed, since crossroads now have a non-null area when streets have been enlarged, there is a positive probability to find a Cox point inside the area of the geometric figure delimiting a crossroad. The surface of that geometric figure can thus be interpreted as the *useful* surface for signal relaying, in the sense that it delimits the region of the plane where relaying between adjacent streets can effectively be ensured.

Main assumptions of the geometric model for crossroads

For convenience or the sake of simplicity, we make the following assumptions in this new geometric model for crossroads.

First, in order to simplify the derivation of the formula for the (typical) crossroad surface, we propose two scenarios. The first one corresponds to the case where the crossroad is defined to be the triangle delimited by the intersections of the streets limits, see the triangle ABC filled with red on Figure 5.4. We denote the surface of this triangle by $\mathscr{S} := \mathscr{S}(l, \alpha, \beta)$. In the second scenario, the resulting surface is the circumcircle of the previous triangle, see the triangle filled with green on Figure 5.4. We denote by $\mathscr{S}' := \mathscr{S}'(l, \alpha, \beta)$ the surface of this circumcircle. Compared to what would be the surface of a crossroad in a real-world street system, the triangle case corresponds to a rather small useful surface for signal relaying, while the circle case corresponds to a rather large surface. Also, it will be much more likely to find a Cox point in the circle than in the

triangle, due to the fact that the ratio between the area of a triangle and the one of its circumcircle is considerably small[1].

Secondly, we assume that relaying can be done on the whole crossroad surface within the LOS range r, as if we were in the canyon shadowing case. In other words, we consider that a Bernoulli point or a Cox point located *anywhere* in the useful surface of the crossroad can be connected by a D2D link to a network node within distance r and located on one of the streets incident to the crossroad.

Finally, we assume that when the relaying between two adjacent streets of S is done by a D2D user (i.e. a Cox point), it is done *directly*, in the sense that it only involves only one network node located on the crossroad. We thus do not consider cases where the crossroad is occupied by two or more Cox points intervening in relaying the signal between the neighbouring streets.

5.3.3 Crossroad surface and occupation probability

We now compute the useful surface of a crossroad in the circle and in the triangle case, and deduce the probability that signal relaying at the typical crossroad can effectively be ensured.

The triangle case

We first begin by deriving the triangle surface $\mathscr{S} = \mathscr{S}(l, \alpha, \beta)$. Recall this corresponds to the case when the useful surface for signal relaying at the crossroad is given by the triangle delimited by the intersections of the enlarged streets (see the triangle ABC filled with red on Figure 5.4). Note first that the surface of the whole triangle ABC can be decomposed as the sum of the surfaces of the three triangles OAB, OBC and OCA, in other words:

$$\mathscr{S} = S_{OAB} + S_{OBC} + S_{OCA}.$$

By symmetry arguments, it is easy to see that the line (OA) is the bissector of the angle spanned by the segments e_1 and e_2, see Figure 5.5. By standard trigonometry arguments, we thus have:

$$OA = \frac{l}{2\sin\frac{\alpha}{2}} \quad ; \quad OB = \frac{l}{2\sin\frac{\beta}{2}}.$$

[1]By standard Euclidean geometry arguments, it is relatively easy to show that such a ratio is maximal when the underlying triangle is equilateral. In such a case, the (maximal) ratio is then equal to $\frac{3\sqrt{3}}{4\pi} \approx 0.413$. Therefore, the area of the circumcircle of a (non-degenerate) triangle is always more than $1/0.413 \approx 2.421$ times larger than the area of the corresponding triangle.

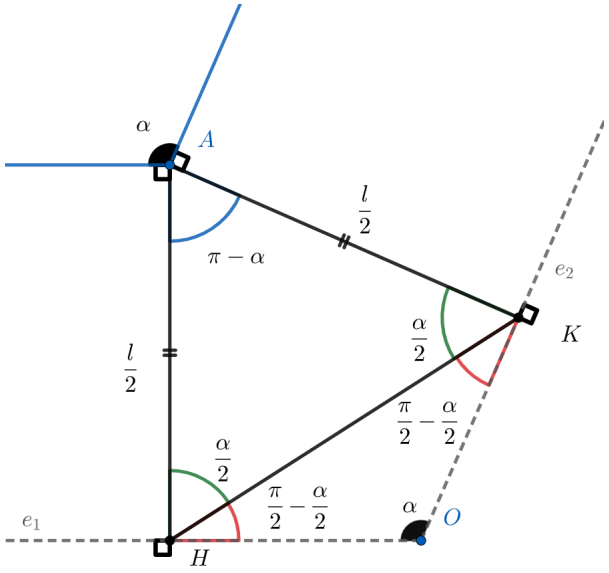


Figure 5.5 – Zoom of Figure 5.4. Consider the orthogonal projection H (respectively K) of A on the street e_1 (respectively on the street e_2). By construction, the triangle AHK is isoceles in A and A belongs to the bissector of the segment $[H, K]$. This, with the measure of the blue angle allows to find the measure of the green angles and then the measure of the red angles. As a consequence, the triangle OHK is isoceles in O. Thus O also belongs to the bissector of the segment $[H, K]$. Hence, the line (AO) is the bissector of the segment $[H, K]$ and, since OHK is an isoceles triangle, is also the bissector of the angle spanned by the streets e_1 and e_2.

Hence, by the cross-product formula, the area of the triangle OAB is given by:

$$\begin{aligned} S_{OAB} &= \frac{1}{2} OA\, OB\, \sin\left(\widehat{OA, OB}\right) \\ &= \frac{l^2}{8 \sin\frac{\alpha}{2} \sin\frac{\beta}{2}} \sin\left(\frac{\alpha+\beta}{2}\right) \\ &= \frac{l^2}{8}\left[\frac{\sin\frac{\alpha}{2}\cos\frac{\beta}{2} + \sin\frac{\beta}{2}\cos\frac{\alpha}{2}}{\sin\frac{\alpha}{2}\sin\frac{\beta}{2}}\right] \\ &= \frac{l^2}{8}\left[\cot\frac{\beta}{2} + \cot\frac{\alpha}{2}\right]. \end{aligned}$$

In the same way, by a circular permutation of α, β and δ, we get the following analogous formulae for the surfaces of triangles OBC and OCA:

$$\begin{aligned} S_{OBC} &= \frac{l^2}{8}\left[\cot\frac{\delta}{2} + \cot\frac{\beta}{2}\right] \\ S_{OCA} &= \frac{l^2}{8}\left[\cot\frac{\alpha}{2} + \cot\frac{\delta}{2}\right]. \end{aligned}$$

Noting that $\delta = 2\pi - \alpha - \beta$, we therefore get the surface $\mathscr{S}(l, \alpha, \beta)$ of the triangle ABC:

$$\begin{aligned}\mathscr{S}(l,\alpha,\beta) &= S_{OAB} + S_{OBC} + S_{OCA} \\ &= \frac{l^2}{4}\left[\cot\frac{\alpha}{2} + \cot\frac{\beta}{2} + \cot\frac{\delta}{2}\right] \\ &= \frac{l^2}{4}\left[\cot\frac{\alpha}{2} + \cot\frac{\beta}{2} - \cot\left(\frac{\alpha+\beta}{2}\right)\right]. \end{aligned} \tag{5.3.1}$$

The circle case

Now that we have derived a formula for the surface $\mathscr{S} = \mathscr{S}(l,\alpha,\beta)$ of the triangle ABC, we can derive a formula for the surface $\mathscr{S}' = \mathscr{S}'(l,\alpha,\beta)$ of its circumcircle. First, note that $\mathscr{S}' = \pi R^2$, where R is the circumcircle radius of the triangle ABC, given by:

$$R = \frac{AB\,BC\,AC}{4\mathscr{S}(l,\alpha,\beta)} \tag{5.3.2}$$

We already have a formula for $\mathscr{S}$ and it only remains to compute the side lengths AB, AC and BC of the triangle. This is easily done using the law of cosines. For instance, we have:

$$\begin{aligned} AB^2 &= OA^2 + OB^2 - 2\,OA\,OB\,\cos\left(\widehat{OA,OB}\right) \\ &= \frac{l^2}{4\sin^2\frac{\alpha}{2}} + \frac{l^2}{4\sin^2\frac{\beta}{2}} - 2\,\frac{l^2}{4\sin\frac{\alpha}{2}\sin\frac{\beta}{2}}\cos\left(\frac{\alpha+\beta}{2}\right) \\ &= \frac{l^2}{4}\left[\frac{1}{\sin^2\frac{\alpha}{2}} + \frac{1}{\sin^2\frac{\beta}{2}} - 2\,\frac{\cos\left(\frac{\alpha+\beta}{2}\right)}{\sin\frac{\alpha}{2}\sin\frac{\beta}{2}}\right]. \end{aligned}$$

Using the classical trigonometric identity $1 = \cos^2 x + \sin^2 x$ on the numerators of the first two fractions and expanding the cosine in the numerator of the last fraction yields:

$$\begin{aligned} AB^2 &= \frac{l^2}{4}\left[1 + \cot^2\frac{\alpha}{2} + 1 + \cot^2\frac{\beta}{2} - 2\cot\frac{\alpha}{2}\cot\frac{\beta}{2} + 2\right] \\ &= \frac{l^2}{4}\left[\left(\cot\frac{\alpha}{2} - \cot\frac{\beta}{2}\right)^2 + 4\right]. \end{aligned} \tag{5.3.3}$$

Again, by circularly permutating α, β and δ, and using the fact that $\delta = 2\pi - \alpha - \beta$, we get similar formulae for BC^2 and AC^2:

$$BC^2 = \frac{l^2}{4}\left[\left(\cot\frac{\beta}{2} - \cot\frac{\delta}{2}\right)^2 + 4\right] = \frac{l^2}{4}\left[\left(\cot\frac{\beta}{2} + \cot\left(\frac{\alpha+\beta}{2}\right)\right)^2 + 4\right] \tag{5.3.4}$$

$$AC^2 = \frac{l^2}{4}\left[\left(\cot\frac{\delta}{2} - \cot\frac{\alpha}{2}\right)^2 + 4\right] = \frac{l^2}{4}\left[\left(\cot\left(\frac{\alpha+\beta}{2}\right) + \cot\frac{\alpha}{2}\right)^2 + 4\right]. \tag{5.3.5}$$

By (5.3.2), the surface $\mathscr{S}'(l,\alpha,\beta)$ of the circumcircle is given by:

$$\mathscr{S}'(l,\alpha,\beta) = \pi R^2 = \pi \frac{AB^2\, AC^2\, BC^2}{16\mathscr{S}(l,\alpha,\beta)^2},$$

where $\mathscr{S}(l,\alpha,\beta)$, AB, BC and AC are respectively given by equations (5.3.1), (5.3.3), (5.3.4) and (5.3.5). That is to say, putting everything back together:

$$\mathscr{S}'(l,\alpha,\beta) = \frac{\pi l^2}{64} \frac{\left[\left(\cot\frac{\alpha}{2}-\cot\frac{\beta}{2}\right)^2+4\right]\left[\left(\cot\frac{\beta}{2}+\cot\left(\frac{\alpha+\beta}{2}\right)\right)^2+4\right]\left[\left(\cot\left(\frac{\alpha+\beta}{2}\right)+\cot\frac{\alpha}{2}\right)^2+4\right]}{\left[\cot\frac{\alpha}{2}+\cot\frac{\beta}{2}-\cot\left(\frac{\alpha+\beta}{2}\right)\right]^2}. \tag{5.3.6}$$

Crossroad occupation probability

We now propose a formula for the probability that the signal can effectively be relayed between two streets that are incident to the typical crossroad of the enlarged street system. Note that this is equivalent to compute the probability that the useful surface of the typical crossroad (triangle or circumcircle, according to which scenario is considered) is occupied either by a physical relay (Bernoulli point) or by a D2D user. We call this probability the *crossroad occupation probability*.

