

Use of Device-to-Device communications for efficient cellular networks

Rita Ibrahim

▶ **To cite this version:**

Rita Ibrahim. Use of Device-to-Device communications for efficient cellular networks. Signal and Image processing. Université Paris Saclay (COmUE), 2019. English. NNT : 2019SACLC002 . tel-02050989

HAL Id: tel-02050989
https://theses.hal.science/tel-02050989

Submitted on 27 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

université PARIS-SACLAY

CentraleSupélec

NNT: 2019SACLC002

Use of Device-to-Device Communications for Efficient Cellular Networks

Thèse de doctorat de l'Université Paris-Saclay
préparée à CentraleSupélec

École doctorale n° 580 Sciences et technologies de l'information et de la communication (STIC)
Spécialité de doctorat: réseaux, information et communications

Thèse présentée et soutenue à Gif-sur-Yvette, le 4 Février 2019, par

Mme Rita Ibrahim

Composition du jury:

M. Raymond Knopp Professeur, Eurecom	Président
Mme. Elza Erkip Professeur, NYU Tandon School of Engineering	Rapporteur
M. Dieter Fiems Professeur Associé, Ghent University	Rapporteur
M. Thomas Bonald Professeur, Telecom ParisTech	Examinateur
M. Georgios Paschos Ingénieur de recherche, Huawei Technologies	Examinateur
M. Mohamad Assaad Professeur, CentraleSupélec	Directeur de Thèse
Mme Berna Sayrac Ingénieure de recherche, Orange	Co-encadrante de Thèse

Thèse de doctorat

"Humans are free beings and are able to not only transform the world themselves, but to cooperate in order to transform the world in more sophisticated and helpful ways" K. Marx

Abstract

This thesis considers Device-to-Device (D2D) communications as a promising technique for enhancing future cellular networks. Modeling, evaluating and optimizing D2D features are the fundamental goals of this thesis and are mainly achieved using the following mathematical tools: queuing theory, Lyapunov optimization and Partially Observed Markov Decision Process (POMDP). The findings of this study are presented in three parts.

In the first part, we investigate a D2D mode selection scheme. We derive the queuing stability regions of both scenarios: pure cellular networks and D2D-enabled cellular networks. Comparing both scenarios leads us to elaborate a D2D vs cellular mode selection design that improves the capacity of the network.

In the second part, we develop a D2D resource allocation algorithm. We observe that D2D users are able to estimate their local Channel State Information (CSI), however the base station needs some signaling exchange to acquire this information. Based on the D2D users' knowledge of their local CSI, we provide an energy efficient scheduling framework that shows how distributed approach outperforms centralized one. In the distributed approach, collisions may occur between the different CSI reporting; thus a collision reduction algorithm is proposed. We give a description on how both centralized and distributed algorithms can be implemented in practice.

In the third part, we propose a mobile relay selection policy in a D2D relay-aided network. Relays' mobility appears as a crucial challenge for defining the strategy of selecting the optimal D2D relays. The problem is formulated as a constrained POMDP which captures the dynamism of the relays and aims to find the optimal relay selection policy that maximizes the performance of the network under cost constraints. We develop a system-level simulator to corroborate our claims.

Keywords: Device-to-Device (D2D) communications, cellular networks, mode selection, resource allocation, relay selection, queuing theory, Lyapunov optimization, Partially Observed Markov Decision Process (POMDP).

Acknowledgement

The underlying thesis would not have been possible to complete without receiving the support and help of many people. First, I would like to express my sincerest gratitude to my supervisors Berna Sayrac and Mohamad Assaad. I am extremely fortunate to work with my Orange Labs supervisor Berna who is now a very good friend. I would like to truly thank her for always being there to advice, support and encourage me with all her kindness and generosity. I had the honor and pleasure to work with her. I am grateful to my academic supervisor Mohamad for his relevant guidance and his availability. I am pleased to do my thesis under his supervision and to learn from his scientific rigor.

I would like to thank all the members of Orange Labs RIDE team for this exceptional experience among them. I would like to especially thank my manager Pierre Dubois for all his efforts during these three years of PhD, for being highly supportive and for having confidence in me. I am really lucky to have a great environment to come to every day thanks to my dear PhD colleagues: Stefan, Thomas, Mohamad, Julien, Tony, Felipe, Fatma, Marie, Romain, Ayat, Nivine, Imène, Abdulaziz, Ahlem. Thank you for all the great moments we spent together.

I would like to express my gratitude to the jury members of my PhD defense: Elza Erkip, Dieter Fiems, Raymond Knopp, Thomas Bonald and Georgios Paschos for their valuable review of my research work.

I would certainly like to thank all my friends for giving me the courage throughout these three years of PhD. You have greatly simplified the accomplishment of this challenging mission.

Last but not least, I want to thank my little family who constantly supported me with love and understanding. Very special recognition to my mother Najibeh and my two sisters Lola and Farah for always believing in me and to whom I would like to dedicate this thesis. Thank you for everything.

Contents

Abstract **i**

Acknowledgement **iii**

Acronyms **vii**

Nomenclature **xi**

List of Figures **xii**

List of Tables **xiv**

1 Introduction **1**
1.1 Background and motivation 1
1.2 Contributions 9
1.3 Organization 11

2 Theoretical Background **13**
2.1 Queuing Theory: Stability analysis 13
2.2 Lyapunov optimization 14
2.3 POMDP 17
2.4 Submodularity 21

3 D2D Mode Selection **23**
3.1 Concept and Related Work 23
3.2 System Model 27
3.3 Three-Users scenario 33
3.4 Multi-Users scenario 43
3.5 Numerical Results 50

3.6 Conclusion 55

4 D2D Resource Allocation **57**
4.1 Concept and Related Work 57
4.2 System model 60
4.3 Centralized approach 63
4.4 Distributed approach 65
4.5 Probability of collision 69
4.6 Implementation 71
4.7 Numerical Results 74
4.8 Conclusion 78

5 Mobile D2D Relay Selection **79**
5.1 Concept and related work 79
5.2 System Model 83
5.3 Problem formulation 84
5.4 Relay Selection Policies 90
5.5 Extension to Multi player scenario 98
5.6 Numerical Results 101
5.7 Simulation Results 107
5.8 Conclusion 112

6 Conclusion and Perspective **113**
6.1 Conclusion 113
6.2 Perspectives 114

Résumé **117**

Appendices **119**
7.3 D2D Mode Selection Proofs 119
7.4 D2D Resource Allocation Proofs 151
7.5 Mobile D2D Relay Selection Proofs 158
7.6 Simulation Settings 163

List of Publications **171**

Bibliography **173**

Acronyms

5G	Fifth-Generation
4G	Fourth-Generation
eMBB	Enhanced Mobile BroadBand
mMTC	Massive Machine Type Communications
URLLC	Ultra-Reliable Low Latency Communications
NR	New Radio
MIMO	Multiple Input Multiple Output
LTE	Long Term Evolution
D2D	Device-to-Device
M2M	Machine-to-Machine
V2V	Vehicle-to-Vehicle
V2X	Vehicle-to-Everything
ProSe	Proximity Services
PS	Public Safety
IoT	Internet of Things
CSI	Channel State Information
HARQ	Hybrid Automatic Repeat Request

AMC	Adaptive Modulation Coding
TDD	Time Division Duplex
QoS	Quality of Service
FDD	Frequency Division Duplex
DTMC	Discrete Time Markov Chain
DRX	Discontinuous Reception
1D	One Dimensional
SS	Simple Scenario
SR	Stability Region
UB	Upper Bound
LFP	Linear-Fractional Programming
LP	Linear Problem
DL	Downlink
UL	Uplink
3GPP	3rd Generation Partnership Project
SNR	Signal-to-Noise Ratio
SINR	Signal-to-Interference-plus-Noise Ratio
OP	Optimization Problem
RB	Resource Block
RE	Resource Emplacement
PUCCH	Physical Uplink Control Channel
BS	Base Station
UB	Upper Bound

UE	User Equipment
RI	Rank Index
PMI	Precoding Matrix Indicator
CQI	Channel Quality Index
DCI	Downlink Control Information
AWGN	Additive White Gaussian Noise
QPSK	Quadrature Phase-Shift Keying
PDCCH	Physical Downlink Control Channel
PUSCH	Physical Uplink Shared Channel
RRC	Radio Resource Control
E-UTRA	Evolved Universal Mobile Telecommunications System Terrestrial Radio Access,
MANET	Mobile Ad-Hoc Networks
OFDMA	Orthogonal Frequency Division Multiple Access
SC-FDMA	Single Carrier Frequency Division Multiple Access
EC	Energy Consumption
EE	Energy Efficiency
MU	Master User Equipment
TTI	Time Transmission Interval
RMAB	Restless Markov Multi-armed Bandit
MDP	Markov Decision Process
CMDP	Constrained Markov Decision Process
POMDP	Partially Observed Markov Decision Process
CPOMDP	Constrained Partially Observed Markov Decision Process

PWLC Piecewise-Linear Convex

MILP Mixed-Integer Linear Program

PBVI Point-Based Value Iteration

CPBVI Constrained Point-Based Value Iteration

GCPBVI Greedy Constrained Point-Based Value Iteration

ISD Inter-Site Distance

UAV unmanned aerial vehicle

CDF Cumulative Distribution Function

Nomenclature

$\mathbb{E}$	Expectation operation
$\sum$	Sum
$\prod$	Product
$\|.\|$	Cardinal for the sets and Absolute value for the scalars
$\mathbb{1}_{\{\}}$	Indicator function
$:=$	Definition
$[x]^+$	Equal to x if $x > 0$ and 0 otherwise
$\oplus$	Cross-Sum
$\boldsymbol{a}.\boldsymbol{b}$	Scalar product of vector $\boldsymbol{a}$ and vector $\boldsymbol{b}$
$\|\|\boldsymbol{v}\|\|_p$	$\left(\sum_{i=1}^{n} \|v_i\|^p\right)^{1/p}$, p-norm of vector $\boldsymbol{v} = (v_1, v_2, ..., v_n)$
$\|\|\boldsymbol{v}\|\|_\infty$	$\max_{i=\{1,..,n\}} \|v_i\|$, infinity norm of vector $\boldsymbol{v} = (v_1, v_2, ..., v_n)$
$\lceil (x) \rceil$	Closet integer higher or equal to scalar x
$\mathbb{Z}$	Set of all integers $1, 2, ...$
$\mathbb{R}$	Set of real numbers
$\gamma(z)$	Gamma function $\gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$
$A \bigcup B$	Union of the two sets A and B
$A \bigcap B$	Intersection of the two sets A and B
$A \subseteq B$	A is subset of B
$n!$	Factorial of a non-negative integer n, $n! = 1 \times 2 \times 3... \times n$
$O(f(x))$	Limiting behavior of function $f(x)$ when x tends to infinity
$s.t.$	Subject to

List of Figures

1.1 Use Cases of D2D communications 4
1.2 Status of D2D communications in 3rd Generation Partnership Project (3GPP) standards 7

3.1 System Model 28
3.2 The queues model of both scenarios 30
3.3 System Model of the Three User Equipment (UE)s scenario 34
3.4 The upper Markov Chain is the exact modeling of the queue Q_{BS} and the lower Markov Chain is the approximated modeling of the queue Q_{BS} . . 40
3.5 Example of one priority policy for $K = 3$ and $U = 2$. 44
3.6 Comparison between the stability region for the cellular scenario $\mathcal{R}_c^{ss}$ and its approximation $\tilde{\mathcal{R}}_c^{ss}$ 51
3.7 Stability regions for both D2D and cellular scenarios for UE to Base Station (BS) distance d in {100, 200, 350, 500}m and two different D2D distance d^d m 52
3.8 Stability region of the cellular scenario (with coupling between the users' queues and the state empty or not of Q_{BS}) vs. rate region (without coupling hence Q_{BS} is considered always not empty) 53
3.9 Max D2D distance for which D2D outperforms cell. comm. 53

4.1 An example of D2D scenario where $N = 6$ 61
4.2 Example of channel indexing feedback for 4 users 67
4.3 The resource blocks allocated for users' feedback (Physical Uplink Control Channel (PUCCH) Format 1 and 2) 73
4.4 Uniform Random Localization of $N = 50$ pairs of D2D 75
4.5 EC as function of the Signal-to-Noise Ratio (SNR) threshold γ_{th} 76
4.6 EE as function of the SNR threshold γ_{th} 77

5.1 D2D relaying scenario for numerical results 101

5.2 Complexity study of the proposed relay selection policies 103
5.3 Comparison between the performance of following relay selection policies: (i) the Constrained Point-Based Value Iteration (CPBVI) policy (see 5.4.2) and (ii) the Greedy Constrained Point-Based Value Iteration (GCPBVI) policy (see 5.4.3) 104
5.4 For single Master User Equipment (MU), performance comparison between both scenarios: with and without D2D relaying 105
5.5 Complexity study of the proposed relay selection policies 106
5.6 Comparison between the performance of the centralized approach (see 5.5.1) and the distributed approach (see 5.5.2) of a multi-MU scenario . . 107
5.7 Flow Chart of D2D relaying procedures in the system level simulator . . 109
5.9 The spectral efficiency of Downlink (DL) communications as function of the number of users N for both scenarios 111
5.10 The relative enhancement of the spectral efficiency of DL communications as function of the number of users N 111

7.1 Markov Chain model of the queue Q_{BS} 120
7.2 Approximated Markov Chain model of the BS queue: $\tilde{Q}_{BS}$ 129
7.3 Markov chain model of each $Q_{i,BS}$ in the multi-UE scenario 142
7.4 The hexagonal grid with 7 tri-sector macro sites 164
7.5 The layouts of users drop that are implemented in the system-level simulator 164
7.8 Traffic generation of FTP Model 2 168
7.9 Cumulative Distribution Function (CDF) of the D2D Signal-to-Interference-plus-Noise Ratio (SINR) 169

List of Tables

3.1 Service rate parameters 41

4.1 Numerical Settings of Chapter 4 75

5.1 Numerical Settings 102

7.1 Notation to make expressions simpler 121
7.2 D2D Scenario: Probability of transmission over link 1 and 2 for different policies and different states of channel 135
7.3 The values of the belief state $\boldsymbol{b}^e$ and the probability $Pr\left(z|\boldsymbol{a},\boldsymbol{b}\right)$ of the relay e in the case where the action at epoch t contains the relay e 161
7.4 The values of the belief state $\boldsymbol{b}^e$ and the probability $Pr\left(z|\boldsymbol{a},\boldsymbol{b}\right)$ of the relay e in the case where the action at epoch t does not contain the relay e 162
7.5 Precoding vectors 167

1 | Introduction

In this chapter, we define the concept of Device-to-Device (D2D) communications and highlight the different use cases that benefit from the advantages of D2D technology. We describe how D2D communications is integrated to 3rd Generation Partnership Project (3GPP) standardization. The various technical challenges of D2D are outlined. Finally, the thesis contributions are summarized.