Denote by $\mathscr{A} \coloneqq \mathscr{A}(l,\alpha,\beta)$ the useful surface of the crossroad:

$$\mathscr{A}(l,\alpha,\beta) = \begin{cases} \mathscr{S}(l,\alpha,\beta) & \text{in the triangle case, see (5.3.1).} \\ \mathscr{S}'(l,\alpha,\beta) & \text{in the circle case, see (5.3.6).} \end{cases} \tag{5.3.7}$$

Keeping the same models as before for the users (Cox process) and for the relays (Bernoulli process) and using the conditional independence of the users and relays given the street system, we pose:

$$\mathbb{P}(\text{typical crossroad is occupied}) \coloneqq 1 - \mathbb{P}(\text{no user in } \mathscr{A}) \times \mathbb{P}(\text{no relay in } \mathscr{A}).$$

Now, note the initial linear intensity λ of users in the Cox process when streets did not have any width can be thought of to be equivalent to a surface density $\lambda' = \lambda/l$ in our model where streets have width $l > 0$. Indeed, in the initial model where streets had null width $l = 0$, the mean number of users per street e is a Poisson random variable with parameter $\lambda|e|$. Now that streets are enlarged and have a width $l > 0$, the accessible surface for a user on a given street e is now equal to $l|e|$ and so, if λ' denotes the new surface intensity of users for an equivalent Cox process on the enlarged street system, the number of users on e would now be a Poisson random variable with parameter $\lambda' l|e|$. The two models can be seen as equivalent if $\lambda' l|e| = \lambda|e|$, i.e. $\lambda' = \lambda/l$.
We keep the same models as before and assume that the users are now distributed on the enlarged streets as a Cox process with surface intensity $\lambda' = \lambda/l$ while the relays are still distributed like as a Bernoulli point process on the crossroads with parameter p. Thus, we get:

$$\mathbb{P}(\text{no user in } \mathscr{A}) \coloneqq \mathbb{E}\left[e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)}\right] \tag{5.3.8}$$

$$\mathbb{P}(\text{no relay in } \mathscr{A}) \coloneqq 1 - p. \tag{5.3.9}$$

Therefore, we have:

$$\mathbb{P}(\text{typical crossroad is occupied}) = 1 - (1-p)\mathbb{E}\left[e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)}\right].$$

Recall that α and β are random variables and denote the angles spanned by the streets of the typical crossroad of the PVT S. To compute the expectation appearing in the above equation, we use the joint probability density for the random vector (α, β). This joint probability density has been computed by Muche [104]:

$$f(\alpha,\beta) = -\frac{8}{3\pi}\sin\alpha\sin\beta\sin(\alpha+\beta)\mathbb{1}\{0<\alpha<\pi\}\mathbb{1}\{\pi-\alpha<\beta<\pi\}.$$

Therefore, the crossroad occupation probability is given by:

$$\begin{aligned} F(\lambda,p,l) &\coloneqq \mathbb{P}(\text{typical crossroad is occupied}) \\ &= 1-(1-p)\int_{\mathbb{R}^2} e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)} f(\alpha,\beta)d\alpha d\beta \\ &= 1+\frac{8}{3\pi}(1-p)\int_{\alpha=0}^{\pi}\int_{\beta=\pi-\alpha}^{\pi} e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)}\sin\alpha\sin\beta\sin(\alpha+\beta)d\beta d\alpha, \end{aligned} \tag{5.3.10}$$

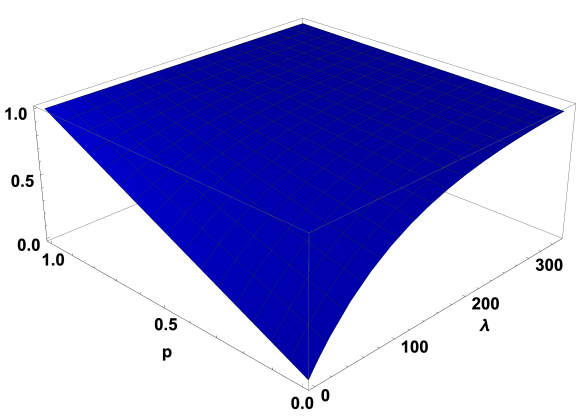
where $\mathscr{A}(l,\alpha,\beta)$ is the useful crossroad surface (either the one of the triangle or the one of the circumcircle according to the chosen scenario) defined in (5.3.7).

Figure 5.6 shows the plotting of the occupation probability of the typical crossroad $F(\lambda,p,l)$ as a function of the user density λ and the physical relay proportion p, when the street width l is equal to 20 metres, a typical value for a classical European city centre, estimated via statistical methods suggested in [59]. The blue surface (Figure 5.6a) corresponds to the triangle case ($\mathscr{A} = \mathscr{S}$, as defined in (5.3.1)), the orange surface (Figure 5.6a) corresponds to the circle case ($\mathscr{A} = \mathscr{S}'$, as defined in (5.3.6)). The crossroad occupation probabilities corresponding to both cases have been plotted together on Figure 5.6c: as stated above, for fixed network parameters (l,λ,p) it is clear that the occupation probability can become much larger in the circle case than in the triangle case, due to the fact that the ratio between the area of a triangle and the one of its circumcircle is rather small.

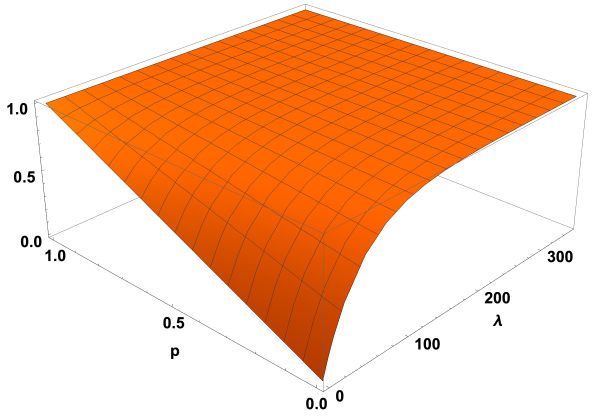
5.3.4 Numerical approach of relay deployment

Thanks to the previous geometric modelling for crossroads, we now propose a method to estimate a key quantity for network design and for operators: the minimal relay proportion needed for large-scale connectivity of the network. Indeed, the relay porportion p in our model is of critical importance, due to the fact that it is the network parameter which is the most directly impacted by an operator's strategical decisions. More precisely:

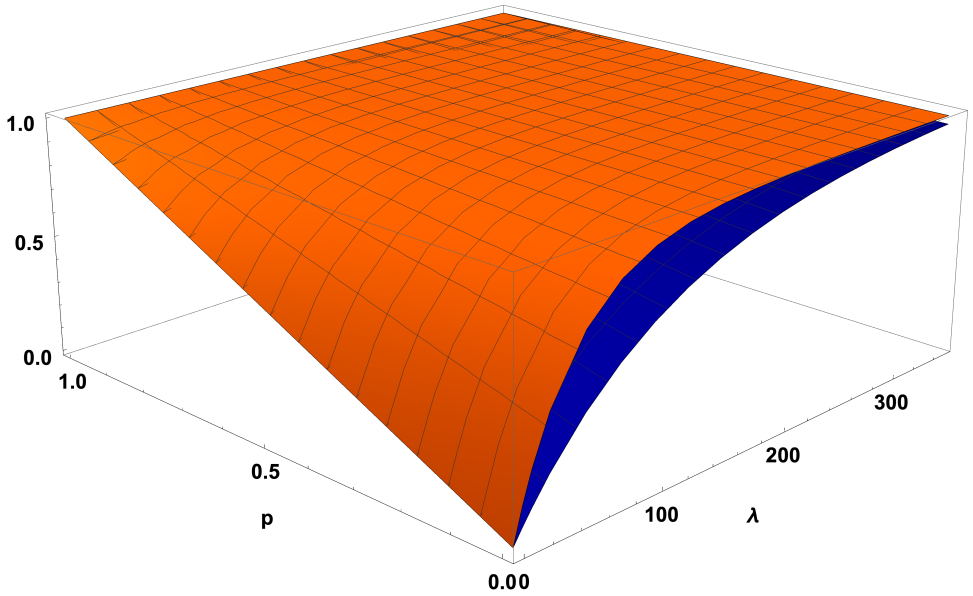
- The geometry of streets is given and so the operator cannot modify it. In our model, the geometry of the streets is given by the street density L_A and the street width $l > 0$.
- Regarding the (LOS) D2D range r, different communication ranges are possible according to the chosen technology (e.g. WiFi or 5G). To enhance D2D performance, and in particular the D2D range, a research effort is needed. This requires some time. Therefore, it is difficult for the operator to have leverage on the D2D range.



(a) Triangle case



(b) Circle case



(c) Triangle and circle case plotted together

Figure 5.6 – Crossroad occupation probability $F = F(\lambda, p, l)$ as a function of the user density λ and of the relay proportion p for a street width $l = 20\,\mathrm{m}$. Top left: Case where the useful surface for relaying at the crossroad is the surface of the triangle delimited by the intersection of the enlarged streets. Top right: Case where the useful surface for relaying at the crossroad is the surface of the circumcircle of the aforementioned triangle. Bottom: Plotting both cases together. Evaluation and plotting of the crossroad occupation probability F have been performed using the software `Mathematica`.

- Through advertising and marketing analyses, the operator can incentivise potential customers to subscribe to its network. Doing so, the operator would gain customers and this would result in a greater user density λ. However, the main problem is that the effects of such marketing policies on λ is hard to predict.
- However, the number of relays deployed over the network completely depends on the operator's choices.

Furthermore, deploying relays implies considerable costs for an operator, and so determining via statistical estimations the minimal number of such relays needed for large-scale

connectivity of the network appears even more crucial.

For convenience and avoiding heavy notations, we fix the street density $L_A > 0$ once and for all and omit the implicit dependencies on L_A from now onwards. Proceeding as in Chapter 4, we consider the network connectivity graph $\mathcal{G} = \mathcal{G}_{p,\lambda,r}$ in the canyon shadowing case and interpret percolation of this graph as large-scale connectivity of the modelled D2D network. Recall, as explained in Section 5.2.5, the main reason why we stick to the canyon shadowing case for numerical estimations is that r' has not much influence on critical network quantities in practice, due to the fact that $r' \ll r$ in real-world networks. Assuming our enlarged street model where crossroads have a non-null surface, crossroads can now either be occupied by *physical relays* or by a D2D user relaying the signal. The reason why we now use the terminology "physical relays" is that such relays would, in practice, correspond to fixed antennas, but *not* to D2D users passing by and relaying the signal. Hence, we define the *minimal proportion of physical relays* needed to ensure large-scale connectivity of the network as:

$$p_c(\lambda, r, l) \coloneqq \inf\left\{p \in [0,1]\,,\, \mathbb{P}\left(\mathcal{G}_{F(\lambda,p,l),\lambda,r} \text{ has an unbounded connected component}\right) > 0\right\},$$

where $F(\lambda, p, l)$ is the *crossroad occupation probability* previously defined in (5.3.10) and we set $p_c(\lambda, r, l) := \infty$ if $\mathbb{P}\left(\mathcal{G}_{F(\lambda,p,l),\lambda,r} \text{ has an unbounded connected component}\right) = 0$ for all $p \in [0,1]$. In other words, assuming that all crossroads of the network graph are statistically equivalent to the typical crossroad, we replace the probability p of finding a (physical) relay at a crossroad in the network graph by the crossroad occupation probability F. We then assess the minimal parameter p ensuring percolation of this new network graph, and hence large-scale connectivity of the network. Note that the quantity $p_c(\lambda, r, l)$ represents the necessary investment in relays for an operator, given the D2D user density λ, the D2D technology given by the D2D range r and the street width l.

To estimate $p_c(\lambda, r, l)$ we proceed as follows. First, using similar algorithms and statistical methods as the ones exposed in Section 4.3, we estimate the minimal proportion of *occupied crossroads* ensuring large-scale connectivity, defined as:

$$p_{\text{occupied}}(\lambda, r) \coloneqq \inf\left\{p \in [0,1]\,,\, \mathbb{P}\left(\mathcal{G}_{p,\lambda,r} \text{ has an unbounded connected component}\right) > 0\right\}, \tag{5.3.11}$$

with $p_{\text{occupied}}(\lambda, r) := \infty$ if $\mathbb{P}\left(\mathcal{G}_{p,\lambda,r} \text{ has an unbounded connected component}\right) = 0$ for all $p \in [0,1]$. Note that $p_{\text{occupied}}(\lambda, r)$ can be seen as the analogue of $p_c(\lambda, r, 0)$ when $l = 0$, i.e. in the original model where PVT streets had null width and where PVT crossroads were punctual. Indeed, the crossroad area can be written as $\mathscr{A}(l, \alpha, \beta)) = l^2 \times g(\alpha, \beta)$, where $g(\alpha, \beta)$ only depends on the random angles (α, β) in both cases, see (5.3.1) and (5.3.6). Therefore, we have $e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)} \uparrow 1$ as $l \downarrow 0$. By monotone convergence, it is thus easy to show that $\mathbb{E}\left[e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)}\right] \uparrow 1$ as $l \downarrow 0$. Hence:

$$\lim_{l \downarrow 0} F(\lambda, p, l) =: \lim_{l \downarrow 0} 1 - (1-p)\mathbb{E}\left[e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)}\right] = p,$$

so that the crossroad occupation probability in the initial model where $l = 0$ is equal to the probability p of having a relay at a crossroad. This result was expected, as a punctual

crossroad can only be occupied by a relay.

Now, fix λ, r, $l \neq 0$ and note that $F(\lambda, p, l)$ is increasing as a function of p, with:

$$F(\lambda, 0, l) = 1 - \mathbb{E}\left[e^{-\frac{\lambda}{l}\mathscr{A}(l,\alpha,\beta)}\right] =: \mathbb{P}(\text{there is at least one user in } \mathscr{A})$$
$$F(\lambda, 1, l) = 1.$$

Thus, if we consider the following equation

$$F(\lambda, p, l) = p_{\text{occupied}}(\lambda, r) \tag{5.3.12}$$

with the variable being p, we are in either one of the following two cases:

- $p_{\text{occupied}}(\lambda, r) \notin [F(\lambda, 0, l), F(\lambda, 1, l)]$: In that case, (5.3.12) has no solution and, by analogy with the definition of p_{occupied}, we set $p := \infty$.
- $p_{\text{occupied}}(\lambda, r) \in [F(\lambda, 0, l), F(\lambda, 1, l)]$: In that case, (5.3.12) has a unique solution p.

Denoting by p the previous solution of (5.3.12) (in the large sense), we then estimate the minimal proportion of physical relays needed for large-scale connectivity by taking

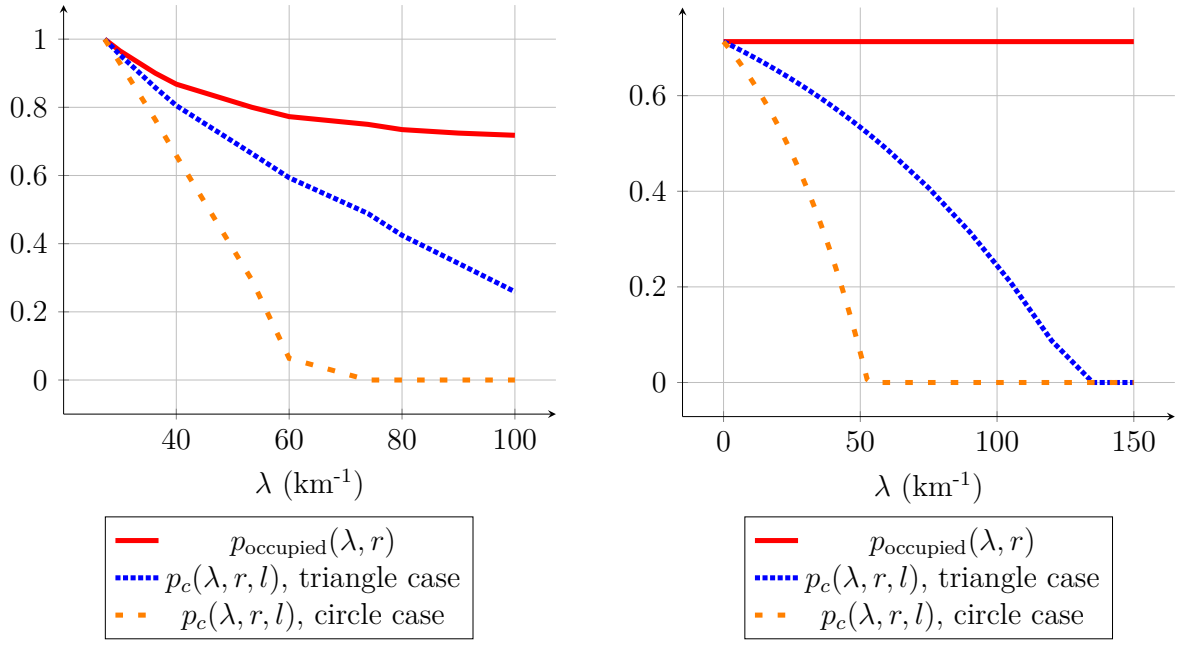
$$p_c(\lambda, r, l) \approx \max(1, \min(0, p)). \tag{5.3.13}$$

In other words, knowing the proportion of occupied crossroads required for good connectivity of the network and to what extent occupation of crossroads is due to users through λ, we deduce the remaining part $p \in [0, 1]$ due to physical relays. Also, note that if large-scale connectivity of the network is not possible even by deploying relays on all crossroads of the street system S, (5.3.12) ensures $p_c(\lambda, r, l) = 1$. In the same way, if the user density λ is sufficient to compensate the relays to allow for large-scale connectivity of the network, $p_c(\lambda, r, l) = 0$.

Figure 5.7 illustrates two examples of the estimations for this minimal physical relay proportion $p_c(\lambda, r, l)$ required for connectivity of the D2D network at large scale. In both examples, two facts are of noticeable importance.

On the one hand, taking into account the presence of D2D users being able to act as relays on crossroads considerably reduces the proportion of physical relays needed to ensure good connectivity of the network. On the other hand, the influence of the chosen geometry for the modelling of crossroads is crucial: the estimated physical relay proportion required for large-scale connectivity varies by at least a factor two in high user density scenarios. This is, once again, due to the fact that the ratio between the area of a triangle and its circumcircle is considerably small. As a matter of fact, it is much more likely to find mobile users at crossroads in the circumcircle case than in the triangle case, which results in less physical relays needed at a global scale.

Figure 5.7a presents the estimates when the D2D range $r = 50\,\text{m}$, which corresponds to a scenario where the radio technology supporting D2D would be WiFi [7]. In such a case, note that a sufficiently high density of users (mainly $\lambda \geq 60\,\text{km}^{-1}$), can fully compensate the relays in the circle case. Note that such a threshold is not unrealistic, as $\lambda \approx 60\,\text{km}^{-1}$ equivalently means that there are in average 60 D2D users per kilometre



(a) $r = 50\,\text{m}$ (D2D technology corresponding to WiFi) and $l = 20\,\text{m}$.

(b) $r = 200\,\text{m}$ (D2D technology corresponding to 5G) and $l = 20\,\text{m}$.

Figure 5.7 – Estimation of the minimal proportion of physical relays needed for large scale connectivity of the network, as a function of user density λ, for $r = 50\,\text{m}$ and $r = 200\,\text{m}$. In both cases, $L_A = 20\,\text{km}^{-1}$ (street density) and $l = 20\,\text{m}$ (street width). As suggested in [59], these are typical values for the fitting of a classical European city centre street system by a PVT.

of streets, which means less than one user every fifteen metres in average. This is easily reached in urban environments.

The case of larger connectivity radii is for instance depicted in Figure 5.7b, where $r = 200\,\text{m}$. Interestingly, we approach an asymptotic situation where the total proportion of occupied crossroads $p_{\text{occupied}}(\lambda, r)$ is independent of λ and equal to the PVT site percolation threshold $p^* \approx 0.713$ defined and estimated in Chapter 4 (see the the horizontal red line on Figure 5.7b). In practice, $r = 200\,\text{m}$ is quite a high D2D range, which can yet only be obtained by using so-called *millimeter-Wave* frequencies [113] for the D2D communications[2]. This will be the case in 5G networks, where one will thus approach the asymptotic situation depicted in Figure 5.7b.

5.4 Conclusion

In this chapter, noting that the canyon shadowing assumption studied in Chapter 4 is too restrictive for accurately modelling real-world networks, we have studied two extensions of our network model. In particular, our main goal was to show that, in practice, large-scale connectivity of the network may require a much smaller proportion of crossroads equipped with physical relays (i.e. antennas) than the previously estimated proportion of

[2]The terminology "millimeter-Wave" comes from the fact that, for such waves, the order of magnitude of the wavelength is around the millimeter.

71.3%.

On a theoretical perspective, by considering the case $r' > 0$ where connections due to scattering or reflection between adjacent streets are possible, we have proven that connectivity at large scale is theoretically possible for a relay proportion $p < 71.3\%$ and, in fact, for whatever small $p \in [0, 1]$. This may however require very large ranges for the user density λ and for r' that cannot be achieved in real-world networks.

This is why we introduced another more applied approach with new geometric models for streets and crossroads. Doing so, crossroads have a non-null surface and some part of the relaying can be ensured by users, as would be the case in real-world networks. This in turn allowed us to considerably reduce predictions of the minimal proportion of crossroads that have to be equipped with relays so as to ensure good connectivity of the network.

Our results bring new quantitative arguments to the question of relay deployment in D2D networks, which has a considerable strategical importance for telecommunications operators.

Chapter 6

Cost modelling and analysis of large-scale urban D2D networks

Good fortune is what happens when opportunity meets with planning.

Thomas Edison

In the previous chapters, our results showed that relays constitute a bottleneck for the connectivity of the network. From a telecommunications operator's perspective, deploying such relays comes with considerable costs and thus represents a necessary investment, which has to be compensated in some way. However, thanks to the D2D paradigm, operators can also count on their density of already existing subscribers to do some part of the signal relaying at crossroads and ensure a better connectivity of the network at a global scale: this is what we called *crowd-networking* in Section 2.2.4. Note that this may only be feasible in environments where the density of D2D users is sufficiently large: this is typically the case in *urban environments*. In such a context, D2D-aided crowd-networking can offer great economic opportunities for new actors willing to enter the telecommunications market by setting up a fully functional D2D network while limiting heavy investments in network infrastructure.

In this chapter, we will be interested in modelling and analysing the costs associated to relay deployment for ensuring good connectivity of D2D networks. This will in turn allow us to study crowd-networking scenarios through D2D from an economic perspective. In particular, and using our previous results, we will study as an example the case of a neo-operator willing to enter the telecommunications market by entirely relying on D2D to set up its network.

This chapter is based on the publication [88].

6.1 Introduction

In Section 2.2.4, we underlined that the explosion of the number of connected devices and the possibility of multihop D2D communications pave the way to *crowd-networking* scenarios. The main idea behind crowd-networking is that network users themselves will

take part in ensuring a better quality of service and connectivity of the network globally. For instance, in the D2D paradigm, this could consist in asking users to serve as relays. However, most of the time, users have their own behaviour and may not tend to cooperate, due to the fact that they have limited resources at their disposal (data amount of their mobile subscription, battery, ...). Therefore, operators need to incentivise users to cooperate: this can for instance be done by economic means.

D2D-aided crowd-networking scenarios bear high economic stakes for operators as well as for new actors willing to enter the telecommunications market. Indeed, resorting to D2D and using the density of their own customers, traditional operators could benefit from a coverage extension and a better quality of service. In sparser areas or in geographical locations that are less accessible and hence less served by the network, this is particularly interesting. Conversely, one can think of new actors (we call them *neo-operators*) willing to enter the telecommunications market and set up a fully functional network by entirely relying on D2D. By proposing a network service at a very low price compared to the subscription prices proposed by traditional operators, these neo-operators could attract a sufficient amount of customers to ensure large-scale connectivity of a D2D network and thus offer a service comparable to the one proposed by traditional operators. In exchange of paying a very low fee for the use of the network, users could agree to serve as D2D relays in the network and collaborate to get good connectivity at a global scale. If such scenarios were to be feasible, neo-operators would be capable of providing telecommunications network services without having properly invested or even owning network equipment: this is what we call *uberisation*[1]. Such uberisation scenarios could do great economic harm to historical operators, who, unlike neo-operators, had to massively invest in core network infrastructure at the beginning of their existence.

As a matter of fact, analysing the economic opportunities and threats of D2D in mobile networks is of great interest for actors of the telecommunications market, or for new actors willing to enter this market. Building on our previous D2D network model introduced in Section 3.3, we propose to introduce a related *economic model* to perform a cost analysis of D2D deployment in mobile networks. While our previous results allowed to study the technical feasibility of crowd-networking scenarios through D2D, our cost model will allow us to study their *economic feasibility*.

6.2 Economic model description

In this section, we present our economic model, which, coupled with the geometric models for the spreading of users and relays over the network, will allow to evaluate costs related to D2D deployment. Throughout our model, the (discrete) variable $t \in \mathbb{N}_0$ will denote the time in months elapsed since an initial instant $t_0 = 0$ defined according to our needs.

6.2.1 Cost parameters

In our economic model, we need to take into account the fact that deploying relays implies expenditures for operators. In return, by offering a telecommunications service to its

[1]The terminology comes from the company *Uber*, which deployed a ride-hailing service only by cooperating with its network of drivers, accepting to sign up and drive for Uber with their own cars.

Table 6.1 – Relevant cost parameters and their signification.

Parameter	Description
c_{CAPEX}	CAPEX cost of one physical relay
c_{OPEX}	Yearly OPEX cost of one physical relay
η	$c_{\text{OPEX}}/c_{\text{CAPEX}}$
G	Operator's monthly revenue per D2D user
T_{DEP}	Depreciation period of one physical relay (in years)
$\mathbf{A}$	Area covered by the D2D network

customers, the operator can expect a revenue. Finally, parameters for dimensioning the network (and hence ruling the strategical decisions of the operator while deploying relays) also need to be specified. Table 6.1 presents all the relevant parameters in a condensed way.

Parameters related to the costs of relays

Regarding the costs of relays, two types of costs are to be paid by an operator:

- *Capital expenditures* (CAPEX), representing the money spent by the operator to buy new relays or renewing existing relays.
- *Operational expenditures* (OPEX), representing money recurrently spent by the operator in its day-to-day operations to ensure the network's exploitation. In our case, these OPEX costs cover for instance maintenance of the relays or energy costs.

We denote by c_{CAPEX} the CAPEX cost of one physical relay and by c_{OPEX} the OPEX cost of one relay per year. As a matter of fact, as long as a relay is present in the network, the monthly expense done by the operator for this relay is equal to $c_{\text{OPEX}}/12$. The parameter $\eta = c_{\text{CAPEX}}/c_{\text{OPEX}}$ denotes the ratio between the CAPEX cost of one relay and the yearly OPEX cost of one relay. Finally, note that relays depreciate over time: that is to say, they lose their economic value over time (e.g. because of ageing). The operator will thus have to replace its relays after some time. To capture this, we consider a relay depreciation period T_{DEP} of years, after which relays will successively be replaced. Note that after this period, a relay may still be functional but has no economic value anymore for the operator.

Operator's revenue

Regarding the operator's revenue on the use of its D2D network, several sources are possible. For instance, the use of the D2D service can be billed to network users as a monthly subscription fee. Another source of revenue for the operator can consist in funding the deployment of the network by advertisement. For simplicity, we do not take

into account the nature of the revenue for the operator and model it by a parameter $G \geq 0$ corresponding to the monthly revenue received by the operator per D2D user.

Parameters for dimensioning the network

Now that we have specified the CAPEX and OPEX costs of one relay as well as the operator's revenue per D2D user, we need to specify the number of relays and D2D users in the network. In this regard, we assume that an area $\mathbf{A}$, expressed in squared kilometres (km^2) is covered by the considered D2D network. Moreover, we use our previous models for the street system, the D2D users and the relays. Recall the corresponding parameters:

- The street density L_A, expressed in inverse kilometres (km^{-1}) denotes the mean street length per unit area. If we consider the enlarged version of the street system, as in Chapter 5, we denote again by $l > 0$ the width of all streets.
- The *D2D users' linear intensity* $\lambda > 0$, also expressed in km^{-1}, denotes the mean number of D2D users per unit length of street. If we consider the enlarged version of the street system, this is equivalent to considering a surface density of D2D users equal to λ/l, representing the mean number of D2D users per unit area.
- The *relay proportion* p, dimensionless, denotes the probability to find a physical relay at a crossroad of the street system. In frequentist terms, p represents the proportion of crossroads equipped with a relay.