1.1 Background and motivation

The growth of mobile broadband traffic and the number of connected devices, the demand of higher data rate as well as the emergence of various new applications have been pushing the investigation of new architectures and paradigms to fulfill the requirements of future cellular networks. In order to achieve higher performance and new capabilities, Fifth-Generation (5G) technology of cellular networks aims to integrate all the means of accessing the internet (e.g. fiber, Fourth-Generation (4G), New Radio (NR), Wi-Fi). The key expectations of 5G networks are: high minimum throughput (almost 50Mbps), high capacity and data rates ($\times 3$ spectral efficiency compared to 4G), high reliability (up to 10^{-6}), low latency (1 to 10 ms latency), high energy efficiency (energy consumption divided by 2) etc. These expectations are required for enabling new generation of applications, services and opportunities (e.g. virtual reality, augmented reality, vehicle automation, industry 4.0 etc.). Due to network virtualization, the wide variety of 5G services will be delivered by the same network. There is mainly three families of 5G services: (i) Enhanced Mobile BroadBand (eMBB), (ii) Massive Machine Type Communications (mMTC), (iii) Ultra-Reliable Low Latency Communications (URLLC). At the level of the physical layer, the main levers considered to answer the ambitious goals of 5G networks are the following: wider bandwidth (400MHz to almost 1GHz), higher frequencies (mmWave spectrum) and higher number of antennas (massive Multiple Input Multiple Output (MIMO)).

1.1.1 D2D Concept

The 5G cellular networks are expected to meet these high demands by the use of different new paradigms including D2D communications. D2D communications is defined as direct communication between spatially close mobile users without the need of passing through the Base Station (BS). D2D communications is considered as a promising component for enhancing next generation mobile networks. In high data rate services, mobile users could potentially be close enough to have D2D communications that highly increase the spectral efficiency of the network. Since D2D technique enables the realization of massive ad hoc mesh networks, D2D was firstly introduced in 3GPP community as a competitive technology for public safety networks (i.e. especially when cellular coverage fails or is not available). Nevertheless, the advantages of D2D communications can be categorized in two main groups: (i) enhancing network performance and (ii) enabling new proximity services. D2D technique boosts the performance of the network by: improving energy and/or spectrum efficiency, increasing the throughput, reducing delays, extending network coverage and offloading network traffic etc. In addition, public safety services as well as different commercial and social proximity services are enabled thanks to D2D communications. The emergence of 5G networks pushes for revisiting D2D communications from a 5G network point of view. New research trends in D2D networks are arriving, especially: Vehicle-to-Vehicle (V2V) communications, UE-to-Network relaying for Internet of Things (IoT) and wearables, millimeter wave technology as well as social trust based networks. For detailed survey on D2D communications, one can refer to [1] and [2].

D2D communication can occur on cellular spectrum (i.e. *inband*) or unlicensed spectrum (i.e. *outband*). For inband case, D2D links can reuse cellular resources (i.e. *underlay*) causing D2D-cellular interference or D2D links can use dedicated cellular resources (i.e. *overlay*) to avoid interference. Majority of the literature considers underlay inband D2D and proposes some resource and/or power allocation algorithms as well as mode selection schemes to mitigate interference. The key advantage of inband D2D is improving spectrum efficiency and ensuring the Quality of Service (QoS). Outband D2D communications can be managed by the network (i.e. *controlled*) or can operate on their own (i.e. *autonomous*). Eliminating interference is the main motivation behind outband D2D. However, the uncontrolled nature of the unlicensed spectrum induces uncertain QoS and immoderate D2D-D2D interference. Even though the majority of the literature studies inband D2D, the high integration of smartphones (with multiple wireless interfaces) in cellular networks push researchers to explore outband D2D schemes.

1.1.2 D2D vs. Ad-Hoc Networks

Before broaching the D2D analysis, it is crucial to compare D2D to Mobile Ad-Hoc Networks (MANET)s, which have been largely investigated over about three decades (i.e. one can see [3] for more details). The main particularity of D2D is the existence of network infrastructure (i.e. BS) that provides the assistance for different control functions (e.g. resource allocation, session setup and synchronization etc.) which are highly expensive in MANET. Furthermore, D2D technique consists of local single or two hops communications that are enabled in situations where they are beneficial for cellular networks. However, multi-hop routing may be required in MANET which generates a degradation in the network performance. In public safety scenarios, infrastructure support may not be available and D2D may look more like a MANET. Whereas, in such scenarios, D2D is only used for providing elementary service (i.e. not like a full MANET which can be deployed for video streaming). Furthermore, in out-of-coverage scenarios, the devices are often clustered where the cluster head can play the role of BS. Although simpler than a MANET, adding D2D features to cellular networks still faces several challenges.

1.1.3 D2D Use Cases

D2D technique is involved in various use cases summarized in diagram 1.1. The main scenarios and services that profit from the advantages of D2D communications can be categorized into the following groups:

- **Proximity services**
  There exists interesting applications of D2D communications in location-aware services: both commercial and social ones. Among the new use cases of D2D communications for commercial proximity services, we enumerate the following: smart cities [4], Machine-to-Machine (M2M) communications, personalized services (local advertising, streaming services [5]), Vehicle-to-Everything (V2X) communications [6]. Several social proximity services can benefit from D2D communications such that: social networking [7], gaming, content sharing and video distribution [8].

- **Network enhancement**
  We describe some D2D scenarios that help the operators to enhance the performance of their cellular networks. First, user to network relaying is enabled due to D2D communications. The main advantage of this scenario is enhancing the capacity [9], extending the coverage of cellular networks [10] as well as reducing the energy consumption of low power devices (e.g. IoT and wearables). Second, D2D technique can be deployed in the following scenarios for offloading the network: content delivery [11], multicasting [12] and direct communications (voice, messages, content sharing). Third, the BS profits from D2D based localization to enhance the accuracy of users' positioning.Finally, the management of various cellular procedure, e.g. handover [13], can benefit from D2D technique.

- **Public safety services**
  D2D communications are expected to have an important role for public safety scenarios. In case of natural disaster, for example, cellular infrastructure can get damaged. Therefore, direct communcations can be setup between mobile users based on D2D communications in order to assure public protection, disaster relief and public safety services [14]. 3GPP has introduced D2D communications especially for public safety scenarios. However, the advantages of this technique seems to be also exploited in different other scenarios.

1.1.4 D2D Standardization

In addition to industrial and academic works, 3GPP introduced D2D communication in Release 12 especially for the Public Safety (PS) use case of Proximity Services (ProSe). [15] gives a brief overview of 3GPP standardization activities concerning D2D communications. In particular, 3GPP working groups have studied the feasibility of ProSe and its use cases (see [16]) as well as investigated the architectural and protocol enhancements required for the application of ProSe use cases (see [17]). The main D2D features that have been examined by 3GPP are: ProSe management, device discovery, synchronization, and direct communication etc. 3GPP activities in Release 12 (see [18]) were focused on basic discovery and communication procedures for PS scenarios (i.e. in coverage discovery and broadcast-based communication). Multiple enhancements were included in Release 13 and 14. The D2D features supported by 3GPP are summarized in 1.2. Sidelink is the term assigned to D2D communications in 3GPP Release 12. From a radio access perspective, the main design aspects that have been discussed in 3GPP are highlighted in the sequel. Higher layer operations, e.g. authentication, authorization, provisioning, lawful Interception and charging, are discussed in [16].

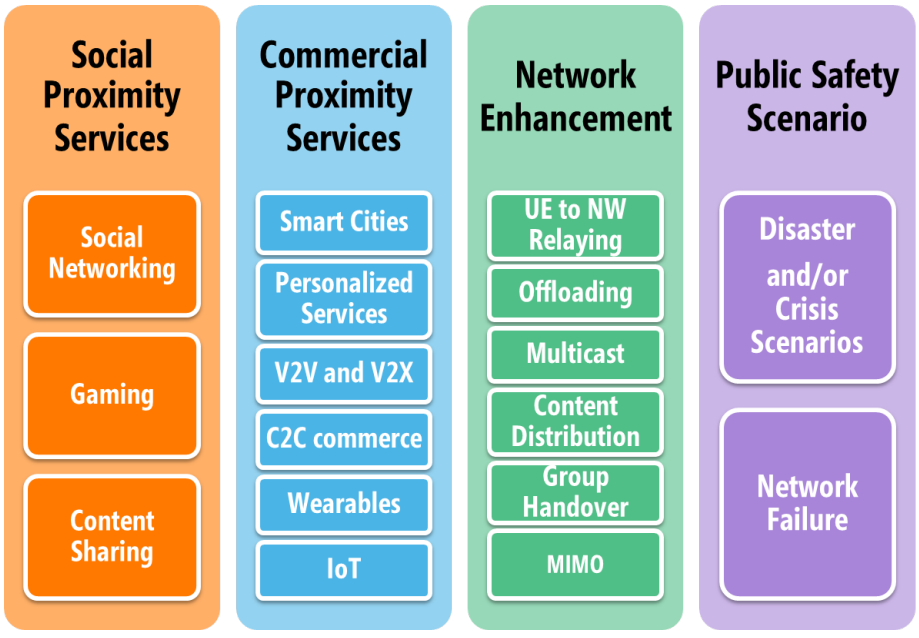


Figure 1.1: Use Cases of D2D communications

D2D Management

- The **Control Plane** of D2D communications is challenging since full network control is not guaranteed especially when devices enter an out-of-coverage area. In the latter case, two alternative topologies exist: ad hoc topology (i.e. each device controls its own behavior) and/or cluster head-based topology (i.e. a cluster head is assigned within a group of devices and plays the role of BS).
- The D2D **spectrum access** can be done either on unlicensed or licensed spectrum. While PS devices may have access to dedicated spectrum, commercial D2D services can share the radio resources with existing cellular devices. Moreover, the use of Uplink (UL) resources for D2D communications is more favorable than using Downlink (DL) resources for the following reasons. In existing Long Term Evolution (LTE) using UL (resp. DL) resources for D2D communications means that D2D UE should be equipped with Single Carrier Frequency Division Multiple Access (SC-FDMA) receiver (resp. Orthogonal Frequency Division Multiple Access (OFDMA) transmitter). In addition to regulatory concerns, the latter is more complicated in terms of hardware design. Furthermore, UL resources are often less utilized than DL resources.
- For D2D communications, two modes of **resource allocation** exist: (i) scheduled mode, called *mode 1*, where the BS schedules sidelink resources and (ii) autonomous mode, called *mode 2*, where each device selects its sidelink resources required for the transmission of sidelink data and sidelink control channel from a resource pool. This resource pool may be static (preconfigured) or dynamic (depending on D2D traffic). A D2D device can exceptionally switch between these two modes.

D2D Synchronization

D2D transmitter and receiver must be synchronized in time and frequency. In the case of in-coverage D2D, the synchronization is provided by the network. Nevertheless, the synchronization between the D2D users is not guaranteed since the devices of a D2D pair may have different timing advance adjustments (i.e. when devices are associated to two unsynchronized cells or when devices are in the same cell but at different distances from the BS). When devices are out-of-coverage, then they are asked to send periodic transmission of synchronization which is an energy consuming procedure.For sidelink communications, new synchronization channels and signals were introduced in [19].

D2D Discovery

The discovery procedure of detecting nearby devices is required for the different use cases of D2D communications. Two types of D2D discovery exist. The first type, called *type 1*, is direct discovery where the device selects autonomously radio resources from a pool of resources dedicated for discovery signals. Thus, this type of discovery is applicable to both in and out-of-coverage devices. In this case, D2D device participates periodically in the discovery process by transmitting or receiving discovery signals. For this type of discovery, two mechanisms are possible: (i) *Model A* where the device announces its presence and (ii) *Model B* where the device searches for discoverable devices. The second type is evolved packet core-level discovery, called *type 2*, where the discovery procedure starts after receiving information from the network. When permission is required for discovering a device then we talk about *restricted* discovery otherwise it is an *open* discovery.

D2D Communication

The D2D communication procedure consists of exchanging voice and data between nearby mobile users. The main D2D communication modes that have been defined are: unicast, groupcast, and broadcast. Existing LTE features (e.g., frequency parameters and frame structure) were recycled for D2D communications. However, many physical aspects need further efforts such that: modulation format, power control (open or closed loop), HARQ operation (indirect or direct), channel estimation (design of reference signals) etc.

UE-to-Network Relay

The ProSe User Equipment (UE) to network relay is defined in 3GPP Release 13 as follows: it is a device that provides layer 3 connectivity to other devices (called remote users). The relay user is always in-coverage which is not the case of the remote users. Therefore, this relay is able to extend the coverage of the network by relaying layer 3 traffic between the remote users and the network. A D2D communication is established between the relay and the remote user however existing cellular communications are used by the relay to connect to the network. The main operations that UE-to-Network relay will proceed are the following: 1. Relay UE Initiation: the BS specifies whether relay operation are or are not supported by this UE, 2. Relay Discovery: model A or B discovery are applied in order to discover nearby candidate relays, 3. Relay Selection: the selection of the relay is done based on the radio link quality of the candidate relays, however detailed criteria of relay selection require further studies, 4. Connection Establishment: in order to provide unicast traffic, the relay establishes 1-to-1 D2D connection with the remote user, 5. Reporting Remote UE information: once D2D link is established, the relay reports to the network some information concerning the remote UE (e.g. remote UE identifier, its IP address etc.).

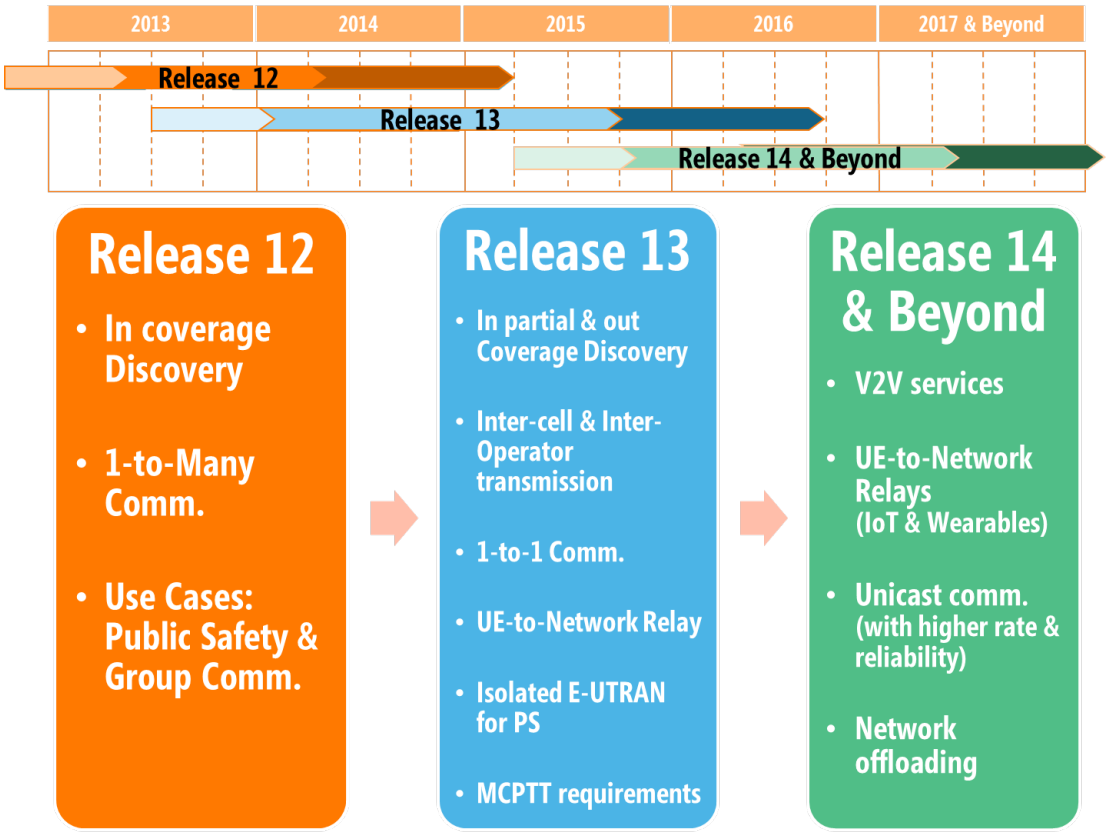


Figure 1.2: Status of D2D communications in 3GPP standards

1.1.5 D2D Challenges

The revealing promise of D2D communications has motivated the industry, academia and standardization bodies to explore the potentials of this technique and to work on solving its technical and business challenges in order to be implemented in cellular networks in the few next years. The major challenges faced by D2D communications are the following:

- **Technical challenges**: spectrum access (inband, outband), resource and/or power allocation, energy consumption, interference management, modulation format (OFDMA or SC-FDMA), channel measurements (Channel State Information (CSI) acquisition), Hybrid Automatic Repeat Request (HARQ) (robustness), signaling techniques (overhead), channel model (indoor, outdoor), architecture procedures (synchronization, discovery ...), privacy, authentication and security etc.
- **Business challenges**: inter-operators D2D communications, lawful interception, charging etc.

D2D communication is a promising technique with several open research issues that deserve further exploration in terms of simulations, demos, evaluation methods and analytical models. In this thesis, we mainly consider the following research problems:

Mode Selection: When a communications is occurring between nearby devices, an algorithm of mode selection is crucial for choosing between cellular mode (i.e. using traditional UL and DL communications) and D2D mode (i.e., the traffic is directly transmitted to the receiver without passing through the BS). Mode selection plays an important role in increasing the spectrum efficiency of cellular networks as well as mitigating the effect of interference between D2D and cellular users. Various criteria were considered when dealing with mode selection scheme such as: path loss [20], distance [21], channel quality of D2D and cellular links [22], interference [23], load of the BS [24], and energy efficiency [25]. An interesting way to select D2D vs. cellular mode is by not limiting the selection algorithm to a physical layer criteria but to take into account the state of the queues in the networks. Therefore, a queuing theory approach for dealing with mode selection schemes draws our attention.

Resource Allocation: Resource allocation is one of the most crucial features of D2D communications due to its advantage in improving the system performance in terms of power consumption, throughput, interference mitigation etc. The majority of the works in the literature have been proposing resource allocation techniques based on centralized approaches (e.g. [26]). However, these centralized approaches are very challenging in D2D-enabled cellular networks since a large number of devices will be required to feedback its CSI to the network. This CSI reporting highly increases the signaling overhead and is limited by the amount of resources available for CSI exchange. Therefore, distributed resource allocation schemes (e.g. [27]) were proposed to overcome the limitation of these centralized solutions. The comparison between centralized and distributed scheduling in terms of both performance and signaling overhead remain an open research question.

Relay Selection: In cooperative D2D networks, having an efficient and optimized relay selection design is a challenging issue. In general, in order to avoid the overload of the BS, the choice of which relay to select is done at the level of the D2D user (e.g. [28]). In this case, choosing the optimal relay is done in a distributed manner and with a passive role of the BS. In the other hand, some works in the literature (e.g. [29]) consider centralized relay selection methods for D2D-enabled networks. In most of the interesting practical cases, relays are in mobility, thus dynamic relay selection policies are necessary to investigate and interesting to explore.

1.2 Contributions

The main result of this thesis is presented in the upcoming three chapters. Here, we summarize the three main contributions of this thesis.

1.2.1 D2D Mode Selection

In chapter 3, we provide an analytical characterization of the cases where D2D outperforms traditional cellular communications. We address this problem of mode selection based on a queuing analysis, i.e. the stability region as the performance metric. When two nearby devices communicate with each other using tradition cellular network then a dependency is detected at the BS level between UL and DL queues (i.e. arrival of the DL queue corresponds to the departure of the UL queues). Indeed, a queuing based analysis is required for capturing the coupling that exists between the different queues in the cellular network. We have characterized the stability region of the following network scenarios: (i) pure cellular networks where the users can communicate with each other only by passing through the BS and (ii) overlay D2D-assisted networks where direct communication between the users are enabled. Comparing the performance of both scenarios leads to a distance based mode selection that improves the throughput of the network. The principle contributions of this chapter are the following:

- Characterization of the exact stability region of both scenarios as a convex polytope with a finite number of vertices and discussion of the complexity of these exact solutions.
- Proposition of approximated models and capturing the existing trade-off between their complexity and preciseness compared to exact stability regions.
- Comparison between the stability regions of both cellular and D2D scenarios and deducing a distance-based mode selection scheme (cellular vs. D2D) that improves the performance of cellular networks.

1.2.2 D2D Resource Allocation

As we mentioned before, resource allocation of D2D communications is still an open research issue especially when it comes to choose between centralized or distributed scheduling approaches. A channel adaptive D2D scheduling is crucial for improving the performance of D2D communications since it enables the transmitter to dynamically adapt its transmission scheme to the D2D channel quality. However, applying centralized scheduling requires that theBS is kept informed about the channel state of D2D links. This information is sent to the BS via CSI feedback transmissions. Therefore, centralized strategies will suffer from the limited amount of resources available for CSI exchange and will increase the signaling overhead. From that comes

the idea of having a distributed strategy that benefits from the users knowledge of their local D2D CSI. In chapter 4, we propose a resource allocation framework that aims to minimize the D2D users' transmission power while maintaining predefined throughput constraints. The main contributions of this chapter are summarized as follows:

- We propose an energy efficient centralized scheduling based on Lyapunov optimization and discuss how the performance of this algorithm is limited by the resources available for CSI exchange between D2D users and the BS.
- Using the D2D users' knowledge of their local CSI, we propose a new CSI reporting mechanism that leads us to deduce an energy efficient distributed scheduling. We show that this distributed approach outperforms the centralized one.
- Considering a distributed approach, collisions may occur during the transmission of CSI reporting messages, therefore we develop an algorithm that reduces this kind of collision.
- We describe in details how both distributed and centralized scheduling can be simply implemented in practice (i.e. mainly in existing LTE standards).

1.2.3 D2D Relay Selection

In D2D enabled cellular networks, an interesting scenario is user to network relaying based on D2D communications between the users and the relays and traditional cellular communications between the relays and the base station. In D2D cooperative cellular networks, relay selection is a challenging problem particularly when the relays are in mobility (e.g. pedestrian, vehicle, etc.). Since mobility of the relays is considered, then a dynamic relay selection policy is inescapable. Indeed, UE-to-Network relaying leads to important gains in terms of network performance (i.e. extending the coverage and enhancing the capacity) as well as enabling new services (e.g. serving the community). However, remote users are asked to be charged in a certain way in order to motivate the D2D users to play the role of relay. In chapter 5, we propose a dynamic relay selection policy that maximizes a certain performance metric of the network (i.e. throughput, reliability, coverage) under cost constraints of the relays (i.e. transmission power, data consumption, battery level). The fundamental results of this chapter are based on a simple scenario of one Master User Equipment (MU) selecting appropriate relays. Then, this result is extended to multiple MUs scenario. The findings of chapter 5 are outlined as follows:

- Formulation of the dynamic relay selection policy as a Constrained Partially Observed Markov Decision Process (CPOMDP).
- Discussing the complexity of the exact solution.

- Verifying the submodularity of the reward and cost value functions and proposition of a greedy point based value iteration solution that turns to well approximate the exact solution with a lower computational complexity.
- Extending the results to multiple MU scenario where both centralized and distributed solutions are exposed.
- Developing a system-level simulator to validate the proposed relay selection policy for both UL and DL scenarios.

1.3 Organization

The remainder of this thesis is organized as follows. In chapter 2, we introduce the mathematical tools used in this work: stability analysis in queuing theory, Lyapunov optimization, Partially Observed Markov Decision Process (POMDP), submodularity of set functions. In chapter 3, we study the stability region of both scenarios: pure cellular networks and overlay D2D-enabled cellular networks. We compare the analytical expressions of both scenarios in order to deduce a **mode selection** scheme that improves the capacity of the network. In chapter 4, based on Lyapunov optimization, a distributed **resource allocation** algorithm is developed. Since centralized approach suffers from the limited number of resources available for channel state feedback, this distributed scheduling turns out to outperform the centralized one. In chapter 5, we study the scenario of mobile UE-to-Network relaying. Based on a Constrained Partially Observed Markov Decision Process (CPOMDP) formulation of the problem, we find the optimal **relay selection** policy that maximizes the performance of the network under cost constraints. Chapter 6 concludes the thesis by giving a summary and some possible perspectives of the current work.

2 | Theoretical Background

In this chapter, we describe the mathematical tools that have been used in this thesis. A stability region analysis is considered in chapter 3. The study in chapter 4 is based on Lyapunov optimization. Finally, POMDP tool and submodularity properties are the main basis of chapter 5.

2.1 Queuing Theory: Stability analysis

Let $Q(t)$ represents the content of a discrete time queuing system defined over time slots $t \in \{0, 1, 2, ...\}$ (with the initial state $Q(0) > 0$). $\lambda(t)$ denoted the arrival process and $\mu(t)$ the service one, then the dynamic equation of the queue is the following:

$$Q(t+1) = [Q(t) - \mu(t)]^{+} + \lambda(t) \text{ for } t \in \{0, 1, 2, ...\} \tag{2.1}$$

In chapter 3, the considered definition of queue's stability is the following.

Definition 2.1.1. ***(Queue's stability)*** *A Queue $Q(t)$ is called stable if:*

$$\lim_{q \to \infty} \left[\lim_{t \to \infty} \mathbb{P}(Q(t) < q) \right] = 1 \tag{2.2}$$

The following [30] theorem proposes a characterization of the queue's stability when the arrival and service processes are assumed stationary.

Theorem 2.1.2. ***(Loynes Theorem)*** *When the arrival and service process of a queue are strictly jointly stationary then this queue is stable if its average arrival rate is less than its average service rate.*

A network of multiple queues $(Q_1(t), Q_2(t), ..., Q_N(t))$ is characterized by its stability region which corresponds to the region where all the queues in the system are stable. We respectively denote by λ_i and μ_i the averages of the arrival and service processes of queue $Q_i(t)$. Therefore, both the vector of arrival and service rate are given by:

$$\boldsymbol{\lambda} = [\lambda_1, \lambda_2, ..., \lambda_N] \tag{2.3}$$

$$\boldsymbol{\mu} = [\mu_1, \mu_2, ..., \mu_N] \tag{2.4}$$

Definition 2.1.3. ***(Stability region for a scheduling policy)*** *For a given scheduling policy ξ, the $\mathcal{R}(\xi)$ is defined as the closure set of arrival rate vectors for which all the queues are stables:*

$$\mathcal{R}(\xi) := \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^N | \boldsymbol{\lambda} \prec \boldsymbol{\mu}(\xi) \right\}$$

where $\prec$ is the component-wise inequality, $\boldsymbol{\lambda}$ is the vector of the queues' average arrival rates and $\boldsymbol{\mu}(\xi)$ is the vector of the queues' average service rates under the scheduling policy ξ.

Definition 2.1.4. ***(Stability Region)*** *The stability region $\mathcal{R}$ is defined as the union of the stability region for all the feasible scheduling policies (denoted by Ω_ξ).*

$$\mathcal{R} := \bigcup_{\xi \in \Omega_\xi} \mathcal{R}(\xi) \Rightarrow$$

$$\mathcal{R} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^N \mid \exists \left[\gamma_1, ..., \gamma_{|\Omega_\xi|} \right] \in [0,1]^{\Omega_\xi} \ \ with \ \sum_{i=1}^{\Omega_\xi} \gamma_i \leq 1 \, , \, \boldsymbol{\lambda} \prec \sum_{i=1}^{\Omega_\xi} \gamma_i \boldsymbol{\mu}(\xi_i) \right\} \quad (2.5)$$

2.2 Lyapunov optimization

In stochastic queuing networks, Lyapunov optimization is a fundamental tool for optimizing time average of certain performance objective (e.g. maximizing average throughput, minimizing average energy etc) under stability constraints. The drift-plus-penalty theorem in [31] provides a control design that maximizes the time average of some performance metric subject to queue stability. A trade-off is derived between the utility maximization and the queue size. The bounds in terms of time average expectations are provided for the drift-plus-penalty theorem. To understand this result, we consider N queues and denote by $\boldsymbol{Q}(t) = [Q_1(t), Q_2(t), ..., Q_N(t)]$ the vector of the queue backlogs at time t. For each queue $Q_n(t)$ the arrival process is $\lambda_n(t)$ with an average $\bar{\lambda}_n$ and the service process is $\mu_n(t)$ with an average $\bar{\mu}_n$.

Based on random events (e.g. channel variation, packet arrivals), a control action is taken every time slot and affects by that the arrival and service variables of the queue and provokes a penalty $p(t)$. The optimization problem is to minimize $\bar{p}$ the time average of $p(t)$ subject to the stabilization of all the network queues:

$$\min \bar{p} \text{ s.t. all queues are stable} \quad (2.6)$$

Assuming that the problem is feasible, then drift-plus-penalty theory specifies a class of algorithms that verifies:

$$\lim_{t \to \infty} \sup \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[p(\tau)] \leq p^* + O\left(\frac{1}{V}\right) \quad (2.7)$$

$$\lim_{t \to \infty} \sup \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}[|Q_n(\tau)|] \leq O(V) \, \forall n \in \{1, ..., N\} \quad (2.8)$$

where the constant parameter $V \geq 0$ is chosen in a way to achieve a trade-off between the performance in terms of utility metric and queues' size, p^* is the optimal time average of $p(t)$. These bounds mean that all queues $Q_n(t)$ are strongly stable with average backlog $O(V)$ and that the time average expected penalty is within $O\left(\frac{1}{V}\right)$ of optimality. The penalty gap $O\left(\frac{1}{V}\right)$ can be diminished by choosing a finite large constant V parameter, at the expense of increasing the linear bound of the queues size $O(V)$. In the sequel, we give some insights concerning the drift-plus-penalty theory since this optimization tool has been used in chapter 4. For more information one can refer to [31].