Recall from [108] that since the street system is a planar PVT, the intensity λ_0 of vertices of the PVT, or, in other words, the mean number of crossroads per unit area, is related to the street density L_A in the following way: $\lambda_0 = L_A^2/2$. Thus, we can express the mean number of relays in the network as $pL_A^2\mathbf{A}/2$. In the same way, since λ denotes the linear intensity of D2D users and L_A denotes the density of streets, the mean number of D2D users in the network can be expressed as $\lambda L_A \mathbf{A}$.

6.2.2 Relay investment strategy and dynamics of user density

To perform economic analyses of D2D deployment in mobile networks, it is not sufficient to consider the previous cost parameters. Indeed, one also needs to consider the variations over time of the amount of relays and of the number of D2D users in the network.

The evolution of the number of relays that are deployed in the network is entirely governed by the strategical decisions of the operator. Indeed, based on the current conjecture, the operator chooses how much and when relays shall be invested in and deployed over the network. Hence, along with the previous cost parameters, the operator's strategy for relay investment shall be considered in the model. Note that such a strategy may strongly be influenced by estimations of critical network quantities available to the knowledge of the operator. For instance, our previous results on the minimal proportion of relays needed for large-scale connectivity of the network may help an operator to elaborate its relay deployment strategy, ruling how much relays will be invested in over time. In this regard, the relay proportion $p = p(t)$ can thus be seen as a time-dependent function. In the same way, we denote by $N_B(t)$ the number of relays that need to be bought by the operator at the beginning of month t. Also, we write $N(t)$ for the total number of relays present in

the network at the end of month t.

The number of D2D users in the network, i.e. customers of the operator, may fluctuate over time, due to several factors, among which are the economic context, the competition coming from other operators or the users' own behaviour (for instance when they choose to switch from one operator to another). To capture this phenomenon, we allow the user density $\lambda = \lambda(t)$ to depend on time. Note that an operator may also predict how its density of customers may vary over time, e.g. by marketing analyses.

6.2.3 Quantities of interest for cost analysis

On an economic perspective, in order to analyse in more detail the costs related to relay deployment in our D2D network model, we introduce three quantities of interest for an operator's strategical decisions.

First, the *cash flow* $CF : t \mapsto CF(t)$ is a function of time whose value at a given month t corresponds to the difference between the money earned and the money spent by the operator at month t. In other words:

$$CF(t) \coloneqq G\lambda(t)L_A\mathbf{A} - N_B(t)c_{\text{CAPEX}} - N(t)\frac{\eta c_{\text{CAPEX}}}{12} \tag{6.2.1}$$

In (6.2.1), the term $G\lambda(t)L_A\mathbf{A}$ corresponds to the revenue of the operator at month t, while the terms $N_B(t)c_{\text{CAPEX}}$ and $N(t)\frac{\eta c_{\text{CAPEX}}}{12} = N(t)\frac{c_{\text{OPEX}}}{12}$ respectively represent the CAPEX spent at month t (corresponding to the acquisition of new relays) and the OPEX spent at month t (costs related to the exploitation of the network). If the cash flow CF is non-negative at a given month, it means that, during this month, the operator has earned more money than it has spent and has thus generated benefits. On the contrary, a negative cash flow means that the monthly expenses were not compensated by the monthly earnings and so the operator undergoes an economic loss. In such a case, the operator has to take money from its savings to keep running its service.

It is also easy to see that the cash flow is deeply related to the operator's ability to generate value and be profitable in the long run. This is captured by the *cumulated revenue* $CR : t \mapsto CR(t)$ of the operator, a function of time giving the amount of money earned by the operator up to month t. In other words:

$$CR(t) \coloneqq \sum_{s=0}^{t} CF(s) \tag{6.2.2}$$

So to speak, the cumulated revenue can be seen as the balance of the operator's bank account at month t: when the cumulated revenue is negative at a given time, it means that the operator has lost more money than it has earned up to that time. When the cumulated revenue becomes non-negative, it means that the operator has performed a return on its initial investment in relays and has earned more money than spent since the initial instant $t_0 = 0$. Most of the time, this initial instant t_0 corresponds to the entrance of the operator on the telecommunications market.

In the end, the operator wishes to have a return on investment ROI by counting on the revenue coming from its customers and a limited investment in renewing relays. This return on investment is reached when the cumulated revenue CR becomes positive:

$$ROI \coloneqq \min\{t \geq 0 \,:\, CR(t) > 0\}. \tag{6.2.3}$$

Note that CR may not be a monotone and thus the cumulated revenue CR may become negative again once the return on investment is reached. This situation must of course preferably be avoided by operators, as their goal is to be profitable in the long run.

6.3 An example: analysis of an uberising neo-operator's strategy

6.3.1 Main assumptions

We now use our economic model to study an example of uberising scenario where a neo-operator is willing to enter the telecommunications market by deploying a network relying on D2D technology only, so as to limit infrastructure investments in relays and base stations.

Regarding the dimensioning of the network, we assume that the territory covered by the network is a dense urban environment and consider, as in Chapter 5, that the street system is modelled by an enlarged PVT with street density L_A and where all streets have a positive width $l > 0$ (see Figure 5.4). The D2D users are still modelled by a Cox process X^λ and the physical relays by a Bernoulli process Y. Regarding the connectivity model, we use the same model as the one studied in Chapter 5, namely the canyon shadowing case ($0 < r < \infty, r' = 0$) with possible relaying of the signal either by D2D users of by physical relays at crossroads of streets.

Moreover, we assume that all customers of the neo-operator will accept to serve as relays using multihop D2D. They may for instance be incentivised to do so by being offered a very low subscription price compared to those offered by traditional operators. The fact that all customers of the neo-operator agree to serve as D2D relays implies that the density of customers is simply equal to the D2D user density.

6.3.2 Relay deployment strategy

We now need to specify the relay deployment strategy of the studied neo-operator. The main idea is to spread CAPEX investments over time and not massively invest in buying relays too quickly. In this regard, we assume that the deployment of relays will be done in two steps: a first network phase before the commercial launch of the neo-operator's service and a second one after the commercial launch, when the neo-operator starts to gain customers. The relay deployment strategy of the neo-operator is illustrated by Figure 6.1. Table 6.2 presents relevant parameters for understanding the uberising strategy of the neo-operator. The initial time $t_0 = 0$ will from now onwards denote the first month of relay deployment, i.e. the moment when the operator starts to buy its relays.

In more detail, an initial investment in relays (CAPEX) has to be made before the commercial launch of the service ($t < T_{\text{LAUNCH}}$). During this initial period of network deployment, no customers have subscribed to the service and so the user density is null ($\lambda(t) = 0$). Note that once a relay has been installed, operational costs (c_{OPEX}) are to be paid for this relay by the operator. We denote by $p_{\min}$ the proportion of crossroads that the neo-operator wishes to equip with relays at the end of this first phase. The choice of a value for $p_{\min}$ is done by the neo-operator and can be oriented by the results about the

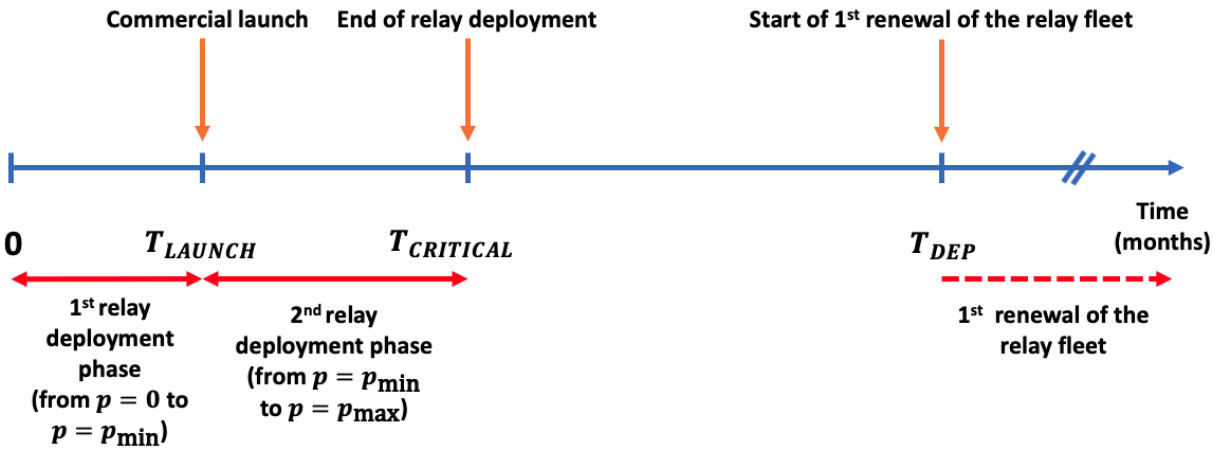


Figure 6.1 – Relay deployment strategy of the neo-operator to set up its D2D network.

Table 6.2 – Relevant quantities for the neo-operator's relay deployment strategy.

Parameter	Description
t	Time (in months)
$\lambda(t)$	User (linear) intensity at month t
$N_B(t)$	Number of relays bought by the neo-operator at the beginning of month t
$N(t)$	Total number of relays in the network at the end of month t
T_{LAUNCH}	Time at which the service is commercially launched
T_{CRITICAL}	Time at which all relays have been deployed
T_{DEP}	Depreciation period of one relay and time at which relays will start to be successively replaced

minimal relay proportion needed for large-scale connectivity of the network, see Chapter 5. We also assume that the relays are deployed in equal amount during each month of the initial deployment period. In other words, during the first phase of relay deployment, the number of relays N_B bought each month is constant and the total number of relays N in the network is a linear function of time. Recalling that the mean number of relays in the network is given by $pL_A^2\mathbf{A}/2$ and assuming for simplicity that $\frac{p_{\min}L_A^2\mathbf{A}}{2} \in \mathbb{N}$, we can thus write:

$$\begin{cases} N_B(t) &= \frac{p_{\min}L_A^2\mathbf{A}}{2T_{\text{LAUNCH}}} \qquad \text{for } 0 \leq t \leq T_{\text{LAUNCH}} - 1 \\ N(t) &= \frac{p_{\min}L_A^2\mathbf{A}}{2T_{\text{LAUNCH}}} + \frac{p_{\min}L_A^2\mathbf{A}}{2T_{\text{LAUNCH}}}t \quad \text{for } 0 \leq t \leq T_{\text{LAUNCH}} - 1 \end{cases} \tag{6.3.1}$$

At month $t = T_{\text{LAUNCH}}$, the service of the neo-operator is commercially launched. From that moment, the operator gets a revenue G per month for each D2D user. Worth noticing is that the neo-operator might not have sufficiently many customers yet to ensure a good

quality of service. Over time and by contagion effect, new customers will be attracted and will allow to reach the critical mass needed to ensure large-scale connectivity of the network.

When the D2D service is commercially launched, the neo-operator will start its second relay deployment phase ($T_{\text{LAUNCH}} \leq t < T_{\text{CRITICAL}}$), until reaching a proportion $p_{\max}$ of crossroads equipped with relays. Again, we assume that the relays are deployed in equal amount during each month of this second phase. Thus, similarly to (6.3.1) and assuming that $\frac{p_{\max} L_A^2 \mathbf{A}}{2} \in \mathbb{N}$, we write for the second phase of relay deployment:

$$\begin{cases} N_B(t) &= \frac{(p_{\max}-p_{\min})L_A^2\mathbf{A}}{2(T_{\text{CRITICAL}}-T_{\text{LAUNCH}})} & \text{for } T_{\text{LAUNCH}} \leq t \leq T_{\text{CRITICAL}}-1 \\ N(t) &= \frac{p_{\min}L_A^2\mathbf{A}}{2} + \frac{(p_{\max}-p_{\min})L_A^2\mathbf{A}}{2(T_{\text{CRITICAL}}-T_{\text{LAUNCH}})}(t - T_{\text{LAUNCH}} + 1) & \text{for } T_{\text{LAUNCH}} \leq t \leq T_{\text{CRITICAL}}-1. \end{cases} \tag{6.3.2}$$

In practice, the choice of values for $p_{\max}$ and T_{CRITICAL} can also be oriented by the results about the minimal relay proportion needed for large-scale connectivity of the network. Typically, the previous values are chosen by the neo-operator in such a way that when all relays have been deployed (i.e. in proportion $p_{\max}$ at time T_{CRITICAL}), the critical mass of users ensuring large-scale connectivity of the network has almost been reached. In other words, this means that the neo-operator tunes $p_{\max}$ and T_{CRITICAL} so as to have:

$$p_{\max} \approx p_{\text{occupied}}(\lambda(T_{\text{CRITICAL}}), r), \tag{6.3.3}$$

where, $p_{\text{occupied}}(\lambda, r)$ was defined in (5.3.11). Recall that $p_{\text{occupied}}(\lambda, r)$ is the proportion of occupied crossroads (either by a D2D user or by a relay) needed to ensure large-scale connectivity of the network under a user density λ and a D2D range r.