Let us consider a quadratic Lyapunov function $L_{\boldsymbol{Q}}(t)$ as a scalar measure of the vector $\boldsymbol{Q}(t)$:

$$L_{\boldsymbol{Q}}(t) := \frac{1}{2}\sum_{n=1}^{\mathrm{N}} Q_n^2(t)$$

$L_{\boldsymbol{Q}}(t)$ is a non negative function. The conditional Lyapunov drift $\Delta(\boldsymbol{L}(t))$ at time slot t is defined as follows:

$$\Delta(\boldsymbol{L}(t)) := \mathbb{E}\left[L_{\boldsymbol{Q}}(t+1) - L_{\boldsymbol{Q}}(t) \,|\boldsymbol{Q}(t)\right]$$

$$\text{Since, } L_{\boldsymbol{Q}}(t+1) - L_{\boldsymbol{Q}}(t) = \frac{1}{2}\sum_{i=1}^{\mathrm{N}}\left[\left(\max\left[Q_i(t) - \mu_i(t), 0\right] + \lambda_i(t)\right)^2 - Q_i(t)^2\right]$$

and $\left(\max\left[Q_i(t) - \mu_i(t), 0\right] + \lambda_i(t)\right)^2 \leq Q_i^2(t) + \lambda_i^2(t) + \mu_i^2(t) + 2Q_i(t)\left(\lambda_i(t) - \mu_i(t)\right)$

$$\Delta(\boldsymbol{L}(t)) \leq \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} \frac{\lambda_i^2(t) + \mu_i^2(t)}{2} |\boldsymbol{Q}(t)\right] + \sum_{i=1}^{\mathrm{N}} Q_i(t)\,\bar{\lambda}_i - \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} Q_i(t)\,\mu_i(t)\,|\boldsymbol{Q}(t)\right] \tag{2.9}$$

We consider that the second moment of both the arrival and service process are finite, thus there exists a constant C such that $\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} \frac{\lambda_i^2(t)+\mu_i^2(t)(t)}{2} |\boldsymbol{Q}(t)\right] \leq C$. This yields to:

$$\Delta(\boldsymbol{L}(t)) \leq C + \sum_{i=1}^{\mathrm{N}} Q_i(t)\,\bar{\lambda}_i - \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} Q_i(t)\,\mu_i(t)\,|\boldsymbol{Q}(t)\right] \tag{2.10}$$

Instead of finding the control decision that minimizes a bound of $\Delta(\boldsymbol{L}(t))$, the minimization is done at the level of the following drift-plus-penalty expression:

$$\Delta(\boldsymbol{Q}(t)) := \Delta(\boldsymbol{L}(t)) + V\mathbb{E}\left[p(t)\,|\boldsymbol{Q}(t)\right]$$

where $V \geq 0$ a non negative weight parameter that induces a trade-off between minimizing the time average of $p(t)$ and the queues' size. Thus, the upper bound of $\Delta(\boldsymbol{Q}(t))$ is:

$$\Delta(\boldsymbol{Q}(t)) \leq C + \sum_{i=1}^{\mathrm{N}} Q_i(t)\,\bar{\lambda}_i - \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} Q_i(t)\,\mu_i(t)\,|\boldsymbol{Q}(t)\right] + V\mathbb{E}\left[p(t)\,|\boldsymbol{Q}(t)\right]$$

The control actions are taken in such a way that the Lyapunov function is drifted into negative direction and by that pushing the queue backlog towards a lower congestion state while minimizing the time average of the utility function $p(t)$. Thus, the decision is based on the weighted sum expressed above.

Now, assuming that the stability of the queues is verified, thus the vector of the arrival processes' averages $\boldsymbol{\lambda}$ belongs to the stability region. Considering an $\epsilon > 0$ such that $\boldsymbol{\lambda} + \epsilon$ belongs to the stability region, there exists a control algorithm $\alpha^*(t)$ such that:

$$\mathbb{E}\left[\mu_i\left(\alpha^*(t)\right)|\boldsymbol{Q}(t)\right] \geq \bar{\lambda_i} + \epsilon \ \forall i = 1, ..., N \tag{2.11}$$

$$\sum_{i=1}^{N} \mathbb{E}\left[p\left(\alpha^*(t)\right)|\boldsymbol{Q}(t)\right] \leq B \text{ with } B \text{ a finite constant} \tag{2.12}$$

Considering the expectation over the values of the queues $\boldsymbol{Q}(t)$ yields:

$$\mathbb{E}\left[\Delta\left(\boldsymbol{Q}(t)\right)\right] \leq C - \epsilon \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[Q_i(t)\right] + VB \tag{2.13}$$

$$\Rightarrow \mathbb{E}\left[L_{\boldsymbol{Q}}(t+1)\right] - \mathbb{E}\left[L_{\boldsymbol{Q}}(t)\right] + V\mathbb{E}\left[p(t)\right] \leq C - \epsilon \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[Q_i(t)\right] + VB \tag{2.14}$$

Summing the above over $t \in \{0, 1, ..., T-1\}$ for some integer $T > 0$ yields:

$$\mathbb{E}\left[L_{\boldsymbol{Q}}(T)\right] - \mathbb{E}\left[L_{\boldsymbol{Q}}(0)\right] + V \sum_{t=0}^{T-1} \mathbb{E}\left[p(t)\right] \leq CT - \epsilon \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[Q_i(t)\right] + VTB \tag{2.15}$$

Neglecting non-negative quantities, rearranging the terms and dividing by ϵT yields:

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[Q_i(t)\right] \leq \frac{C}{\epsilon} + \frac{\mathbb{E}\left[L_{\boldsymbol{Q}}(0)\right]}{\epsilon T} + \frac{B}{\epsilon} \tag{2.16}$$

Neglecting non-negative quantities, rearranging the terms and dividing by VT yields:

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\left[p(t)\right] \leq B + \frac{C}{V} + \frac{\mathbb{E}\left[L_{\boldsymbol{Q}}(0)\right]}{VT} \tag{2.17}$$

Assuming that $\mathbb{E}\left[L_{\boldsymbol{Q}}(0)\right] < \infty$ and considering the superior limits as $T \to \infty$ proves that the drift-plus-penalty theory generates control algorithms that achieve performance given in equations (2.7) and (2.8).

2.3 POMDP

2.3.1 Definition

Partially Observed Markov Decision Process (POMDP) is a generalization of Markov Decision Process (MDP) that models a sequential agent decision process in a system that has a Markovian evolution and where the system states are not fully observable. The system dynamics are determined by a MDP where the underlying states of the system cannot be observed directly by the agent. However, the agent benefits from previous observations for maintaining a probability distribution over the set of possible states. Thus, in order to make optimal actions, the agent should take into account all the previous observations and actions. This makes exact solutions suffering from high computational cost. For a full comprehension of POMDP and their existing solutions the reader can refer to [32].

In a sequential decision process, the agent interacts with the environment and takes optimal actions that maximize certain reward function. A finite-horizon POMDP is formally described by the following tuplet $< S, A, T, R, Z, O, H, \gamma >$:

- S is the set of environment states that describe the environment at a particular point. We generally assume discrete state space.
- A is the set of actions that the agent may take.
- T is the state transition function. Since the process is Markovian, then the current state is sufficient for predicting the next one with $T(s, a, s')$ the probability of a transition from state s to state s' after the execution of action a.
- $R : S \times A \to \mathbb{R}$ is the reward function that assigns a numeric value for each state and action.
- Z is the set of observations an agent can receive.
- $O : S \times A \to Z$ is the set of conditional observation probabilities where $O(s', a, z)$ specifies the probability that observation z will be recorded after performing action a and landing in state s.
- H is the horizon over which the decisions are made (with the possibility to have infinite horizon).
- $\gamma \in [0, 1]$ is the discount factor that aims to discount future rewards and reduce by that their contribution compared to current rewards.

At each discrete time $t \in \mathbb{N}$ the system is in some state $s_t \in S$. The agent takes action $a_t \in A$ which transits the system to state $s'_{t+1} \in S$ with probability $T(s'|s, a)$. This leads the agent to observe o_{t+1} with probability $O(o|s', a)$ and to receive a reward $R(s_t, a_t)$. The aim of the agent is to maximize its expected discounted rewards collected through time $\mathbb{E}\left[\sum_{t=1}^{H} \gamma^t R(s_t, a_t)\right]$.

2.3.2 Belief State MDP

Based on the history of previous observations and actions, the agent chooses the optimal action to take. Since the agent has not a full knowledge of the system state, it maintains a certain probability distribution over the system states called belief states. The result of [33] demonstrates that selecting upcoming actions based on the belief states is sufficient for defining the optimal policy. The previous observations and actions are gathered in the belief state defined as it follows:

$$b_t(s) := P(s_t = s | z_t, a_{t-1}, z_{t-1}, ..., a_0) \, \forall s \in S \tag{2.18}$$

Therefore, a POMDP can be formulated as a fully observable MDP with a continuous belief states (i.e. there exists infinite belief states). This equivalent belief state MDP is described by the following quadruple $< B, A, \tau, \rho >$:

- B is the continuous belief simplex.
- A is the set of actions that the agent may take.
- $\tau : B \times A \to B$ is the belief transition function. The belief states pursue a Markov process where the belief state at a given epoch depends on the belief, action and observation of the previous epoch with the following transition function:

$$\tau(b, a, b') = Pr(b'|b, a) = \sum_{z \in Z} Pr(b'|b, a, z) \, Pr(z|b, a) \tag{2.19}$$

$$= \sum_{z \in Z} Pr(b'|b, a, z) \sum_{s' \in S} O(s', a, z) \sum_{s \in S} T(s, s') \, b(s)$$

  where

$$P(b'|b, a, z) = \begin{cases} 1 & \text{iff } b' = b_z^a \\ 0 & \text{otherwise} \end{cases}$$

  When action a is taken and observation z is received, the previous belief b is updated as it follows:

$$b_z^a(s') = \frac{O(s', a, z) \sum_{s \in S} T(s, a, s') \, b(s)}{Pr(z|a, b)}$$

- $\rho : B \times A \to \mathbb{R}$ is the belief-based reward

$$\rho(\boldsymbol{b}, \boldsymbol{a}) = \sum_{\boldsymbol{s} \in S} b(\boldsymbol{s}) R(\boldsymbol{s}, \boldsymbol{a}) \tag{2.20}$$

A policy $\pi : B \to A$ is a stationary mapping from the belief space to the action space. For an horizon H and initial belief b_0, each policy π is characterized by a value function $V^\pi(b_0)$ that evaluates the expected cumulative reward from $t = 1$ to $t = H$:

$$V^\pi(b_0) = \sum_{t=1}^{H} \gamma^t \rho(b_t, \pi(b_t)) = \mathbb{E}_\pi \left[\sum_{t=1}^{H} \gamma^t R(s_t, a_t) \, | b_0, \pi \right] \tag{2.21}$$

The optimal policy π^* is obtained by optimizing the value function:

$$\pi^* = \underset{\pi}{\operatorname{argmax}} V^{\pi}(b_0) \tag{2.22}$$

The action-value function $Q_t^{\pi}(b,a)$ is the reward of taking action a at time t and following policy π thereafter:

$$Q_t^{\pi}(b,a) = \rho(b,a) + \gamma \sum_{z \in Z} Pr(z|a,b) V_{t+1}^{\pi}(b_z^a) \tag{2.23}$$

The objective of the agent decision is to maximize the expected cumulative reward. There exists an optimal Markov policy which defines the action to be executed for each belief state supposing that the upcoming actions will be chosen in an optimal manner. The optimal value $V_t^*(b)$ satisfies the Bellman optimal equation:

$$V_t^*(b) = \max_a Q_t^{\pi}(b,a) = \max_a \left[\rho(b,a) + \gamma \sum_{z \in Z} Pr(z|a,b) V_{t+1}^*(b_z^a) \right] \tag{2.24}$$

2.3.3 Optimal Policy

An exact solution to a POMDP consists of determining, for each belief state $b \in B$, the optimal action that maximizes the expected reward of the agent over a given horizon (that may be infinite). The optimal policy corresponds to the sequence of optimal actions. Since any POMDP can be formulated as continuous belief-state MDP, value iteration (given in Algo. 1 for a finite horizon H) can be used for computing the optimal decision policy of POMDP problem.

Algorithm 1 POMDP Value Iteration

1: $t = 0$ and $V_0(b) = 0 \, \forall \, b \in B$
2: **for** $t = 1, .., H$ **do**
3: **for** $b \in B$ **do**
4:

$$V_t(b) = \max_{a \in A} \left[\rho(b,a) + \gamma \sum_{b' \in B} \tau(b,a,b') V_{t-1}(b') \right]$$

$$= \max_{a \in A} \left[\sum_{s \in S} b(s) R(s,a) + \gamma \sum_{z \in Z} Pr(z|a,b) V_{t-1}(b_z^a) \right]$$

5: **end for**
6: **end for**

In practice, the exact solution of finite-horizon POMDP has been proven to be PSPACE-complete [34], and their existence for infinite-horizon POMDP is undecidable [35]. Therefore, several approximated optimal solutions have been constructed especially those that use dynamic programming (e.g. [33], [36], [37], [38]).

For developing efficient algorithms, important properties of the value function are explored. The optimal finite-horizon value function is Piecewise-Linear Convex (PWLC) and can be represented as a finite set of hyperplans $\mathcal{V}_t = \{\alpha_t^i\}, i = 1, ..., |\mathcal{V}_t|$ (see [33]). Each $\alpha-$vector defines a region of the belief state B where this vector represents the highest value of V_t. The value function at a belief state b and time t is deduced from the set of α-vectors $\mathcal{V}_t$ as it follows:

$$V_t(b) = \max_{\alpha_t^i \in \mathcal{V}_t} \sum_{s \in S} b(s)\, \alpha_t^i(s)$$

Therefore, the essential task of POMDP value-iteration algorithms is to find the set of $\alpha-$vectors $\mathcal{V}_{t+1}$ (corresponding to value function V_{t+1}) given the previous set of α-vectors $\mathcal{V}_t$ (corresponding to value function V_t). Various algorithms that manipulate α-vectors in order to deduce the value function have been proposed in the literature. The most naive way to construct the set of hyperplans $\mathcal{V}_{t+1}$ is by listing all the pair of observations and actions that are possibly mapped to the set $\mathcal{V}_t$ (see [33]). In the sequel, we describe the required operations for constructing the projections' set $\mathcal{V}_{t+1}$ based on the previous set $\mathcal{V}_t$:

- *(Step 1)* Compute the vectors associated with each action $a \in A$:

$$\Gamma^{a,*} \leftarrow \alpha^{a,*}(s) = R(s,a) \tag{2.25}$$

- *(Step 2)* Compute the projection for each pair action-observation $a \in A$ and $z \in Z$:

$$\Gamma^{a,z} \leftarrow \alpha_i^{a,z}(s) = \gamma \sum_{s' \in S} T(s,s')\, O(z,s',a)\, \alpha_t^i(s')\, \forall\, \alpha_t^i \in \mathcal{V}_t \tag{2.26}$$

- *(Step 3)* Apply the cross-sum operator $\oplus$ in order to complete the generation of the set of α-vectors associated to the time $t+1$ and all actions $a \in A$:

$$\Gamma^a = \Gamma^{a,*} + \oplus_{z \in \mathcal{Z}} \Gamma^{a,z} \tag{2.27}$$

- *(Step 4)* Finally, we consider the union of all the sets Γ^a for all actions $a \in A$:

$$\mathcal{V}_{t+1} \leftarrow \mathcal{L}\mathcal{V}_t = \bigcup_{\boldsymbol{a} \in A} (\Gamma_c^{\boldsymbol{a}}, \Gamma_r^{\boldsymbol{a}}) \tag{2.28}$$

Thus, the size of $\mathcal{V}_{t+1}$ is equal to $|A||\mathcal{V}_t|^{|Z|}$ and considering the state space it gives $|S|^2|A||\mathcal{V}_t|^{|Z|}$. Many vectors in $\mathcal{V}_t$ might be dominated in the sense that they are never the optimal vector in any region of the belief simplex. These dominated vectors are called useless vectors. Therefore, one can propose some pruning algorithms that exclude these useless vectors and represent the optimal value function at time t by a subset of smaller size. As an example, algorithms [39], [33] and [40] have proposed a way to identify these useless vectors in order to ignore them. Such pruning operations are computationally expensive since a linear programming problem is needed to be solved for each α-vector.