Finally, once all relays have been deployed ($t \geq T_{\text{CRITICAL}}$), the neo-operator will begin to replace them at time $t = T_{\text{DEP}}$, in such a way that the whole relay fleet will be entirely replaced within another depreciation period $[T_{\text{DEP}}, 2T_{\text{DEP}}[$, and so on. In other words, after the critical time T_{DEP} the total number of relays N in the network remains constant, while the number of bought relays N_B is null before T_{DEP} and constant after. We thus write:

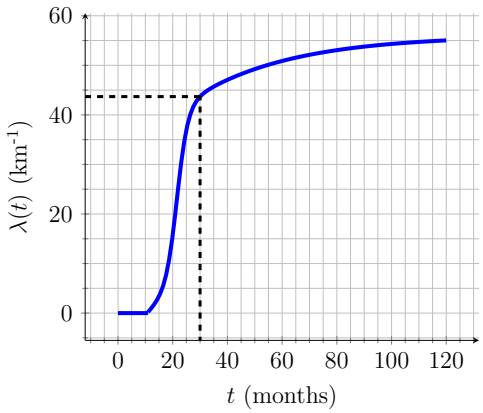
$$N(t) = \frac{p_{\max} L_A^2 \mathbf{A}}{2} \quad \text{for } t \geq T_{\text{CRITICAL}} \tag{6.3.4}$$

$$N_B(t) = \begin{cases} 0 & \text{for } T_{\text{CRITICAL}} \leq t \leq T_{\text{DEP}} - 1 \\ \frac{p_{\max} L_A^2 \mathbf{A}}{2T_{\text{DEP}}} & \text{for } t \geq T_{\text{DEP}}. \end{cases} \tag{6.3.5}$$

Putting equations (6.3.1), (6.3.2), (6.3.4) and (6.3.5) together, we thus obtain general expressions for the total number of relays N and the number of relays to be bought each month N_B. This in turn allows to compute the cash flow CF, the cumulated revenue CR and the return on investment ROI defined by equations (6.2.1) to (6.2.3).

6.3.3 Simulation parameters

We performed a numerical evaluation of the uberising strategy presented in Section 6.3.2 with parameters coming from internal data provided in Table 6.2b and a D2D user density



(a) User density as a function of time.

Simulation parameter	Numerical value
c_{CAPEX}	€1200
η	10%
G	€3
T_{DEP}	7 years = 84 months
T_{LAUNCH}	1 year = 12 months
T_{CRITICAL}	2 years 1/2 = 30 months
$p_{\min}$	10%
$p_{\max}$	20%
$\lambda(t)$	As in Fig. 6.2a
L_A	20 km/km^2
$\mathbf{A}$	25 km^2

(b) Numerical values of other simulation parameters.

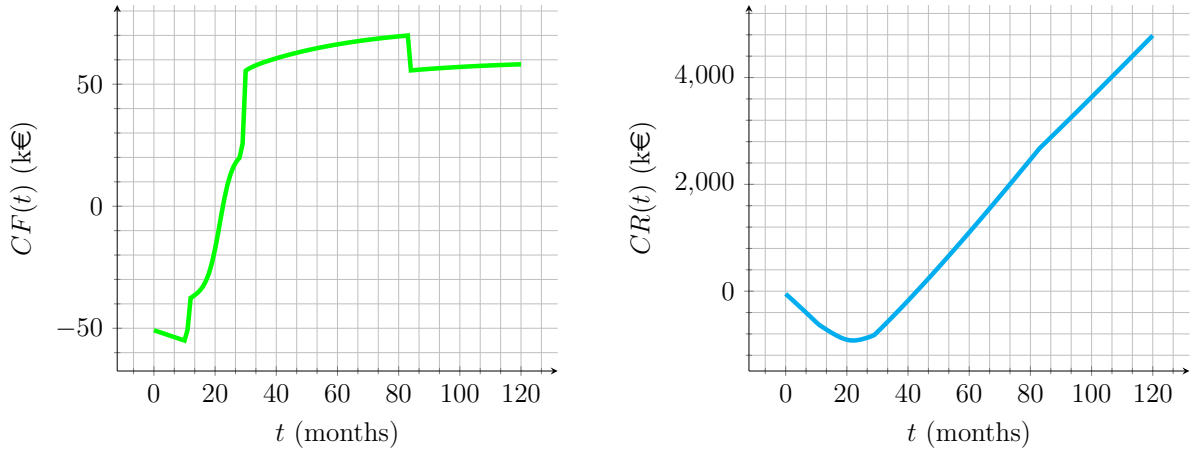
Figure 6.2 – Parameters used for numerical evaluation of the neo-operator's uberising strategy.

$\lambda(t)$ illustrated by Figure 6.2a.

The parameters dimensioning the typology of the territory covered by the network are the street density $L_A = 20\,\text{km}^{-1}$ and the area covered by the network $\mathbf{A} = 25\,\text{km}^2$. As suggested by fitting against real data [58], such values are typical for the city centre of a major European city.

We chose a user density $t \mapsto \lambda(t)$ which is increasing as a function of time, thus assuming that the neo-operator does not lose customers. Note the user density is null before commercial launch, i.e. for $t \leq T_{\text{LAUNCH}} = 1\,\text{year}$, then undergoes a steep increase corresponding to early adopters and finally undergoes a slower growth once the early adopters effect is over.

In our numerical example, we took $p_{\min} = 10\%$ and $p_{\max} = 20\%$. Moreover, the critical time is equal to $T_{\text{CRITICAL}} = 30\,\text{months}$. Note that this choice of parameters roughly satisfies (6.3.3). Indeed, $\lambda(T_{\text{CRITICAL}}) \approx 45\,\text{km}^{-1}$ (see the black dashed lines on Figure 6.2a). In a 5G millimeter-Wave context where $r = 200\,\text{m}$ (see Figure 5.7b), assuming circle crossroads, this density of users requires a minimal relay proportion of about $p_c(\lambda = 45\,\text{km}^{-1}, r = 200\,\text{m}) \approx 20\%$. Thus, at the critical time T_{CRITICAL}, the neo-operator has fully deployed its relays and ensured large-scale connectivity of its D2D network.



(a) Cash flow CF as a function of time.

(b) Cumulated revenue CR as a function of time.

Figure 6.3 – Cash flow CF and cumulated revenue CR in thousands of euros (k€) as a function of time. The return on investment is reached at $ROI = 43$ months.

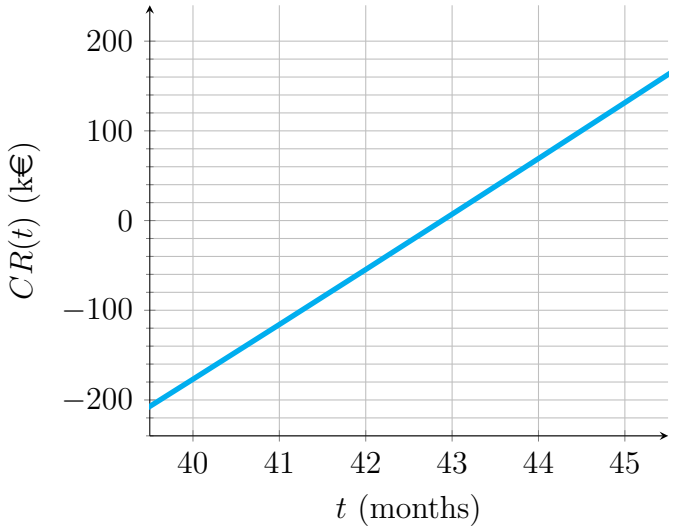


Figure 6.4 – Zoom of Figure 6.3b. The return on investment is reached at $ROI = 43$ months.

6.3.4 Results

Figure 6.3 shows the evolution of the cumulated revenue and of the cash flow over time. Before the commercial launch, only CAPEX and OPEX are spent in relays, and so CR is decreasing. From the commercial launch, the operator starts to get a revenue: this is the first abrupt increase of the cash flow CF. At the beginning of the commercial exploitation of the network, the revenue coming from users may not compensate the CAPEX and OPEX spent in relays, which is why the cumulated revenue CR stays negative. At the time T_{CRITICAL} denoting the end of the second relay deployment phase, the cash flow CF increases abruptly for the second time, due to the fact that the neo-operator does not have to buy new relays anymore (CAPEX investments thus drop to 0). Once all relays

have been deployed, only OPEX costs are spent and the growth of the user density allows to increase the cumulated revenue of the operator. Going further in time, the time T_{DEP} at which relays start to be replaced corresponds to the sudden drop of the cash flow CF. This drop is due to the fact that the neo-operator has to buy new relays (CAPEX) to replace the old ones.

With the parameters values prescribed by Table 6.2b, we obtain a return on investment $ROI = 43$ months i.e. 31 months after the commercial launch of the D2D service (see Figure 6.4). In practice, note that such a value is more than satisfying for potential investors: the return on the initial investment is realised in less than 3 years after the commercial launch of the service.

6.4 Conclusion

In this chapter, we have presented a cost model for analysing the deployment of D2D in mobile networks from an economic perspective. This model allows one to more reliably quantify the necessary investments in relays required to deploying a relay-assisted D2D network. As a main application, we studied a practical uberisation scenario where a neo-operator, willing to enter the telecommunications market and disrupt conventional operators, relies on D2D only to set up its network. In such a case, note that the number of D2D users may be too small at the beginning of the service to get a fully connected network. The neo-operator could then start by offering proximity services, until large-scale connectivity of its network is ensured by sufficiently many D2D users. Another possibility to reach this critical mass of users could consist in resorting to customers of existing operators through negociated costs.

Finally, funding the D2D service from its very beginning is essential, otherwise the return on investment would happen much later in time, and thus the uberisation scenario described in our model might not be profitable for a new actor in the telecommunications sector. Our cost model is also sufficiently generic so that any operator willing to launch a D2D service can replace our simulation parameters with its own internal data. This allows one to assess when setting up a relay-assisted D2D network will be profitable and paves the way to interesting discussions regarding deployment of relay-assisted D2D networks.

General conclusion and research perspectives

In this thesis, we focused on new stochastic models of multihop device-to-device (D2D) networks in urban environments allowing one to assess the technical and economic feasibility of such networks at large scale. After having presented the mathematical background of our work and having given an overview of wireless and D2D communications, we have addressed the following research objectives.

Reviewing the different approaches adopted in the stochastic geometry literature to model telecommunications networks, we have underlined that appropriate models for D2D networks, especially in urban environments, shall take into account that the underlying topology of the network environment is of critical importance. Moreover, realistic connectivity conditions have to be considered. Based on these points, we thus built and proposed new models where the support of the network is given by a planar Poisson-Voronoi tessellation (PVT), thought of as a street system of the urban environment covered by the D2D network. Doubly stochastic point processes supported by the PVT prescribe the spatial distributions of network users and relays in the network. Representing the network by a random connectivity graph, we then interpreted percolation of this graph as good connectivity of the network.

In a first approach, we studied the case of a network environment with many obstructions, resulting in particularly restrictive connectivity conditions where only line-of-sight (LOS) connections along the edges of the PVT within a certain range are possible. We called this the canyon shadowing assumption. In this case, we studied percolation of the network connectivity graph both from a theoretical and a numerical point of view and proved that phase transitions between several connectivity regimes with very different properties exist. Particular cases corresponding to percolation models that had not been studied in the literature before have also been considered. What is more, our numerical simulations allowed us to estimate the frontiers between the various connectivity regimes, which is of great practical interest for economic planning of large-scale D2D networks. In particular, we were able to estimate critical quantities such as the minimal density of users allowing good connectivity of the network.

In the canyon shadowing case, the presence of relays at crossroads of streets (i.e. vertices of the PVT) is essential to relay the signal between adjacent streets (edges of the PVT). As a matter of fact, relays constitute a bottleneck for good connectivity of the network and a

necessary investment for operators. Moreover, our simulations allowed us to predict that under an absolute minimal proportion $p^* \approx 0.713$ of crossroads equipped with a relay, good connectivity of the network is not possible, no matter what. Such a high value for the relay proportion may not be feasible for operators under reasonable costs and is mainly due to the fact that, in our model, the probability that some Cox point is located at a vertex of the PVT is exactly 0. In other words, relaying of the signal at crossroads can only be done by relays. In real-world networks, however, one may expect some part of the network users themselves to perform this relaying. We thus proposed two variants to deal with the previous issue:

- On the one hand, we considered more general connectivity conditions, allowing for supplementary non-line-of-sight (NLOS) connections due to scattering and/or reflections of the signal at crossroads. This gives rise to a new network connectivity graph, whose percolation was studied.

- On the other hand, we considered a new geometrical model where streets are given a positive width and crossroads have a positive surface. As a matter of fact, the probability of finding a user at a crossroad and able to relay the signal becomes positive. This in turn allowed us to revise the critical network estimates that had been estimated in the previous canyon shadowing case. We particularly focused on the minimal proportion of crossroads that have to be equipped by a relay to ensure good connectivity of the network. This quantity indeed represents the necessary investment that operators have to make in relays before being able to set up a functional D2D network at large scale.

Our next step consisted in introducing an operational cost model related to our network stochastic model. This cost model aims at quantifying the costs related to relay deployment in D2D networks and investigate possible economic consequences of the feasibility of large-scale D2D networks for operators. As an application, we studied an uberisation scenario where a neo-operator enters the telecommunications market by setting up a functional telecommunications network relying only on D2D. The idea behind this scenario is to offer a telecommunications service to users at a very low price, while counting on them to act as D2D relays. In the end, the neo-operator wishes to have a return on investment (ROI) by counting on the revenue coming from its customers and a limited investment in relays. Using our previous numerical results about large-scale connectivity, we studied an investment strategy in relays making such a neo-operator become profitable at a reasonable time horizon, even if the revenue earned per D2D user is very low compared to those of traditional operators. This study outlines the great economic opportunities and threats that could arise from the arrival of D2D in the future cellular networks.

The results presented in this thesis can be used as a first basis towards a more complete analysis of the connectivity of large D2D networks, of their technical feasibility and of the associated economic consequences. Several improvements can be considered for future research directions.

Regarding our D2D network stochastic model, many perspectives can be explored.
First, other choices of random tessellations for the street system could be considered.