In practice, optimal solutions for POMDP are intractable to find for realistic size problems. Thus, researchers have been motivated to develop approximated solutions for solving POMDP. Myopic solutions [41] and [38], i.e. where the action is taken independently from the belief state and the observations, define a lower bound of the POMDP optimal solution. A MDP approximation of the POMDP was proposed in [42]. A modification of MDP approximations is to use the Q-functions. Moreover, the value function on the continuous belief space can be approximated by an interpolation or extrapolation rule based on a finite number of belief states. The most commonly used interpolation approximation in the literature is linear interpolation [41]. This is proven to be an upper bound of the optimal value function [43].

2.4 Submodularity

The submodularity is an important property of set functions that leads to interesting theoretical results and arises in various practical applications. From global view, this property shows similarity to the convexity of real-valued functions. $f : 2^V \to \mathbb{R}$ is a set function that assigns each subset $S \subset V$ a value function $f(S)$ (i.e. assuming that empty set carries no value $f(\emptyset) = 0$). We start by defining the the notion of discrete derivative from which both monotonicity and submodularity properties are deduced.

Definition 2.4.1. *(Discrete derivative) For a set function $f : 2^V \to \mathbb{R}$, subset $S \subset V$ and element $e \in V$, the discrete derivative of f at S with respect to e is:*

$$\Delta_f(e|S) := f(S \cup e) - f(S) \tag{2.29}$$

Definition 2.4.2. *(Monotonicity) A set function $f : 2^V \to \mathbb{R}$ is monotone if for every $A \subset B \subset V$, $f(A) \leq f(B)$ which is equivalent to $\forall\, e \in V$ it holds: $\Delta_f(e|A) \geq 0$.*

Definition 2.4.3. *(Submodularity) A set function $f : 2^V \to \mathbb{R}$ is submodular if for every $A \subset B \subset V$ and any element $e \in V \setminus B$, it holds that:*

$$f(A \cap B) + f(A \cup B) \leq f(A) + f(B) \Rightarrow \Delta_f(e|A) \geq \Delta_f(e|B) \tag{2.30}$$

2.4.1 Submodular function Maximization

We focus on the problem of maximizing submodular functions of the form:

$$\max_{S \subset V} f(S) \text{ s.t. some constraints on } S \tag{2.31}$$

Unfortunatly, even for the simplest example of cardinality constraints ($|S| \leq K$) the problem above remains NP-hard. Thus, for the case of cardinality constraints , a simple greedy algorithm was proposed in [44]. It consists of starting with the empty

set $S_0 = \emptyset$ and in iteration $i = 1, 2, ..., K$ the added element is the one that maximizes the discrete derivative $\Delta_f(e|S_{i-1})$:

$$S_i = S_{i-1} \cup \{\underset{e \in V \setminus S_{i-1}}{\operatorname{argmax}} \Delta_f(e|S_{i-1})\} \tag{2.32}$$

The following theorem shows that this greedy algorithm is a good approximation of the optimal solution of problem 2.31 with cardinality constraints.

Theorem 2.4.4. *(Performance of greedy maximization [44]) Considering a nonnegative monotone submodular function $f : 2^V \to \mathbb{R}_+$ and let S_K be the greedily selected sets defined in equation (2.32) after K iterations, then:*

$$f(S_K) \geq \left(1 - \frac{1}{e}\right) \max_{S \subset V : |S| \leq K} f(S) \tag{2.33}$$

2.4.2 Submodular Knapsack

Moreover, maximizing a submodular function subject to a submodular upper bound constraint (submodular knapsack) is investigated in [45]. This class of discrete optimization problems have the following form with f and g submodular functions:

$$\max_{S \in V} f(S) \text{ s.t. } g(S) \leq b \tag{2.34}$$

The special case that we use in chapter 5 is the one where $g(S) = \sum_{e \in S} g(e)$ (called modular function). In this case, authors in [46] propose the greedy algorithm described in Algo. 2. This algorithm provides a $\left(1 - \frac{1}{e}\right)$ approximation .

Algorithm 2 Greedy Algorithm with knapsack constraint

Input: Constraint b, set V, set functions f and g
$S = \emptyset$, $V' = V$
while $V' \neq \emptyset$ **do**
  for $v \in V'$ **do**
    $\Delta_f(e|S_{i-1}) = f(S \cup e) - f(S \cup e)$
  end for
  $e^* = \underset{e \in V \setminus S_{i-1}}{\operatorname{argmax}} \frac{\Delta_f(e|S_{i-1})}{g(i)}$
  if $g(S \cup e^*) \leq b$ **then**
    $S = S \cup e^*$
  end if
  $V' = V' \setminus e^*$
end while

For the general problem of maximizing a submodular function subject to a submodular upper bound constraint, a greedy algorithm was discussed in the extended version of [45] and the performance of this algorithm is derived compared to the optimal solution.

3 | D2D Mode Selection

This chapter presents a fully analytical approach for the performance evaluation of overlaid D2D communication systems. We use a queuing theory approach to provide a comparison between overlaid D2D and cellular communications in Time Division Duplex (TDD) wireless networks. We provide a simple theoretical characterization of the queuing stability region for two wireless network models: the first one is a pure cellular network where any two UEs can communicate with each other only through the BS and the second one is a hybrid network where overlay D2D coexist with cellular communications. If we consider a communication between nearby devices that passes through the BS then a packet cannot be delivered to the destination UE until it is first received by the BS queue from the source UE. Hence, due to the relaying functionality at the BS level, a coupling is created between the UL queue and the DL queue and applying a queuing theory approach allows us to capture this coupling. In addition, we assume a link adaptation model (i.e. multiple rate model) where the bit rate of a link depends on its radio conditions. The coupling between the queues as well as the multiple rate model are the main challenges that highly increase the complexity of the stability region characterization. In this chapter, we provide a fully characterization of the stability region for both cellular and D2D scenarios as convex polytopes with a finite number of vertices. An approximated model is proposed for reducing the computational complexity of the exact stability region. Based on a comparison between both scenarios, we derive new insights in the cases where overlay D2D mode outperforms cellular mode. For the multi-user scenario, a trade-off is established between the complexity and the precision of the approximated stability region compared to the exact one. Furthermore, numerical results are presented to corroborate our claims and to show the gains that overlay D2D can bring to cellular networks.

3.1 Concept and Related Work

The revealing promise of D2D communications has motivated researchers to explore the potentials of D2D communications and to solve the challenges faced by this technique in order to be implemented in communications infrastructure. For this aim, several tools have been used for modeling D2D technology such as: game theory (e.g. [47], [48] and [49]), stochastic geometry (e.g. [50], [51] and [52]) and coupled processors (e.g. [53] and [54]). However, among these works, only couple processors

studies have considered a traffic pattern and did not limit their study to a physical layer approach with saturated queues. These works emphasize the importance of accurately capturing the dependencies that exist between the achievable rates of the different users in the network.

One of the main questions that operators may face is when to enable D2D links for enhancing the overall network capacity. Therefore, in this chapter, we propose a mode selection framework that allows us to choose between D2D and cellular links: in which conditions is it more advantageous (from a capacity point of view) to use D2D than cellular links. Our study is based on the stability region characterization of the following two scenarios: cellular networks with and without enabling direct D2D communications.

3.1.1 Related Work

Recently, there have been several attempts to provide an answer to the aforementioned question: which transmission mode, i.e. via D2D or cellular links, will be used by the devices in order to maximize the performance of the cellular networks? In the literature, more attention has been given to underlay D2D networks where D2D links reuse the resources allocated to cellular communications. Authors of [55] and [56] present a resource allocation and mode selection schemes that aim to maximize the system throughput. The work in [57] considers the power efficiency as performance metric and proposes a centralized power allocation and mode selection algorithm that achieves higher power efficiency based on an exhaustive search over all possible modes of all devices.

Apart from the underlay case, authors in [58] consider an overlay D2D network where dedicated resources are used for D2D communications in order to mitigate interference. A BS-assisted resource allocation and power control algorithm is proposed in order to eliminate interference between D2D users. Based on stochastic geometry and considering both underlay and overlay D2D, authors in [21] studied a distance-based mode selection as well as an optimal spectrum sharing between cellular and D2D transmissions. For more details on the existing work on D2D at the physical layer one can refer to [1] and [2] and the references therein.

Most of these works have studied the problem at the physical layer without considering any traffic pattern. However, we show in this work that the flow-level approach brings important additional outcomes in terms of network capacity evaluation.

3.1.2 Contribution and organization

In this chapter, we attempt to investigate in which cases overlay D2D is more beneficial than cellular links. Contrary to most of aforementioned existing work, the main performance metric used here is the queuing stability. The use of bursty traffic and queuing analysis is motivated by the following scenario. A communication between two nearby devices in a cell can be done directly (i.e. D2D) or through the base

station. Indeed, in TDD cellular networks, when a pair of users communicate with each other by passing through the BS, the UL and DL parts of this communication compete for the same spectrum. The BS plays the role of a relay that receives packets from the first device (source) on the UL, stacks them in a buffer and then transmits them to the second device (destination) on the DL using an opportunistic scheduling algorithm. In a TDD system, the UL and DL transmissions are performed on different time-slots, hence a packet cannot be transmitted on the downlink (BS-to-destination) until it is first received by the BS on the uplink (source-to-BS). During some time-slots, even when the channel state of the BS-destination link DL is favorable, it may happen that the buffer at the BS is empty (i.e no need to schedule this DL).

This coupling cannot be captured by a simple performance analysis at the physical layer. Hence, a traffic pattern must be included in the analysis in order to provide a more realistic evaluation of the network capacity. Therefore, we explicitly illustrate in this chapter the difference between the two approaches: performance evaluation (i) taking into account the queuing aspects (coupling queues) and (ii) without considering the coupling between the queues. The latter case is based on assuming full buffer queues where the bit rate is computed at the physical layer with saturated queues. Both the bursty traffic and the relaying role of the BS lead to a coupling between the queues in the system such that the service rate of each queue will depend on the state of the other queues. The stability region characterization of the system of interacting queues has been a challenging problem and has received the attention of researchers (i.e. especially for the case of multiple-access channel networks).

The stability region is defined as the union of all arrival rate vectors to the sources such that all the queues in the network remain bounded (precised technical definition is provided in section 3.2). In order to compare the performance of overlay D2D and cellular communications and to determine the cases where overlay D2D is favorable over cellular links, we consider a bursty traffic and provide a simple characterization of the stability region for the two following network scenarios:

Cellular scenario: Only cellular communications are possible and D2D is not allowed. Two nearby users communicate through UE-to-BS (UL) and BS-to-UE (DL) links. The BS transmits packets to the destination user only if it receives them from the source user. There is a buffer at the BS and this buffer could be empty during some time slots. The coupling between the queues of the system and that of the BS makes the problem challenging.

D2D scenario: This is an hybrid model where both D2D and cellular communication are allowed. This hybrid model is motivated by the fact that D2D will not totally replace the cellular communications and there will always be users exchanging data with the BS. UE-to-UE communication can be established directly between nearby users without passing by the BS. D2D links use different radio resources than the cellular links (i.e. overlay D2D).

In the literature, the stability region analysis of interacting queues with relaying functionality was mainly analyzed either: (i) based on stochastic dominance technique for finding the stability region (mainly for aloha systems with three-user scenario

and ON/OFF transmission model) or at least describing the necessary and sufficient conditions for queuing stability or (ii) via numerical and simulation results. However, in this work we approach the problem in a different way and we transform the multidimensional Markov Chain that models the network to multiple One Dimensional (1D) Markov Chain models (as detailed in the sequel). After evaluating the different 1D Markov Chains, we characterize the exact stability region as a convex polytope with a limited number of vertices. This exact stability region turns out to have a high complexity. Therefore, we propose approximated models characterized by having an explicit analytic form and a low complexity while guaranteeing a high precision compared to the exact stability region. We start by considering the simple scenario that consists of one UE2UE and one UE2BS communication. In theorem 3.3.1, we characterize the exact stability region of this simple scenario which turns out to have a non explicit complex form. Hence, we propose a closed-form tight upper bound of the exact stability region (see theorem 3.3.3). Moreover, for the general scenario of multiple UE2UE and UE2BS communications, we evaluate the exact stability region and we discuss its computational complexity (see theorem 3.4.1). Thus, we propose two techniques to respectively reduce the complexity of the exact stability region (see theorems 3.4.2 and 3.4.5). We deduce the trade-off between the complexity and the precision of the stability region computation. The key particularities of this work are summarized as follows:

1. The majority of previous works does not take the traffic pattern into consideration and consider the queues saturated at all the transmitters. In this work, motivated by capturing the relaying effect at the BS level, a bursty traffic is considered and the performance of the network is evaluated in terms of stability region based on a queuing theory approach.

2. Previous works focus on the underlaid D2D communications including the high complexity of interference mitigation. By contrast, the spectrum access studied in this work is the overlay.

3. We assume a link adaptation model rather than single rate model (ON/OFF model). It corresponds to the matching of the bit rate to the radio conditions (i.e. SNR) of the link. This realistic assumption makes the analysis more complicated as one can see in the sequel.

The remainder of this chapter is organized as follows. Section 3.2 describes the system model. The stability region analysis for a three-UEs simple case is provided in Section 3.3 in order to have a clear perception of the advantages of overlay D2D compared to cellular communications. The stability region of the multi-UEs general case is characterized in Section 3.4 as a simple convex polytope. Numerical results are presented in Section 3.5. Section 3.6 concludes the chapter whereas the proofs are provided in the appendices 7.3.

3.2 System Model

In this study, two scenarios are considered. In the first one, only cellular communications are allowed while in the second one a mix of cellular and D2D communications is assumed. In both scenarios, we consider a set of UE-to-BS communications where users are transmitting packets to the BS (e.g. to contact users in other cells or to have internet connections, etc.) and a set of UE-to-UE communications (between pair of UEs). In the cellular scenario, the UE2UE communications are performed by passing through the BS contrary to the D2D scenario where the users that operates UE2UE communications are connected directly to each other (without passing through the BS). Only UL type of UE2BS communications is considered since D2D is assumed to share UL resources with cellular network. In practice, these two types of communications (UE2UE and UE2BS) coexist with DL communications (i.e. between the BS and the users) However, extending this work to the scenario where the DL traffic is taken into account is a straightforward process.

We assume a single cell scenario with K UE2UE communications and U UE2BS communications. In other terms, we suppose $K+U$ communications and $2K+U$ users in the cell. We denote by $\text{UE}_{i,s}$ and $\text{UE}_{i,d}$ the pair source and destination users corresponding to the i^{th} UE2UE communication (for all $1 \leq i \leq K$) and by $\text{UE}_{j,u}$ the user corresponding to the j^{th} uplink UE2BS communication (for all $1 \leq j \leq U$). Let's describe the cellular and D2D scenarios as illustrated in Figure 3.1:

- **Cellular scenario**: Only cellular communications are possible and D2D is not allowed. Two devices that want to communicate with each other must exchange their packets through the BS. No direct communication is possible between two users. The communication between $\text{UE}_{i,s}$ and $\text{UE}_{i,d}$ is performed through the BS ($\forall 1 \leq i \leq K$) such that the BS transmits packets to the destination user $\text{UE}_{i,d}$ only if it receives them from the source user $\text{UE}_{i,s}$. Therefore, there is a buffer at the base station level that corresponds to each UE2UE communication. This buffer can be empty during some time slots. Therefore, a coupling exists between the queues such that the service rate of the users' queues depend on the state empty or not of the BS which makes the queuing stability analysis challenging. In this scenario the following links exist: $\text{link}_{i,s}$: $\text{UE}_{i,s}$ - BS; $\text{link}_{i,d}$: BS - $\text{UE}_{i,d}$ and $\text{link}_{j,u}$: $\text{UE}_{j,u}$ - BS ($\forall 1 \leq i \leq K, \forall 1 \leq j \leq U$).
- **D2D Scenario**: This is an hybrid model where both D2D and cellular communication are allowed. This hybrid model is motivated by the fact that D2D will not totally replace the cellular communications and there will always be some users that are exchanging data with the BS. We assume that UE2UE communications are established directly between the UEs without passing by the BS while the users of the UE2BS communications are connected to the BS. In this scenario the following links exist: $\text{link}^{d}_{i,sd}$: $\text{UE}_{i,s}$ - $\text{UE}_{i,d}$ and $\text{link}^{d}_{j,u}$: $\text{UE}_{j,u}$ - BS ($\forall 1 \leq i \leq K, \forall 1 \leq j \leq U$). An overlay D2D is considered so that orthogonal resources are allocated for D2D communications and interference is mitigated.

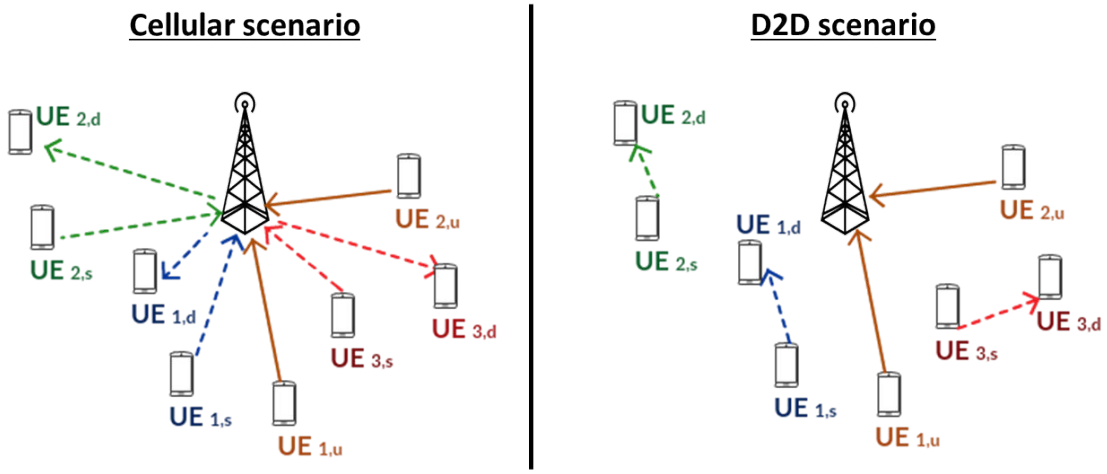


Figure 3.1: System Model

We assume a TDD system (which is a quite common assumption in D2D networks), slotted time and a user scheduling such that only one communication is possible in each time-slot. Thus, the scheduling decision at each time-slot lays on the following factors: (i) scheduling policy, (ii) channel state of the links, (iii) the state of the queues in the network. This approach supposes the knowledge of the global channel state information however the results of this work are provided without addressing the feedback overhead which requires further investigation to be integrated into our model.

3.2.1 Priority Policies

We call *priority policy* the sorting of the communications' priorities according to which the users are chosen for transmission. Therefore, among the users that are able to transmit (i.e. have some packets to transmit and have the required radio conditions to do it), the UE that is chosen to transmit is the one that has the highest priority according to the considered priority policy. Hence, a user is scheduled only when all the higher-priority users are not able to transmit. If we consider K communications then there exists $K!$ possible priority policies (i.e. $K!$ possibility of sorting these K communications). We denote by Ω_Γ the set of all the possible priority policies. $\Gamma \in \Omega_\Gamma$ denotes a priority policy according to which the users are chosen for transmission. One can see that for N communications the number of possible priority policies is given by the number of existing permutations: $N!$.

Note that any other scheduling policy is nothing but a convex combination of these priority policies. For this reason, our work is based on studying these priority policies that characterize the corner points of the stability region. Any other scheduling corresponds to an interior point of the characterized stability region. These priority policies allow us to avoid the multidimensional Markov Chain modeling of the

interacting queues. Indeed, for a given priority policy, each queue can be modeled by a one dimensional 1D Markov chain. Therefore, our approach transforms the multidimensional Markov Chain, that captures the dependency between the queues, to a 1D Markov Chain model for a given priority policy. However, the modeling that we propose remains challenging due to the coupling between the queues. As one can see later, an additional analysis is required for capturing the existing interaction between the queues.

3.2.2 System of Queues

We present the system of queues that describes both cellular and D2D scenario. In both scenarios, $\mathrm{UE}_{j,u}$ (for all $1 \leq j \leq U$) communicate with the BS through an UL cellular communication and the queue of user $\mathrm{UE}_{j,u}$ is represented by $\mathrm{Q}_{j,u}$. In cellular scenario, the communication between $\mathrm{UE}_{i,s}$ and $\mathrm{UE}_{i,d}$ (for all $1 \leq i \leq K$) is represented by the cascade of the UL queue $\mathrm{Q}_{i,s}$ of $\mathrm{UE}_{i,s}$ and the DL queue $\mathrm{Q}_{i,BS}$ (see figure 3.2). The BS does not transmit to $\mathrm{UE}_{i,d}$ unless it has received at least one packet from $\mathrm{UE}_{i,s}$. This coupling between the queues induces that the service rates of all the queues $\mathrm{Q}_{i,s}$ and $\mathrm{Q}_{i,u}$ depend on the state (empty / not empty) of each $\mathrm{Q}_{i,BS}$ which makes the queuing stability analysis challenging. The users' queues are assumed saturated. The traffic arriving to the queues $\mathrm{Q}_{i,s}$ and $\mathrm{Q}_{j,u}$ is time varying, i.i.d. over time and with rate respectively equal to $\lambda_{i,s}$ and $\lambda_{j,u}$ for $1 \leq i \leq K$ and $1 \leq j \leq U$. The traffic arriving to $\mathrm{Q}_{i,BS}$ is nothing but the departure from $\mathrm{Q}_{i,s}$.

The traffic departure from the users' queues is also time varying and depends on the queues' states (empty or not), the scheduling allocation decision and the time varying channel conditions. For a given priority policy Γ, the average service rates of the queues $\mathrm{Q}_{i,s}$, $\mathrm{Q}_{i,BS}$ and $\mathrm{Q}_{j,u}$ are respectively denoted by $\mu_{i,s}(\Gamma)$ and $\mu_{i,d}(\Gamma)$ and $\mu_{j,u}(\Gamma)$. The vector that describes the service rate of the users' queues for a given policy Γ is the following:

$$\boldsymbol{\mu}(\Gamma) = [\mu_{1,s}(\Gamma) \;\; \mu_{2,s}(\Gamma) \;\; \ldots \;\; \mu_{K,s}(\Gamma) \;\; \mu_{1,u}(\Gamma) \;\; \mu_{2,u}(\Gamma) \;\; \ldots \;\; \mu_{U,u}(\Gamma)]$$

In D2D scenario, communication between $\mathrm{UE}_{i,s}$ and $\mathrm{UE}_{i,d}$ is represented by the queue $\mathrm{Q}^d_{i,s}$ (see figure (3.2)). The users' queues are assumed saturated. The traffic arrival at queues $\mathrm{Q}^d_{i,s}$ and $\mathrm{Q}^d_{i,u}$ is time varying, i.i.d. over time and with rates respectively equal to $\lambda^d_{i,s}$ and $\lambda^d_{i,u}$. The traffic departure is also time varying and depends on the scheduling allocation decision and the time varying channel conditions. for a given priority policy Γ, the average service rates for the queues $\mathrm{Q}^d_{i,s}$ and $\mathrm{Q}^d_{j,u}$ are respectively denoted by $\mu^d_{i,s}(\Gamma)$ and $\mu^d_{j,u}(\Gamma)$. The vector that describes the service rate of the users' queues for a given policy Γ is the following:

$$\boldsymbol{\mu^d}(\Gamma) = \left[\mu^d_{1,s}(\Gamma) \;\; \mu^d_{2,s}(\Gamma) \;\; \ldots \;\; \mu^d_{K,s}(\Gamma) \;\; \mu^d_{1,u}(\Gamma) \;\; \mu^d_{2,u}(\Gamma) \;\; \ldots \;\; \mu^d_{U,u}(\Gamma)\right]$$

For both scenarios, the vector that describes the arrival rates of the users' queues is given by:

$$\boldsymbol{\lambda} = [\lambda_{1,s} \;\; \lambda_{2,s} \;\; \ldots \;\; \lambda_{K,s} \;\; \lambda_{1,u} \;\; \lambda_{2,u} \;\; \ldots \;\; \lambda_{U,u}]$$

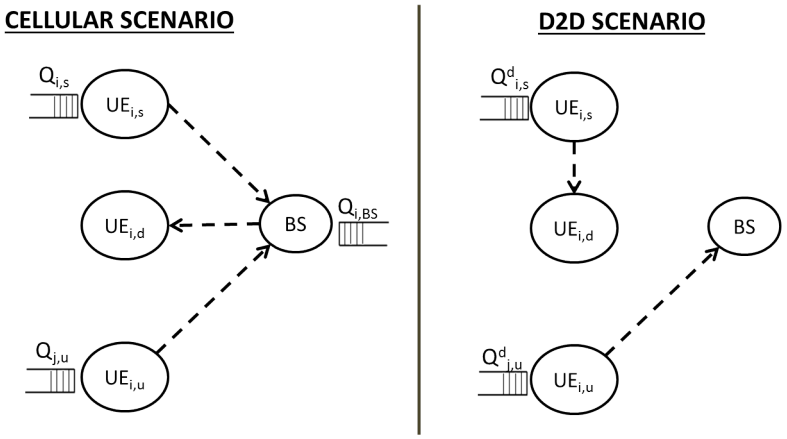


Figure 3.2: The queues model of both scenarios

3.2.3 Set of bit rates

We consider an adaptive modulation scheme such that the transmission rates is improved by exploiting the channel state at the transmitter. The SNR values are divided into a finite set of intervals $[S_1, ..., S_M]$ where the j^{th} interval $S_j = [\gamma_j, \gamma_{j-1}]$ (called hereinafter j^{th} state) is characterized by two SNR thresholds γ_{j-1} and γ_j such that $\gamma_{j-1} > \gamma_j$. Here, the adaptive modulation consists of considering a finite set $[r_1, ..., r_M]$ of bit rates as a mapping of the SNR intervals. So if a link has a SNR within the state S_j then the bit rate if its transmission is r_j. It is worth mentioning that to transmit at a rate r_j in the DL or the UL, the SNR states and thresholds may be different. In order to deal with that we use γ_j^{UL} and γ_j^{DL} to describe the SNR thresholds for respectively the uplinks and downlinks.

3.2.4 Probabilities for channel quality

The channel between any two nodes in the network is modeled by a Rayleigh fading channel that remains constant during one time slot and changes independently from one time slot to another based on a complex Gaussian distribution with zero mean and unit variance. The received SNR for a link i is given by:

$$SNR_i = \frac{|h_i|^2 P_i}{N_0 d_i^\beta}$$

where h_i is the fading coefficient, d_i is the distance between source and destination, P_i is the transmission power, β is the path loss exponent and N_0 is the noise.

We denote by p_i^n the probability that the SNR of the link i is within the n^{th} SNR interval (i.e. in state S_n),

$$p_i^n = \mathbb{P}\left[SNR_i \in S_n\right] = \mathbb{P}\left[\gamma_n \leq SNR_i \leq \gamma_{n-1}\right] = \mathbb{P}\left[SNR_i \geq \gamma_n\right] - \mathbb{P}\left[SNR_i \geq \gamma_{n-1}\right]$$

Given a complex Gaussian distribution of the channel h_i then:

$$\mathbb{P}\left[SNR_i \geq \gamma_n\right] = \mathbb{P}\left[|h_i|^2 \geq \frac{\gamma_n N_0}{d_i^{-\beta} P_i}\right] = \exp\left(-\frac{\gamma_n N_0}{d_i^{-\beta} P_i}\right)$$

$$\Rightarrow p_i^n = \exp\left(-\frac{\gamma_n N_0}{d_i^{-\beta} P_i}\right) - \exp\left(-\frac{\gamma_{n-1} N_0}{d_i^{-\beta} P_i}\right) \tag{3.1}$$

Note that $\mathbb{P}\left[SNR_i \geq +\infty\right] = 0$ and $\mathbb{P}\left[SNR_i \geq -\infty\right] = 1$. We consider the following notation $\bar{p}_i^n = 1 - p_i^n$. We develop the different probabilities that will be considered in this study:

- The probability that the uplink of the i^{th} UE2UE communication transmits at rate r_n is:

$$p_{i,s}^n = \exp\left(-\frac{\gamma_n^{UL} N_0}{d_{i,s}^{-\beta} P_{UL}}\right) - \exp\left(-\frac{\gamma_{n-1}^{UL} N_0}{d_{i,s}^{-\beta} P_{UL}}\right) \tag{3.2}$$

- The probability that the downlink of the i^{th} UE2UE communication transmits at rate r_n is:

$$p_{i,d}^n = \exp\left(-\frac{\gamma_n^{DL} N_0}{d_{i,d}^{-\beta} P_{DL}}\right) - \exp\left(-\frac{\gamma_{n-1}^{DL} N_0}{d_{i,d}^{-\beta} P_{DL}}\right) \tag{3.3}$$

- The probability that the uplink of the j^{th} UE2BS communication transmits at rate r_n is:

$$p_{j,u}^n = \exp\left(-\frac{\gamma_n^{UL} N_0}{d_{j,u}^{-\beta} P_{UL}}\right) - \exp\left(-\frac{\gamma_{n-1}^{UL} N_0}{d_{j,u}^{-\beta} P_{UL}}\right) \tag{3.4}$$

- The probability that the D2D of the i^{th} UE2UE communication transmits at rate r_n is:

$$p_{i,sd}^n = \exp\left(-\frac{\gamma_n^{UL} N_0}{d_{i,sd}^{-\beta} P_{UL}}\right) - \exp\left(-\frac{\gamma_{n-1}^{UL} N_0}{d_{i,sd}^{-\beta} P_{UL}}\right) \tag{3.5}$$

In the aforementioned expressions, $d_{i,s}$ is the distance between $\mathrm{UE}_{i,s}$ and BS, $d_{i,d}$ is the distance between $\mathrm{UE}_{i,d}$ and BS, $d_{j,u}$ is the distance between $\mathrm{UE}_{j,u}$ and BS, $d_{i,sd}$ is the distance between $\mathrm{UE}_{i,s}$ and $\mathrm{UE}_{i,d}$, P_{UL} is the user's transmission power and P_{DL} is the BS transmission power.