The phase transition results of Chapters 4 and 5 can easily be generalised to the *dual tessellation* of a PVT: a *Poisson-Delaunay* tessellation (PDT). More generally, we also believe that our results are still valid for any stabilizing and asymptotically essentially connected tessellation as defined in Definitions 4.4.1 and 4.4.2. Other tessellation models could been considered. In particular, when thinking of African or North-American cities, the so-called Manhattan grids (MG) are often investigated. Nested tessellations could also be considered in anisotropic environments or when the territory covered by the network cannot be considered as morphologically homogeneous, see [32].
Secondly, refined connectivity models closer to the physical reality of radio wave propagation could be investigated. For instance, appropriate models for interference could be considered, as has been done in [128]. The case of random radii for the communication ranges could also be studied. Moreover, these refined connectivity models could be combined with refined geometric models for crossroads and streets, such as the enlarged version of the PVT street system considered in Chapter 5.
Thirdly, and perharps more importantly, appropriate models for user mobility in the network could be introduced and would allow to consider the problem of large-scale connectivity of D2D networks from a dynamic point of view.

Regarding the mathematical part of our work, note that we only limited ourselves to the existence of phase transitions between the different connectivity regimes, which is the primary question in any percolation problem. A next step could consist in investigating the percolation probability θ that some network node belongs to the infinite connected component of the network. As has been done in [29], this percolation probability can be interpreted by the probability that some randomly selected user of the network belongs to the largest communication cluster. Going further, one could look at the stretch factor μ giving the ratio between the number of hops needed to connect two users of the network and their relative Euclidean distance. Though values of this stretch factor μ are accessible via numerical simulations, proving its existence has not been done yet.

Finally, our operational cost model paves the way towards multiple research directions. Rather than investigating threats coming from neo-operators, one could investigate economic opportunities related to coverage extensions by D2D for traditional operators: for instance, relying on users to enhance connectivity in areas less served by the network could avoid massive expenditures in network infrastructure for operators. More complicated scenarios where different neo-operators owning different kinds of relays could also be investigated, and it would be interesting to see whether coalition or competition between them would allow for a better connectivity at a global scale. Another stream of generalisation could also consist in considering users that may not be willing to cooperate and serve as D2D relays, contrary to the uberisation scenario we studied, where all customers of the neo-operators agree to serve as a D2D relay. In real-world networks, users have their own behaviour and have limited resources (e.g. battery) at their disposal. They must thus be incentivised (e.g. by economic ways) to cooperate and serve as relays: this would represent additional costs which must be considered by operators.

Appendix A

0-1 law for the percolation probability of the network connectivity graph

We must stand firm between two kinds of madness: the belief that we can do anything; and the belief that we can do nothing.

Alain

In this appendix, we state and prove an additional result about the percolation probability of the network connectivity graph presented in Section 3.3. More precisely, we have a 0-1 law for the percolation probability of the general network graph $\mathcal{G} = \mathcal{G}_{p,\lambda,r,r'}$, regardless of the values of the network parameters (L_A, p, λ, r, r'). We obtain this result in two steps: first, we prove an ergodicity result on the superposition $Z := X^\lambda \cup Y$ of the point processes of users and relays. Then, as a consequence of this ergodicity result, we obtain a 0-1 law for the percolation probability $P(p, \lambda, r, r')$ of $\mathcal{G}$.

Our ergodicity result for Z is the following one:

Proposition A.0.1. *The superposition Z of the point processes of users and relays is mixing, and hence ergodic.*

Proof. To prove that Z is mixing, we will work on the canonical space and it will thus suffice to show that

$$\lim_{\|x\|_2 \to \infty} \mathbb{P}(A \cap S_x B) = \mathbb{P}(A)\mathbb{P}(B), \tag{A.0.1}$$

for all events $A, B \in \sigma(Z)$ that are measurable with respect to the sigma-algebra $\sigma(Z)$ generated by Z and where $\{S_x\}_{x\in\mathbb{R}^2}$ denotes the natural shift on $\mathbb{R}^2$.

Note first that by [38, Lemma 12.3.II] (and as has been done in the proof of [38, Proposition 12.3.VI]), it suffices to check the mixing condition (A.0.1) for *local events*, i.e. of the form $A \in \sigma(Z \cap W_A)$ and $B \in \sigma(Z \cap W_B)$, where W_A and W_B are compact observation windows in $\mathbb{R}^2$. Thus, let A and B be such events. We will show that for all $\epsilon > 0$, we

can find x with $\|x\|_2$ sufficiently large so that $|\mathbb{P}(A \cap S_x B) - \mathbb{P}(A)\mathbb{P}(B)| \leq \epsilon$.

Take any $\epsilon > 0$. By condition (2) in the definition of stabilization (Definition 4.4.1), we can find sufficiently large $n \geq 1$ such that $\mathbb{P}(R(Q_n) \geq n) \leq \epsilon/3$. Moreover, such n can be chosen so as to satisfy $W_A \subseteq Q_n$ and $W_B \subseteq Q_n$. Fix such n.

Since $A \in \sigma(Z \cap W_A)$, A only depends on the configuration of Z inside W_A. In the same way, B only depends on the configuration of Z inside W_B, and $S_x B$ only depends on the configuration of Z inside $S_x W_B = W_B - x$. Since $B \subseteq Q_n$, we have that $S_x W_B \subseteq Q_n - x =: Q_n(-x)$. Take x with $\|x\|_2 > 6n\sqrt{2}$. Then we have $Q_n \cap Q_n(-x) = \emptyset$ and thus $W_A \cap S_x W_B = \emptyset$, so that the events A and $S_x B$ depend on the configuration of Z in disjoint sets. Since the conditional distribution of Z given the random support Λ is that of a superposition of a Bernoulli process and of a Poisson point process, the events A and $S_x B$ are conditionnally independent given Λ. Hence:

$$\mathbb{P}(A \cap S_x B) = \mathbb{E}\Big[\mathbb{E}\left(\mathbb{1}\{A\}\mathbb{1}\{S_x B\} \,|\, \Lambda\right)\Big] = \mathbb{E}\Big[\mathbb{E}(\mathbb{1}\{A\} \,|\, \Lambda)\mathbb{E}\left(\mathbb{1}\{S_x B\} \,|\, \Lambda\right)\Big] \tag{A.0.2}$$

Now, since $A \in \sigma(Z \cap W_A)$ with $W_A \subseteq Q_n$, we can write $\mathbb{E}(\mathbb{1}\{A\}\,|\Lambda) = f(\Lambda_{Q_n})$ as a bounded deterministic function of Λ_{Q_n}. In the same way, we can write $\mathbb{E}(\mathbb{1}\{S_x B\}\,|\Lambda) = g(\Lambda_{Q_n(-x)})$ as a bounded deterministic function of $\Lambda_{Q_n(-x)}$. By (A.0.2), we thus get:

$$\begin{aligned}
\mathbb{P}(A \cap S_x B) &= \mathbb{E}\left[f(\Lambda_{Q_n}) g(\Lambda_{Q_n(-x)})\right] \\
&= \mathbb{E}\left[f(\Lambda_{Q_n}) g(\Lambda_{Q_n(-x)}) \mathbb{1}\{R(Q_n) < n\}\mathbb{1}\{R(Q_n(-x)) < n\}\right] \\
&+ \mathbb{E}\left[f(\Lambda_{Q_n}) g(\Lambda_{Q_n(-x)}) \mathbb{1}\{(R(Q_n) \geq n) \cup (R(Q_n(-x)) \geq n)\}\right]
\end{aligned} \tag{A.0.3}$$

Let us first deal with the second term appearing in the right-hand side of (A.0.3). Using the fact that both f and g, being conditional expectations of indicator functions, are upper-bounded by 1, we get:

$$\begin{aligned}
&\mathbb{E}\left[f(\Lambda_{Q_n}) g(\Lambda_{Q_n(-x)}) \mathbb{1}\{(R(Q_n) \geq n) \cup (R(Q_n(-x)) \geq n)\}\right] \\
&\qquad \leq \mathbb{P}\left[(R(Q_n) \geq n) \cup (R(Q_n(-x)) \geq n)\right] \\
&\qquad \leq \mathbb{P}\left[R(Q_n) \geq n\right] + \mathbb{P}\left[R(Q_n(-x)) \geq n\right]
\end{aligned} \tag{A.0.4}$$

where we have used the union bound in (A.0.4). Now, by stationarity of the stabilization field $\{R_y\}_{y \in \mathbb{R}^2}$, we get that the right-hand side in (A.0.4) is equal to $2\mathbb{P}\left[R(Q_n) \geq n\right]$. In all, we thus get:

$$\mathbb{E}\left[f(\Lambda_{Q_n}) g(\Lambda_{Q_n(-x)}) \mathbb{1}\{(R(Q_n) \geq n) \cup (R(Q_n(-x)) \geq n)\}\right] \leq 2\epsilon/3 \tag{A.0.5}$$

We now deal with the first term appearing in the right-hand side of (A.0.3). Note that since $\|x\|_2 > 6n\sqrt{2}$, the set $\varphi := \{0, -x\} \subset \mathbb{R}^2$ satisfies $\forall y \in \varphi, \mathrm{dist}_2(y, \varphi \setminus \{y\}) > 3n$ and so, by the condition (3) in the definition of stabilization, the random variables $f(\Lambda_{Q_n})\mathbb{1}\{R(Q_n) < n\}$ and $g(\Lambda_{Q_n(-x)})\mathbb{1}\{R(Q_n(-x)) < n\}$ are independent. Thus, the first term appearing in the right-hand side of (A.0.3) becomes:

$$\begin{aligned}&\mathbb{E}\left[f(\Lambda_{Q_n})g(\Lambda_{Q_n(-x)})\mathbb{1}\{R(Q_n)<n\}\mathbb{1}\{R(Q_n(-x))<n\}\right]\\&=\mathbb{E}\left[f(\Lambda_{Q_n})\mathbb{1}\{R(Q_n)<n\}\right]\mathbb{E}\left[g(\Lambda_{Q_n(-x)})\mathbb{1}\{R(Q_n(-x))<n\}\right].\end{aligned}\tag{A.0.6}$$

Now, using the fact that $f(\Lambda_{Q_n}) =: \mathbb{E}\left(\mathbb{1}\{A\}\,|\,\Lambda\right)$ and noting that the event $\{R(Q_n) < n\}$ is Λ-measurable, we can put everything back into a single expectation and get:

$$\begin{aligned}\mathbb{E}\left[f(\Lambda_{Q_n})\mathbb{1}\{R(Q_n)<n\}\right]&=\mathbb{E}\left[\mathbb{E}\left(\mathbb{1}\{A\}\,|\,\Lambda\right)\mathbb{1}\{R(Q_n)<n\}\right]\\&=\mathbb{E}\left[\mathbb{E}\left(\mathbb{1}\{A\}\mathbb{1}\{R(Q_n)<n\}\,|\,\Lambda\right)\right]\\&=\mathbb{E}\left[\mathbb{1}\{A\}\mathbb{1}\{R(Q_n)<n\}\right]\\&=\mathbb{P}(A\cap\{R(Q_n)<n\}).\end{aligned}$$

In the same way, we get:

$$\mathbb{E}\left[g(\Lambda_{Q_n(-x)})\mathbb{1}\{R(Q_n(-x))<n\}\right]=\mathbb{P}(S_xB\cap\{R(Q_n(-x))<n\}).$$

Thus, (A.0.6) yields:

$$\begin{aligned}&\mathbb{E}\left[f(\Lambda_{Q_n})g(\Lambda_{Q_n(-x)})\mathbb{1}\{R(Q_n)<n\}\mathbb{1}\{R(Q_n(-x))<n\}\right]\\&=\mathbb{P}(A\cap\{R(Q_n)<n\})\mathbb{P}(S_xB\cap\{R(Q_n(-x))<n\})\\&=\mathbb{P}(A\cap\{R(Q_n)<n\})\mathbb{P}\circ S_x(B\cap\{R(Q_n)<n\})\\&=\mathbb{P}(A\cap\{R(Q_n)<n\})\mathbb{P}(B\cap\{R(Q_n)<n\}),\end{aligned}$$

where we have used the stationarity assumption to get the last line. Finally, using the fact that $\mathbb{P}(R(Q_n) < n) \geq 1 - \epsilon/3$, we get

$$|\mathbb{P}(A\cap\{R(Q_n)<n\})\mathbb{P}(B\cap\{R(Q_n)<n\})-\mathbb{P}(A)\mathbb{P}(B)|\leq\epsilon/3$$

and thus

$$\left|\mathbb{E}\left[f(\Lambda_{Q_n})g(\Lambda_{Q_n(-x)})\mathbb{1}\{R(Q_n)<n\}\mathbb{1}\{R(Q_n(-x))<n\}\right]-\mathbb{P}(A)\mathbb{P}(B)\right|\leq\epsilon/3.\tag{A.0.7}$$

Using (A.0.5), (A.0.7) to put everything back together in (A.0.3) and using the triangular inequality, we finally get:

$$|\mathbb{P}(A\cap S_xB)-\mathbb{P}(A)\mathbb{P}(B)|\leq 2\epsilon/3+\epsilon/3=\epsilon\tag{A.0.8}$$

for sufficiently large x, as required. This concludes the proof of Proposition A.0.1. □

Remark. Note that we actually did not need to use the PVT structure nor the asymptotic essential connectedness of the random support S here. We only used the fact that S is a stabilizing random tessellation and the complete independence properties of the point processes of users X^λ and of relays Y given their random support S. As a matter of fact, Proposition A.0.1 can be generalised to any stabilizing random tessellation S in $\mathbb{R}^2$.