Definition 3.2.1. *Consider a finite point set $S \subset \mathbb{R}_+^n$ such that $S = \bigcup_{i=1}^{|S|} \{x_i\}$, we define the corresponding convex hull of $co\,(S)$ as follows:*

$$co\,(S) = \left\{ x \in \mathbb{R}_+^n \mid \forall x_i \in S :\ x \prec x_i \right\} \bigcup \left\{ \sum_{i=1}^{|S|} \gamma_i x_i \mid (\forall i : \gamma_i \geq 0) \wedge \sum_{i=1}^{|S|} \gamma_i \leq 1 \right\}$$

3.2.5 Stability region computation

We denote by Ω_ξ the set of feasible scheduling policies. Based on the stability region definition given in 2.1, we deduce as follows the stability region of the system of queues studied in this work:

$$\mathcal{R} = \bigcup_{\xi \in \Omega_\xi} \mathcal{R}(\xi) = \bigcup_{\xi \in \Omega_\xi} \left\{ \boldsymbol{\lambda} \in \mathbb{R}_+^{K+U} | \boldsymbol{\lambda} \prec \boldsymbol{\mu}(\xi) \right\} \tag{3.6}$$

We based our approach on the fact that any possible scheduling policy $\xi \in \Omega_\xi$ can be written as a convex combination of the priority policies Ω_Γ. From this comes the idea of defining the stability region based on the priority policies. Indeed, the stability region $\mathcal{R}$ can be written as the union of the stability regions achieved by considering only the feasible priority policies (denoted by Ω_Γ). Therefore, we limit the stability region computation to the following:

$$\mathcal{R} := \bigcup_{\Gamma \in \Omega_\Gamma} \mathcal{R}(\Gamma) \tag{3.7}$$

For characterizing the stability region we have to increase as much as possible the service rates of the users' queues. For the i^{th} UE2UE communications, this means increasing the service rate of the UL side ($UE_{i,s}$-BS) which implies the increase of the arrival rate at the DL side (BS-$UE_{i,d}$) and risks the loss of the stability of the BS queue $Q_{i,BS}$. Since the service rates of both UL and DL sides are coupled through the scheduling, instability of $Q_{i,BS}$ means that many packets will not be delivered to $UE_{i,d}$ and the network becomes unstable (i.e. inducing infinite average delay).

The coupling between the stability region and the queues $Q_{i,BS}$ is very challenging in general in queuing theory. Here, the main issue is to characterize the stability region by finding the priority policies that maximize the service rates of the UL side (i.e. achieve the border points of the stability region) while guaranteeing the stability of $Q_{i,BS}$. In fact, finding the policies that achieve the corner points of the stability region, i.e denoted by Ω_Γ^*, turns out to be sufficient for characterizing the stability region of the system. Therefore, for each priority policy $\Gamma \in \Omega_\Gamma^*$, we find the region that captures the coupling between the queues and guarantee their stability. Based on the definition 3.2.1 of the operator *co*, the stability region can be simply defined as follows:

$$\mathcal{R} = co\left(\bigcup_{\Gamma \in \Omega_{\Gamma^*}} \mathcal{R}(\Gamma) \right)$$

3.2.6 Organization

This work is divided into two sections. In the first section, we consider a simple scenario that consists of three users and contains one UE2UE communication and one UE2BS communication. For this three-UEs scenario, we consider a set of three bit rates $[r_1, r_2, r_3]$ such that the transmission rate of a link takes a value within this

set depending on its channel state. We start by providing the exact stability region of both D2D and cellular 3-UE scenarios. The stability region of the cellular scenario does not have an explicit form and turns be computationally complex. In order to reduce complexity, we find the explicit form of an upper bound of this stability region and prove analytically that it is a very close approximation of the stability region. The study of this simple scenario provides a comparison between D2D and cellular 3-UE scenarios. Therefore, a distance based mode selection is deduced by finding the distance threshold for which D2D communications is more advantageous than cellular ones.

In the second section, we consider the general scenario with K UE2UE communications and U UE2BS communications. For this general scenario, a set of two bit rates $[r_1, r_2]$ is considered due to the computation complexity of considering a set of three bit rates for this general case. Similarly to the simple case, we provide the exact stability region of both D2D and cellular multi-UE scenarios and we show the complexity of such computation. Hence, we propose a tight and explicit approximation that decreases the complexity of the stability region. The main novelty of this work is to capture the trade-off between the complexity and the precision of the stability region calculation.

3.3 Three-Users scenario

3.3.1 3-UEs scenario description

The Simple Scenario (SS) consists of considering a single cell wireless network with three users in the cell $\mathrm{UE}_{1,s}$, $\mathrm{UE}_{1,d}$, $\mathrm{UE}_{1,u}$ respectively denoted by UE_s, UE_d, UE_u . Both cellular and D2D scenarios are considered. We assume that user UE_s wants to communicate with UE_d, while UE_u is transmitting packets only to the BS. In the cellular scenario, the communication between UE_s and UE_d is performed through the BS. Hence, we consider the following three links: link_s: UE_s - BS, link_d: BS - UE_d, link_u: UE_u-BS. In the D2D scenario, UE_s communicates directly with UE_d without passing through the BS. Hence, we consider only two links: link_s: UE_s - UE_d and link_u: UE_u - BS. The 3-UEs D2D and cellular scenarios are illustrated in Figure 3.3.

In both scenarios, UE_u communicates with the BS through an uplink cellular communication and the queue of user UE_u is represented by Q_u. The traffic arriving to the queues of UE_s and UE_u is time varying, i.i.d. over time and with rate respectively equal to λ_s and λ_u.

In the cellular scenario, the communication between UE_s and UE_d is represented by the cascade of the UL queue Q_s of UE_s and the DL queue Q_{BS}. The BS does not transmit to UE_d unless it has received at least one packet from UE_s. This coupling between the queues induces that the service rates of Q_s and Q_u depend on the state (empty / not empty) of Q_{BS} which makes the queuing stability analysis challenging. The traffic arriving to Q_{BS} is nothing but the departure from Q_s. The traffic departure

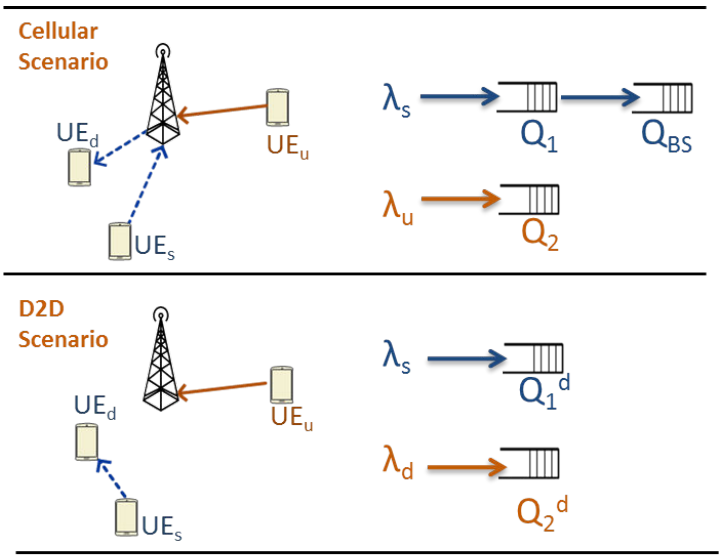


Figure 3.3: System Model of the Three UEs scenario

is also time varying and depends on the scheduling allocation decision and the time varying channel conditions. The departure average rates from queues Q_s and Q_u are respectively denoted by μ_s and μ_d.

In D2D scenario, direct communication between UE_s and UE_d is represented by the queue Q_s^d. The traffic departure is also time varying and depends on the scheduling allocation decision and the time varying channel conditions. The average service rates for both queues are denoted respectively by μ_s^d and μ_u^d.

For this simple scenario, a set of three bit rate is considered $[r_1, r_2, r_3]$ with $r_1 = kr_2$ and $r_3 = 0$ (with $k \in \mathbb{N}^+$). For this set of bit rate corresponds a set of SNR intervals $[S_1, S_2, S_3]$. As we mentioned before, we use the notation p_i^n to describe the probability that the SNR of the link i is within the n^{th} SNR interval S_n. We suppose a complex Gaussian distribution of the channel h_i; hence these probabilities can be simply derived from equation (3.1) as follows:

- For the DL BS-UE_d, p_d^n probabilities for $n = 1, 2$ and 3 are deduced from equation (3.3) as follows:

$$p_d^1 = \exp\left(-\frac{\gamma_1^{DL} N_0}{d_d^{-\beta} P_{DL}}\right)$$

$$p_d^2 = \exp\left(-\frac{\gamma_2^{DL} N_0}{d_d^{-\beta} P_{DL}}\right) - \exp\left(-\frac{\gamma_1^{DL} N_0}{d_d^{-\beta} P_{DL}}\right)$$

$$p_d^3 = 1 - p_d^1 - p_d^2$$

- For both UL UE_s-BS and UE_u-BS, the probabilities p_i^n for $n = 1, 2, 3$ are deduced from equation (3.4) as follows:

$$p_i^1 = \exp\left(-\frac{\gamma_1^{UL} N_0}{d_i^{-\beta} P_{UL}}\right)$$

$$p_i^2 = \exp\left(-\frac{\gamma_2^{UL} N_0}{d_i^{-\beta} P_{UL}}\right) - \exp\left(-\frac{\gamma_1^{UL} N_0}{d_i^{-\beta} P_{UL}}\right)$$

$$p_i^3 = 1 - p_i^1 - p_i^2$$

  with $i = s$ for the UE_s-BS link and $i = u$ for the UE_u-BS link.

- For the D2D link UE_s-UE_d, the probabilities p_{sd}^n for $n = 1, 2, 3$ are deduced from equation (3.5) as follows:

$$p_{sd}^1 = \exp\left(-\frac{\gamma_1^{UL} N_0}{d_{sd}^{-\beta} P_{UL}}\right)$$

$$p_{sd}^2 = \exp\left(-\frac{\gamma_2^{UL} N_0}{d_{sd}^{-\beta} P_{UL}}\right) - \exp\left(-\frac{\gamma_1^{UL} N_0}{d_{sd}^{-\beta} P_{UL}}\right)$$

$$p_{sd}^3 = 1 - p_{sd}^1 - p_{sd}^2$$

In the aforementioned expressions, d_s is the distance between UE_s and the BS, d_u is the distance between UE_u and the BS, d_d is the distance between UE_d and the BS, d_{sd} is the distance between D2D peer UE_s and UE_d. In the sequel, we refer to the three-UEs scenarios with 3 bit rates (r_1, r_2 and r_3) by the *simple scenario* SS .

3.3.2 Organization

For the 3-UEs scenario study, the organization is as follows: section 3.3.3 provides a theoretical analysis of the stability region for the cellular scenario. We start by characterizing the exact Stability Region (SR) for this scenario. We note that deriving this SR is computationally complex. Therefore, we propose an explicit and simple expression of an Upper Bound (UB) for this SR. After analytically evaluating the error induced by this approximation, the latter turns out to be a simple and very close approximation of the SR. In section 3.3.4, we provide the stability region for the 3-UEs D2D scenario. Both scenarios are compared in section 3.3.5 which yields to a mode selection scheme which is function of the UE-BS distance as well as the distance between the D2D peer.

3.3.3 Stability region for 3-UEs cellular scenario

In this section, we characterize the exact stability region for the 3-UEs cellular scenario. Assuming a user scheduling such that only one communication, i.e. UE_u-BS or UE_s-UE_d, is possible in each time-slot. We recall that Γ is the priority policy according to which the links are arranged. We know that in order to characterize the stability region it is sufficient to consider the corner points that correspond to the "extreme policies" where the priority is always given to the same communication or when the priority is always given to the communication that has the better channel state. In this chapter, we denote by Ω_{Γ}^{ss} the set of the priority policies corresponding to the corner points of the SS stability region and that are sufficient for the characterization of the stability region. These corner points correspond to the extreme policies where the priority is always given to the same communication or when the priority is always given to the communication that has the better channel state. Actually we consider two possible rates for each communication (r_1 and r_2) hence the priority policies Ω_{Γ}^{ss} that present the corner points of the stability region are the following:

- **Policy** Γ_1: *UE2UE communication* is given higher priority, hence *UE2BS communication* can only take place when the rate of the UL and DL of the *UE2UE communication* are null.
- **Policy** Γ_2: *UE2BS communication* is given higher priority, hence the users of *UE2UE communication* transmit only when the rate of the *UE2BS communication* is null.
- **Policy** Γ_3: At rate r_1 , the *UE2UE communication* has the priority and at rate r_2, *UE2BS communication* has the priority. Hence, *UE2UE communication* transmits at r_2 only when the rate of the *UE2BS communication* is null.
- **Policy** Γ_4: At rate r_1, the *UE2BS communication* has the priority and at rate r_2, *UE2UE communication* has the priority. Hence, *UE2BS communication* transmits at r_2 only when the rate of the uplink and downlink of the *UE2UE communication* are null.
- **Policy** Γ_5: Communication with highest rate transmits and when both communications have the same rate then *UE2UE communication* is prioritized.
- **Policy** Γ_6: Communication with highest rate transmits and when both communications have the same rate then *UE2BS communication* is prioritized.

Thus, the set of priority policies Ω_{Γ}^{ss} that corresponds to the corner points of the three-UEs stability region is given by: $\Omega_{\Gamma}^{ss} = \{\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5, \Gamma_6\}$.

However, the problem is not simply solved by finding the subset Ω_{Γ}^{ss}. In fact, for each priority policy $\Gamma \in \Omega_{\Gamma}^{ss}$, the challenge remains in capturing the coupling between the queues due to the relaying functionality of the BS between the UL and DL traffic of the UE2UE communications. For a given priority policy Γ, the service rate of the queues depend not only on the SNR states of the links but also on the state empty

or not of the BS. For this reason, we consider a queuing theory approach in order to capture this coupling between the service rate of the queues (Q_s and Q_u) and the state of Q_{BS}. In section 3.5, we expose examples that illustrate the impact of queuing approach on the performance analysis of the network.

Q_{BS} might be empty at some time-slot, therefore when UE_s-UE_d communication is scheduled, the choice between UL (UE_s-BS) or DL (BS-UE_d) depends not only on the SNR states of these two links but also on the state empty or not of Q_{BS}. In order to take that into account, we introduce a new parameter α called fraction vector. It is used to divide the time where the UE_s-to-UE_d communication is scheduled between the UL and DL sides in such a way that the stability of the queue Q_{BS} remains satisfied.