Finally, since the percolation of the connectivity graph $\mathcal{G}$ is a translation-invariant event, a straightforward consequence of the previous result is the following 0-1 law:

Corollary A.0.1.1.

$$\forall L_A > 0, p \in [0,1], \lambda \geq 0, 0 \leq r' \leq r, \quad P(p, \lambda, r, r') \in \{0, 1\}.$$

In other words, percolation of the connectivity graph $\mathcal{G}$ is either an almost sure or almost impossible event.

Bibliography

[1] Nomenclature of the frequency and wavelength bands used in telecommunications. Technical Report ITU-T Recommendation B.15, International Telecommunications Union, October 1996.

[2] Global mobile data traffic forecast update, 2017–2022 white paper. Technical report, Cisco Visual Networking Index, 2019. Available at: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white-paper-c11-738429.html.

[3] Internet of things forecast. Technical report, Ericsson, 2019. Available at: https://www.ericsson.com/en/mobility-report/internet-of-things-forecast.

[4] D. Ahlberg, S. Griffiths, R. Morris, and V. Tassion. Quenched Voronoi percolation. *Advances in Mathematics*, 286:889–911, 2016.

[5] M. Alam, D. Yang, J. Rodriguez, and R. A. Abd-Alhameed. Secure Device-to-Device communication in LTE-A. *IEEE Communications Magazine*, 52(4):66–73, 2014.

[6] R. Alkurd, R. M. Shubair, and I. Abualhaol. Survey on Device-to-Device communications: Challenges and design issues. In *2014 IEEE 12th International New Circuits and Systems Conference (NEWCAS)*, pages 361–364. IEEE, 2014.

[7] A. Asadi, P. Jacko, and V. Mancuso. Modeling D2D communications with LTE and WiFi. *ACM SIGMETRICS Performance Evaluation Review*, 42(2):55–57, 2014.

[8] A. Asadi, Q. Wang, and V. Mancuso. A survey on Device-to-Device communication in cellular networks. *IEEE Communications Surveys & Tutorials*, 16(4):1801–1819, 2014.

[9] L. Babun, A. I. Yürekli, and I. Güvenç. Multi-hop and D2D communications for extending coverage in public safety scenarios. In *2015 IEEE 40th local computer networks conference workshops (LCN Workshops)*, pages 912–919. IEEE, 2015.

[10] F. Baccelli and B. Błaszczyszyn. Stochastic geometry and wireless networks: Volume I – Theory. *Foundations and Trends in Networking*, 3(3–4):249–449, 2009.

[11] F. Baccelli and B. Błaszczyszyn. Stochastic geometry and wireless networks: Volume II – Applications. *Foundations and Trends in Networking*, 4(1–2):1–312, 2010.

[12] F. Baccelli, B. Błaszczyszyn, and M. Karray. *Random Measures, Point Processes, and Stochastic Geometry*. 2020. Available at https://hal.inria.fr/hal-02460214/document.

[13] F. Baccelli, M. Klein, M. Lebourges, and S. Zuyev. Stochastic geometry and architecture of communication networks. *Telecommunication Systems*, 7(1-3):209–227, 1997.

[14] A. J. Baddeley, R. Turner, et al. Spatstat: An R package for analyzing spatial point pattens, 2004.

[15] I. Balberg. Recent developments in continuum percolation. *Philosophical Magazine B*, 56(6):991–1003, 1987.

[16] P. Balister and B. Bollobás. Percolation in the k-nearest neighbor graph. In *Recent Results in Designs and Graphs: a Tribute to Lucia Gionfriddo*, volume 28, pages 83–100. Quaderni di Matematica, 2013.

[17] P. Balister, B. Bollobás, and M. Walters. Continuum percolation with steps in the square or the disc. *Random Structures & Algorithms*, 26(4):392–403, 2005.

[18] P. Balister, A. Sarkar, and B. Bollobás. Percolation, connectivity, coverage and colouring of random geometric graphs. In *Handbook of Large-Scale Random Networks*, Bolyai Society Mathematical Studies 18, pages 117–142. Springer, 2008.

[19] A. M. Becker and R. M. Ziff. Percolation thresholds on two-dimensional Voronoi networks and Delaunay triangulations. *Physical Review E*, 80(4):041101, 2009.

[20] D. Beringer, G. Pete, Á. Timár, et al. On percolation, critical probabilities and unimodular random graphs. *Electronic Journal of Probability*, 22(106):1–26, 2017.

[21] B. Błaszczyszyn, M. Haenggi, P. Keeler, and S. Mukherjee. *Stochastic geometry analysis of cellular networks*. Cambridge University Press, 2018.

[22] B. Blaszczyszyn and D. Yogeshwaran. Connectivity in sub-Poisson networks. In *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pages 1466–1473. IEEE, 2010.

[23] B. Błaszczyszyn and D. Yogeshwaran. Clustering and percolation of point processes. *Electron. J. Probab*, 18(72):1–20, 2013.

[24] B. Bollobás, S. Janson, and O. Riordan. Line-of-sight percolation. *Combinatorics, Probability and Computing*, 18(1-2):83–106, 2009.

[25] B. Bollobás and O. Riordan. The critical probability for random Voronoi percolation in the plane is 1/2. *Probability Theory and Related Fields*, 136(3):417–468, 2006.

[26] B. Bollobás and O. Riordan. *Percolation*. Cambridge University Press, 2006.

[27] B. Bollobás and O. Riordan. Percolation on random Johnson–Mehl tessellations and related models. *Probability Theory and Related Fields*, 140(3-4):319–343, 2008.

[28] S. R. Broadbent and J. M. Hammersley. Percolation processes: I. Crystals and mazes. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 53, pages 629–641. Cambridge University Press, 1957.

[29] E. Cali, N. N. Gafur, C. Hirsch, B. Jahnel, T. En-Najjary, and R. I. Patterson. Percolation for D2D networks on street systems. In *2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 1–6. IEEE, 2018.

[30] Y. Cao, T. Jiang, and C. Wang. Cooperative Device-to-Device communications in cellular networks. *IEEE Wireless Communications*, 22(3):124–129, 2015.

[31] S. N. Chiu, D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic geometry and its applications*. John Wiley & Sons, 2013.

[32] T. Courtat. *Walk on City Maps - Mathematical and Physical phenomenology of the City, a Geometrical approach*. PhD thesis, 2012.

[33] T. Courtat, L. Decreusefond, and P. Martins. Stochastic simulation of urban environments. Application to path-loss in wireless systems. *arXiv preprint arXiv:1604.00688*, 2016. Available at https://arxiv.org/pdf/1912.07895.pdf.

[34] T. Courtat, S. Douady, and C. Gloaguen. Centrality maps and the analysis of city street networks. In *Proceedings of the 5th International ICST Conference on Performance Evaluation Methodologies and Tools*, pages 316–321, 2011.

[35] T. Courtat, C. Gloaguen, and S. Douady. Mathematics and morphogenesis of cities: A geometrical approach. *Physical Review E*, 83(3):036106, 2011.

[36] T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[37] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume I, Elementary Theory and Methods*. Springer, 2003.

[38] D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes: Volume II, General Theory and Structure*. Springer, 2008.

[39] O. Dousse, F. Baccelli, and P. Thiran. Impact of interferences on connectivity in ad hoc networks. *IEEE/ACM Transactions on networking*, 13(2):425–436, 2005.

[40] O. Dousse, M. Franceschetti, N. Macris, R. Meester, and P. Thiran. Percolation in the signal to interference ratio graph. *Journal of applied probability*, 43(2):552–562, 2006.

[41] G. D. Durgin, T. S. Rappaport, and H. Xu. Partition-based path loss analysis for in-home and residential areas at 5.85 GHz. In *IEEE GLOBECOM 1998 (Cat. NO. 98CH36250)*, volume 2, pages 904–909. IEEE, 1998.

[42] V. Erceg, L. J. Greenstein, S. Y. Tjandra, S. R. Parkoff, A. Gupta, B. Kulic, A. A. Julius, and R. Bianchi. An empirically based path loss model for wireless channels in suburban environments. *IEEE Journal on selected areas in communications*, 17(7):1205–1211, 1999.

[43] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[44] D. Feng, L. Lu, Y. Yuan-Wu, G. Y. Li, S. Li, and G. Feng. Device-to-Device communications in cellular networks. *IEEE Communications Magazine*, 52(4):49–55, 2014.

[45] M. Franceschetti, L. Booth, M. Cook, R. Meester, and J. Bruck. Percolation in multi-hop wireless networks. Technical Report UCB/ERL M03/18, EECS Department, University of California, Berkeley, 2003. Available at http://www2.eecs.berkeley.edu/Pubs/TechRpts/2003/4088.html.

[46] M. Franceschetti, O. Dousse, N. David, and P. Thiran. Closing the gap in the capacity of wireless networks via percolation theory. *IEEE Transactions on Information Theory*, 53(3):1009–1018, 2007.

[47] M. Franceschetti and R. Meester. *Random Networks for Communication: From Statistical Physics to Information Systems*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2008.

[48] A. Frieze, J. Kleinberg, R. Ravi, and W. Debany. Line-of-sight networks. *Combinatorics, Probability and Computing*, 18(1-2):145–163, 2009.

[49] H. Frisch and J. Hammersley. Percolation processes and related topics. *Journal of the society for industrial and applied mathematics*, 11(4):894–918, 1963.

[50] Z. Galil and G. F. Italiano. Data structures and algorithms for disjoint set union problems. *ACM Computing Surveys (CSUR)*, 23(3):319–344, 1991.

[51] P. Gandotra and R. K. Jha. Device-to-Device communication in cellular networks: A survey. *Journal of Network and Computer Applications*, 71:99–117, 2016.

[52] P. Gandotra, R. K. Jha, and S. Jain. A survey on Device-to-Device (D2D) communication: Architecture and security issues. *Journal of Network and Computer Applications*, 78:9–29, 2017.

[53] S. Ghosh, M. Krishnapur, Y. Peres, et al. Continuum percolation for Gaussian zeroes and Ginibre eigenvalues. *The Annals of Probability*, 44(5):3357–3384, 2016.

[54] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[55] E. N. Gilbert. Random plane networks. *Journal of the Society for Industrial and Applied Mathematics*, 9(4):533–543, 1961.

[56] I. Glauche, W. Krause, R. Sollacher, and M. Greiner. Continuum percolation of wireless ad hoc communication networks. *Physica A: Statistical Mechanics and its Applications*, 325(3-4):577–600, 2003.

[57] C. Gloaguen. Stochastic modelling of urban access networks. In *Proc. 10th Internat. Telecommun. Network Strategy Planning Symp.(Munich, June 2002)*, pages 99–104. VDE, 2002.

[58] C. Gloaguen, F. Fleischer, H. Schmidt, and V. Schmidt. Fitting of stochastic telecommunication network models via distance measures and Monte-Carlo tests. *Telecommunication Systems*, 31(4):353–377, 2006.

[59] C. Gloaguen, F. Voss, and V. Schmidt. Parametric distance distributions for fixed access network analysis and planning. In *2009 21st International Teletraffic Congress*, pages 1–8. IEEE, 2009.

[60] C. Gloaguen, F. Voss, and V. Schmidt. Parametric distributions of connection lengths for the efficient analysis of fixed access networks. *A nnals of telecommunications - Annales des télécommunications*, 66(1-2):103–118, 2011.

[61] A. Goldsmith. *Wireless communications*. Cambridge University Press, 2005.

[62] G. Grimmett. *Percolation*. Springer, 1999.

[63] N. Gupta, A. D. Singh, P. Shrivastava, and V. A. Bohara. A two-way cooperative D2D communication framework for a heterogeneous cellular network. *Wireless Personal Communications*, 109(1):579–593, 2019.

[64] M. Haenggi. *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.

[65] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti. Stochastic geometry and random graphs for the analysis and design of wireless networks. *IEEE journal on selected areas in communications*, 27(7):1029–1046, 2009.

[66] M. Haenggi, R. K. Ganti, et al. Interference in large wireless networks. *Foundations and Trends in Networking*, 3(2):127–248, 2009.

[67] P. Hall. On continuum percolation. *The Annals of Probability*, pages 1250–1266, 1985.

[68] J. M. Hammersley. Percolation processes: Lower bounds for the critical probability. *The Annals of Mathematical Statistics*, 28(3):790–795, 1957.

[69] T. E. Harris. A lower bound for the critical probability in a certain percolation process. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 56, pages 13–20. Cambridge University Press, 1960.

[70] C. Hirsch, B. Jahnel, and E. Cali. Continuum percolation for Cox point processes. *Stochastic Processes and their Applications*, 129(10):3941–3966, 2019.

[71] P. Jacquet and D. Popescu. Self-similar geometry for ad-hoc wireless networks: Hyperfractals. In *International Conference on Geometric Science of Information*, pages 838–846. Springer, 2017.

[72] P. Jacquet and D. Popescu. Self-similarity in urban wireless networks: Hyperfractals. In *2017 15th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, pages 1–6. IEEE, 2017.

[73] B. Jahnel and W. König. *Probabilistic Methods in Telecommunications*. Compact Textbooks in Mathematics. Springer International Publishing, 2020.

[74] B. Jahnel and A. Tóbiás. SINR percolation for Cox point processes with random powers. *arXiv preprint arXiv:1912.07895*, 2019. Available at https://arxiv.org/pdf/1912.07895.pdf.

[75] B. Jahnel, A. Tóbiás, and E. Cali. Phase transitions for the Boolean model of continuum percolation for Cox point processes. *arXiv preprint arXiv:2003.06206*, 2020. Available at https://arxiv.org/pdf/2003.06206.pdf.

[76] F. Jameel, Z. Hamid, F. Jabeen, S. Zeadally, and M. A. Javed. A survey of Device-to-Device communications: Research issues and challenges. *IEEE Communications Surveys & Tutorials*, 20(3):2133–2168, 2018.

[77] S. Jansen et al. Continuum percolation for Gibbsian point processes with attractive interactions. *Electronic Journal of Probability*, 21, 2016.

[78] O. Kallenberg. *Random measures*. Akademie-Verlag, 1983.

[79] O. Kallenberg. *Random measures, theory and applications*. Springer, 2017.

[80] U. N. Kar and D. K. Sanyal. An overview of Device-to-Device communication in cellular networks. *ICT Express*, 4(4):203 – 208, 2018.

[81] A. Kechris. *Classical descriptive set theory*, volume 156. Springer Science & Business Media, 2012.

[82] W. S. Kendall, G. Last, and I. S. Molchanov. New perspectives in stochastic geometry. *Oberwolfach Reports*, 5(4):2655–2702, 2009.

[83] H. Kesten. *Percolation theory for mathematicians*, volume 423. Springer, 1982.

[84] H. Kesten et al. The critical probability of bond percolation on the square lattice equals 1/2. *Communications in mathematical physics*, 74(1):41–59, 1980.

[85] J. Kingman. *Poisson Processes*. Oxford Studies in Probability. Oxford University Press, 1993.

[86] G. Last and M. Penrose. *Lectures on the Poisson process*, volume 7. Cambridge University Press, 2017.

[87] Q. Le Gall, B. Błaszczyszyn, E. Cali, and T. En-Najjary. The Influence of Canyon Shadowing on Device-to-Device Connectivity in Urban Scenario. In *2019 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–7, 2019.

[88] Q. Le Gall, B. Błaszczyszyn, E. Cali, and T. En-Najjary. Relay-assisted Device-to-Device Networks: Connectivity and Uberization Opportunities. In *2020 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–7, 2020. To appear.

[89] Q. Le Gall, B. Błaszczyszyn, E. Cali, and T. En-Najjary. Continuum Line-of-Sight Percolation on Poisson-Voronoi Tessellations. *arXiv preprint arXiv:1904.10875*, 2019. Submitted. Available at https://arxiv.org/pdf/1904.10875.pdf.

[90] W. C. Lee. *Mobile communications design fundamentals*, volume 25. John Wiley & Sons, 2010.

[91] T. M. Liggett, R. H. Schonmann, and A. M. Stacey. Domination by product measures. *The Annals of Probability*, 25(1):71–95, 1997.

[92] X. Lin, J. G. Andrews, and A. Ghosh. Spectrum sharing for Device-to-Device communication in cellular networks. *IEEE Transactions on Wireless Communications*, 13(12):6727–6740, 2014.

[93] K. R. Liu, A. K. Sadek, W. Su, and A. Kwasinski. *Cooperative communications and networking*. Cambridge university press, 2009.

[94] R. Löffler. Percolation phase transitions for the SIR model with random powers. *arXiv preprint arXiv:1908.07375*, 2019. Available at https://arxiv.org/pdf/1908.07375.pdf.

[95] R. Lyons. Random walks and percolation on trees. *The Annals of Probability*, pages 931–958, 1990.

[96] R. Ma, Y.-J. Chang, H.-H. Chen, and C.-Y. Chiu. On relay selection schemes for relay-assisted D2D communications in LTE-A systems. *IEEE Transactions on Vehicular Technology*, 66(9):8303–8314, 2017.

[97] V. H. Mac Donald. Advanced mobile phone service: The cellular concept. *The Bell system technical Journal*, 58(1):15–41, 1979.

[98] G. Matheron. *Random sets and integral geometry*. Wiley, 1975.

[99] R. Meester and R. Roy. *Continuum percolation*. Cambridge University Press, 1996.

[100] S. Mertens and C. Moore. Continuum percolation thresholds in two dimensions. *Physical Review E*, 86(6):061109, 2012.

[101] I. Molchanov and I. S. Molchanov. *Theory of random sets*, volume 19. Springer, 2005.

[102] J. Møller. Random tessellations in $\mathbb{R}^d$. *Advances in Applied Probability*, 21(1):37–73, 1989.

[103] J. Møller. *Lectures on random Voronoi tessellations*, volume 87. Springer Science & Business Media, 2012.

[104] L. Muche. The Poisson-Voronoi Tessellation: III. Miles' formula. *Mathematische Nachrichten*, 191(1):247–267, 1998.

[105] M. Newman and R. M. Ziff. Efficient Monte-Carlo algorithm and high-precision results for percolation. *Physical Review Letters*, 85(19):4104, 2000.

[106] M. E. Newman and R. M. Ziff. Fast Monte-Carlo algorithm for site or bond percolation. *Physical Review E*, 64(1):016706, 2001.

[107] H. Nishiyama, M. Ito, and N. Kato. Relay-by-smartphone: realizing multihop device-to-device communications. *IEEE Communications Magazine*, 52(4):56–65, 2014.

[108] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu. *Spatial tessellations: concepts and applications of Voronoi diagrams*, volume 501. John Wiley & Sons, 2009.

[109] G. K. Pedersen. The existence and uniqueness of the Haar integral on a locally compact topological group. *Preprint, University of Copenhagen, November*, pages 2004–2, 2000.

[110] R. Peierls. On Ising's model of ferromagnetism. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 32, pages 477–481. Cambridge University Press, 1936.

[111] G. Pike and C. Seager. Percolation and conductivity: A computer study. I. *Physical review B*, 10(4):1421, 1974.

[112] T. S. Rappaport et al. *Wireless communications: principles and practice*. Prentice Hall PTR New Jersey, second edition, 1996.

[113] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez. Millimeter-Wave mobile communications for 5G cellular: It will work! *IEEE Access*, 1:335–349, 2013.

[114] D. Ring. Mobile telephony-wide area coverage - Case 20564. *Bell Telephony Lab. Tech. Memoranda*, 47:1–22, 1947.

[115] L. Russo. A note on percolation. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 43(1):39–48, 1978.

[116] M. Schmittner, A. Asadi, and M. Hollick. Semud: Secure multi-hop Device-to-Device communication for 5G public safety networks. In *2017 IFIP Networking Conference (IFIP Networking) and Workshops*, pages 1–9. IEEE, 2017.

[117] R. Schneider and W. Weil. *Stochastic and integral geometry*. Springer Science & Business Media, 2008.

[118] C. Seager and G. Pike. Percolation and conductivity: A computer study. II. *Physical Review B*, 10(4):1435, 1974.

[119] P. D. Seymour and D. J. Welsh. Percolation probabilities on the square lattice. In *Annals of Discrete Mathematics*, volume 3, pages 227–245. Elsevier, 1978.

[120] E. Spodarev. *Stochastic Geometry, Spatial Statistics and Random Fields: Asymptotic Methods*, volume 2068. Springer, 2013.

[121] D. Stauffer and A. Aharony. *Introduction to percolation theory*. CRC press, 1994.

[122] K. Stucki. Continuum percolation for Gibbs point processes. *Electronic communications in probability*, 18, 2013.

[123] R. E. Tarjan and J. Van Leeuwen. Worst-case analysis of set union algorithms. *Journal of the ACM (JACM)*, 31(2):245–281, 1984.

[124] V. Tassion. Crossing probabilities for Voronoi percolation. *The Annals of Probability*, 44(5):3385–3398, 2016.

[125] M. N. Tehrani, M. Uysal, and H. Yanikomeroglu. Device-to-Device communication in 5G cellular networks: Challenges, solutions, and future directions. *IEEE Communications Magazine*, 52(5):86–92, 2014.

[126] D. Tse and P. Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[127] R. Turner. deldir: Delaunay triangulation and Dirichlet (Voronoi) tessellation. *R package version 0.0-8. URL* `http://cran.r-project.org/web/packages/deldir`, 2009.

[128] A. Tóbiás. Signal-to-interference ratio percolation for Cox point processes. *Latin American Journal of Probability and Mathematical Statistics*, 17:273, 01 2020.

[129] M. Q. Vahidi-Asl and J. C. Wierman. First-passage Percolation on the Voronoi Tessellation and Delaunay Triangulation. In *Random graphs*, volume 87, pages 341–359, 1990.

[130] F. Voss, C. Gloaguen, F. Fleischer, and V. Schmidt. Distributional properties of Euclidean distances in wireless networks involving road systems. *IEEE Journal on Selected Areas in Communications*, 27(7):1047–1055, 2009.

[131] S. Wang, W. Guo, Z. Zhou, Y. Wu, and X. Chu. Outage probability for multi-hop D2D communications with shortest path routing. *IEEE Communications Letters*, 19(11):1997–2000, 2015.

[132] V. Weiss and R. Cowan. Topological relationships in spatial tessellations. *Advances in Applied Probability*, 43(4):963–984, 2011.

[133] S. Wen, X. Zhu, Y. Lin, Z. Lin, X. Zhang, and D. Yang. Achievable transmission capacity of relay-assisted Device-to-Device (D2D) communication underlay cellular networks. In *2013 IEEE 78th Vehicular Technology Conference (VTC Fall)*, pages 1–5. IEEE, 2013.

[134] Y. Wu, W. Guo, H. Yuan, L. Li, S. Wang, X. Chu, and J. Zhang. Device-to-Device meets LTE-unlicensed. *IEEE Communications Magazine*, 54(5):154–159, 2016.

[135] F. Xue and P. R. Kumar. The number of neighbors needed for connectivity of wireless networks. *Wireless networks*, 10(2):169–181, 2004.

[136] S. Ziesche. Bernoulli Percolation on random Tessellations. *arXiv preprint arXiv:1609.04707*, 2016. Available at https://arxiv.org/pdf/1609.04707.pdf.

[137] S. Ziesche. First passage percolation in Euclidean space and on random tessellations. *arXiv preprint arXiv:1611.02005*, 2016. Available at https://arxiv.org/pdf/1611.02005.pdf.

RÉSUMÉ

La cinquième génération de réseaux mobiles devrait être en mesure de servir un nombre jamais vu d'appareils sur de vastes étendues. Un des principaux paradigmes étudiés pour répondre à ce défi est celui des communications Device-to-Device (D2D), c'est à dire de communications directes et de courte portée entre les appareils d'un réseau. Un cas d'usage d'intérêt économique significatif pour les opérateurs est celui de l'übérisation des réseaux : grâce au D2D, un nouvel opérateur n'ayant pas (ou presque pas) d'infrastructures réseau pourrait construire un réseau mobile fonctionnel reposant uniquement sur des terminaux mobiles.

Dans cette thèse, nous nous intéressons à de nouveaux modèles mathématiques de réseaux D2D en environnement urbain. Nous modélisons la voirie d'une ville par une mosaïque de Poisson-Voronoi plane. Les utilisateurs du réseau sont modélisés par un processus ponctuel de Cox sur les arêtes de cette mosaïque tandis que des relais supplémentaires sont modélisés par un processus ponctuel de Bernoulli sur ses sommets. Le réseau D2D est alors représenté par un graphe de connectivité dont les sommets sont les atomes des deux processus ponctuels précédents et où les connexions sont possibles uniquement entre des nœuds du réseau situés sur la même arête ou sur deux arêtes incidentes de la mosaïque de Poisson-Voronoi. Nous interprétons la percolation de ce graphe aléatoire (c'est-à-dire une probabilité positive d'existence d'une composante connexe infinie) comme signe d'une bonne connectivité du réseau. A l'aide de techniques de renormalisation, nous prouvons l'existence de transitions de phases entre différents régimes de connectivité : ceux où la percolation peut être assurée seulement par les relais ou, a contrario, ceux où une densité suffisante d'utilisateurs est nécessaire. A l'aide de simulations numériques et de nouveaux algorithmes de détection de chemins, nous estimons des paramètres critiques (par exemple la densité minimale d'utilisateurs) permettant une connectivité à grande échelle du réseau. Enfin, nous introduisons des modèles de coûts et utilisons les estimations précédentes pour étudier la faisabilité économique de scénarios d'übérisation des réseaux de télécommunications.

MOTS CLÉS

Communication Device-to-Device, Percolation, Crowd-networking, Modèles de voirie, Übérisation des réseaux, Géométrie stochastique

ABSTRACT

The fifth generation of cellular networks is expected to provide coverage for an unprecedented number of devices over large areas. One of the main paradigms investigated to address this challenge, called Device-to-Device (D2D) communication, consists in allowing for short-range direct communications between network devices. An application of significant economic interest for operators is the one of the uberisation of networks, where an operator having no (or very few) network infrastructure could build a mobile network relying only on its end-devices (users).

In this thesis, we study new mathematical models of D2D networks in urban environments. We see the street system of a city as a planar Poisson-Voronoi tessellation (PVT). Network users are given by a Cox process supported by the edges of the PVT while additional network relays are given by a Bernoulli process on the vertices of the PVT. The network is then modelled by a connectivity graph as follows: vertices are the atoms of both these processes and fixed-range connections between them possible only along the PVT edges or between network nodes located on adjacent PVT edges. Percolation of this random graph (existence of an infinite connected component with positive probability) is interpreted as good connectivity of the network. Using renormalisation techniques, we prove the existence of phase transitions between different connectivity regimes, in particular those where percolation can be solely ensured by the relays or, on the contrary, where a sufficient density of users is essential. Performing numerical simulations with original path-finding algorithms, we estimate critical parameters (e.g. the density of relays and users) allowing for good connectivity of the network. Finally, we also introduce appropriate cost models and use our numerical estimates to study the economic feasibility of uberisation scenarios of telecommunications networks.

KEYWORDS

Device-to-Device communication, Percolation, Crowd-networking, Street systems, Uberisation of networks, Stochastic geometry