The UE_s to UE_d communication, modeled as a chain of queues Q_s - Q_{BS}, might be in 4 different SNR states: $(S_1, S_1), (S_1, S_2), (S_2, S_1), (S_2, S_2)$ where the SNR state S_i is defined in section 3.2. For each couple of SNR states for the links (UE_s-BS, BS-UE_d), we define a parameter α_i (with $0 \leq \alpha_i \leq 1$) that corresponds to the fraction of time that the resources are allocated to the UL (UE_s-to-BS) whereas $1 - \alpha_i$ corresponds to the fraction of time that the resources are allocated to the DL (BS-to-UE_d). Thus, $\alpha = (\alpha_1, \alpha_2, \alpha_3, \alpha_4)$. The fraction vector α is considered only for these 4 couples of SNR states where both UL and DL of the UE_s-to-UE_d communication have a SNR state different than S_3 (i.e. a non zero rate). Only for these cases, the concept of fraction parameter makes a sense since resources will be divided between two links able to transmit (UL and DL). However, for the other combinations of SNR states' couple, where at least one SNR state is equal to S_3, this concept does not hold because only the link with a positive rate is able to transmit.

For different priority policies of UE_u-BS and UE_s-UE_d communications, different fraction of resources will be allocated to the UL and DL parts of the chain Q_s - Q_{BS} which corresponds to different values of $\alpha \in [0, 1]^4$. For each priority policy, we find the optimal fraction vector $\alpha^* = (\alpha_1^*, \alpha_2^*, \alpha_3^*, \alpha_4^*)$ that achieves a corner point of the stability region. Finding α^* allows us to avoid the need to vary $\alpha \in [0, 1]^4$ for each priority policy in order to obtain the corresponding corner point. To make expression simpler, we use the following notation for a given priority policy Γ:

- $U(\Gamma)$ and $V(\Gamma)$ as the probabilities that UE_s transmits respectively at rate r_1 and r_2 when Q_{BS} is empty.
- $W(\Gamma)$ and $X(\Gamma)$ as the probabilities that UE_u transmits respectively at rate r_1 and r_2 when Q_{BS} is empty.
- $Y(\Gamma)$ and $Z(\Gamma)$ the probabilities that UE_u transmits respectively at rate r_1 and r_2 when Q_{BS} is not empty.
- $N(\Gamma) = p_s^1 U(\Gamma) + \bar{p}_s^1 V(\Gamma)$ and $M(\Gamma) = p_s^1 U(\Gamma) + p_s^2 U(\Gamma) + p_s^3 U(\Gamma) = U(\Gamma)$.

3.3.3.1 Exact Stability region

Theorem 3.3.1. *The stability region for the 3-UEs cellular scenario is given by* $(\lambda_s, \lambda_u) \in \mathcal{R}_c^{ss}$ *such that:*

$$\mathcal{R}_c^{ss} = co\left(\bigcup_{\Gamma\in\Omega_\Gamma^{ss}}\bigcup_{\alpha\in[0,1]^4}\left\{\mu_s\left(\alpha,\Gamma\right),\mu_u\left(\alpha,\Gamma\right)\right\}\right)$$

where Ω_Γ^{ss} *is the set of the priority policies for the SS scenario that achieves the corner points of the stability region (with* $|\Omega_\Gamma^{ss}| = 6$*). The queues' service rates* $\mu_s\left(\alpha,\Gamma\right)$ *and* $\mu_u\left(\alpha,\Gamma\right)$ *are respectively given by (3.8) and (3.9) for all the priority policies* $\Gamma \in \Omega_\Gamma^{ss}$*.*

$$\mu_s := \mu_s\left(\alpha,\Gamma\right) = \left[r_1 p_s^1 U\left(\Gamma\right) + r_2 p_s^2 V\left(\Gamma\right)\right] \times \Pi_{BS}^0\left(\Gamma\right)$$

$$+ \left[r_1 p_s^1 U\left(\Gamma\right)\left(\alpha_1 p_d^1 + \alpha_2 p_d^2 + p_d^3\right) + r_2 p_s^2\left(\alpha_3 p_d^1 U\left(\Gamma\right) + \alpha_4 p_d^2 V\left(\Gamma\right) + p_d^3 V\left(\Gamma\right)\right)\right] \times \left[1 - \Pi_{BS}^0\left(\Gamma\right)\right] \tag{3.8}$$

$$\mu_u := \mu_u\left(\alpha,\Gamma\right) = \left[r_1 p_u^1 W\left(\Gamma\right) + r_2 p_u^2 X\left(\Gamma\right)\right] \times \Pi_{BS}^0\left(\Gamma\right) + \left[r_1 p_u^1 Y\left(\Gamma\right) + r_2 p_u^2 Z\left(\Gamma\right)\right] \times \left[1 - \Pi_{BS}^0\left(\Gamma\right)\right] \tag{3.9}$$

With $\Pi_{BS}^0\left(\Gamma\right)$ *the probability that the queue* Q_{BS} *is empty and the six values of the parameters* U, V, W, X, Y, Z *corresponding to the priority policies* $\Gamma \in \Omega_\Gamma^{ss}$ *are given in table 3.1.*

Proof. See Appendix-7.3.1. □

We can resume the procedure of finding the stability region by the algorithm 3. Beside the coupling between the queues, the complexity of the following computation as well as the non existence of an explicit form of the exact stability region comes essentially from the consideration of a set of three bit rates which complicates the Markov Chain model of the queue Q_{BS}. In fact, finding the probability that queue Q_{BS} is empty Π_0 is deduced from the solution of a system of linear equations. For this reason, it is difficult to derive an explicit form of the queues' service rates. Therefore, we are not able to find the optimal fraction vector α^* for each priority policy that achieves the corner point corresponding to this policy and avoids the consideration of all the values $\alpha \in [0,1]^4$.

Algorithm 3 Procedure of the exact stability region

```
for all Γ ∈ Ω_Γ^ss do
    for all α ∈ [0,1]^4 do
        Find the probability Π_0 by solving the system of linear equations
        Deduce the service rates of the queues
    end for
end for
```

The computational complexity of the exact SR comes from two factors: (i) the need of varying the fraction vector α within $[0,1]^4$ in order to find the stability region and (ii) the calculation of the probability $\mathit{\Pi}^0_{BS}$ that the queue Q_{BS} is empty for each value of α. In fact, considering a set of three bit rates makes the Markov Chain modeling of the BS queue Q_{BS} complicated as shown in the proof 7.3.1. It yields to a complex computation of the probability $\mathit{\Pi}^0_{BS}$ by solving a linear system of equation for each value of the fraction vector α. On the aim of reducing the complexity and having an explicit form of the SR, we propose the following simple approximation of the exact SR.

3.3.3.2 Approximated stability region

As we mentioned before, finding the exact stability region is computationally complex because of the computation of the probability Π^0_{BS} and the exhaustive variation of $\alpha \in [0,1]^4$. In fact, the exact SR has not an explicit analytic form due to the fact that but is deduced from the numerical solution of a system of equations for each value of $\alpha \in [0,1]^4$. Therefore, it is important to propose an epsilon-close upper bound (ϵ-UB) that has an explicit form and that is simple to compute. An epsilon-close upper bound presents an upper bound of the exact stability region that has a maximum error equal to ϵ.

For this aim, we start by simplifying the Markov Chain modeling of the queue Q_{BS} in order to derive an approximation of the stability region that has a closed-form analytic expression and is simple to compute. Considering three different bit rates generates a complicated Markov Chain model of the queue Q_{BS} (see upper side of Fig. 3.4) makes it challenging to finding the probability Π^0_{BS} (i.e. more details given in the proof of theorem 3.3.1). Therefore, we propose a simple approximated Markov Chain for modeling the queue Q_{BS} (Lower side of Fig. 3.4). This approximated model is a simple birth and death Markov Chain where passing from state x_i to x_{i+1} (receiving a packet) corresponds to the average probability of receiving a packet at both rates r_1 and r_2. Moreover, passing from state x_i to x_{i-1} (transmitting a packet) corresponds to the average probability of transmitting a packet at both rates r_1 and r_2. We recall that the multiple rate model is still taken into consideration at the transition probabilities level of the approximated Markov Chain. Compared to the exact SR, this approximated model permits to avoid the two following main reasons that make the exact SR complex: (i) computation of Π^0_{BS} by applying an explicit formula and not by solving a system of equations and (ii) avoiding the variation of α within $[0,1]^4$ by giving the explicit expression of the optimal fraction vectors that achieve the corner points of the approximated SR.

In the sequel we characterize the approximated stability region by proceeding in two steps: (i) we derive in lemma (3.3.2) the condition that α should satisfy for each priority policy to guarantee the stability of Q_{BS} and then (ii) we provide in theorem (3.3.3) the expression of the corner points of the approximated SR and that correspond to the priority policies $\Gamma \in \Omega^{ss}_{\Gamma}$. We distinguish the results of the approximation from

that of the exact stability region by adding the following indication ˜ over all the used notation. We recall that the set of bit rates is given by $[r_1 = kr_2, r_2, r_3 = 0]$.

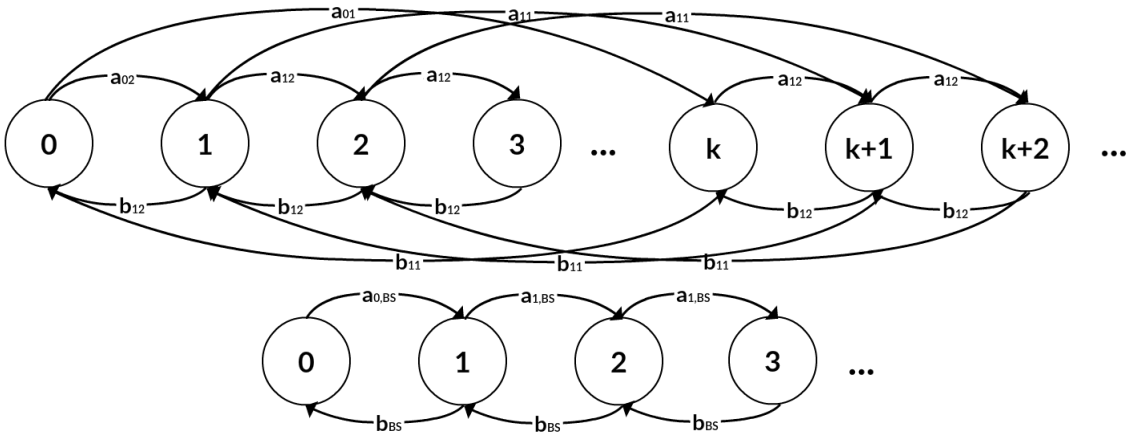


Figure 3.4: The upper Markov Chain is the exact modeling of the queue Q_{BS} and the lower Markov Chain is the approximated modeling of the queue Q_{BS}

Lemma 3.3.2. *For a given priority policy* Γ*, the fraction vector* $(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4)$ *should satisfy:*

$$2k\tilde{\alpha}_1 p_s^1 p_d^1 U + (k+1)\,\tilde{\alpha}_2 p_s^1 p_d^2 U + (k+1)\,\tilde{\alpha}_3 p_s^2 p_d^1 U + 2\tilde{\alpha}_4 p_s^2 p_d^2 V \leq k p_d^1 + p_d^2 N - \left(k p_s^1 U + p_s^2 V\right) p_d^3 \tag{3.10}$$

with U*,* V*,* W*,* X*,* Y*,* Z *depend on the priority policy* Γ *and their values for the border priority policies* $\Gamma \in \Omega_\Gamma^{ss}$ *are given in table 3.1. We recall that* $N = p_s^1 U + \bar{p}_s^1 V$*.*

Proof. See Appendix-7.3.2. □

Theorem 3.3.3. *The approximated stability region for the 3-UEs cellular scenario is the set of* $(\lambda_s, \lambda_u) \in \tilde{\mathcal{R}}_c^{ss}$ *such that:*

$$\tilde{\mathcal{R}}_c^{ss} = co\left(\bigcup_{\Gamma \in \Omega_\Gamma^{ss}} \bigcup_{\tilde{\alpha} \in \mathbb{S}_\alpha} \left\{\tilde{\mu}_s\left(\tilde{\alpha}, \Gamma\right), \tilde{\mu}_u\left(\tilde{\alpha}, \Gamma\right)\right\}\right)$$

where the queues' service rates $\tilde{\mu}_s(\Gamma)$ *and* $\tilde{\mu}_u(\Gamma)$ *are respectively given by (3.11) and (3.12),* $\mathbb{S}_\alpha$ *is a limited subset that is given in (3.13).* $\Gamma \in \Omega_\Gamma^{ss}$ *where* Ω_Γ^{ss} *is the set of priority policies that achieve the corner points of the SR (with* $|\Omega_\Gamma^{ss}| = 6$*). The values of the parameters* $U(\Gamma)$*,* $V(\Gamma)$*,* $W(\Gamma)$*,* $X(\Gamma)$*,* $Y(\Gamma)$ *and* $Z(\Gamma)$ *for the six priority policies* $\Gamma \in \Omega_\Gamma^{ss}$ *are given in table 3.1. We recall that* $N = p_s^1 U + \bar{p}_s^1 V$*.*

$$\tilde{\mu}_s := \tilde{\mu}_s(\Gamma) = \frac{1}{2}\left(r_1 p_s^1 U + r_2 p_s^2 V\right) \times$$
$$\left(1 + \frac{(1-k)\tilde{\alpha}_2^* p_s^1 p_d^2 U + (k-1)\tilde{\alpha}_3^* p_s^2 p_d^1 U - \left(k p_d^1 + p_d^2 N\right) + \left(k p_s^1 U + p_s^2 V\right)\bar{p}_d^3}{2k\tilde{\alpha}_1^* p_s^1 p_d^1 W + (k+1)\tilde{\alpha}_2^* p_s^1 p_d^2 U + (k+1)\tilde{\alpha}_3^* p_s^2 p_d^1 U + 2\tilde{\alpha}_4^* p_s^2 p_d^2 U - \left(k p_d^1 + p_d^2 N\right) - (k p_s^1 U + p_s^2 V)\bar{p}_d^3}\right) \tag{3.11}$$

$$\tilde{\mu}_d := \tilde{\mu}_d(\Gamma) = r_1 p_u^1 W + r_2 p_u^2 X + \frac{\left(r_1 p_u^1 (W-Y) + r_2 p_u^2 (X-Z)\right)\left(k p_s^1 U + p_s^2 V\right)}{2k\tilde{\alpha}_1^* p_s^1 p_d^1 U + (k+1)\tilde{\alpha}_2^* p_s^1 p_d^2 U + (k+1)\tilde{\alpha}_3^* p_s^2 p_d^1 U + 2\tilde{\alpha}_4^* p_s^2 p_d^2 V - \left(k p_d^1 + p_d^2 N\right) - (k p_s^1 U + p_s^2 V)\bar{p}_d^3} \quad (3.12)$$

$$\mathbb{S}_\alpha = \left(\begin{array}{lll} \tilde{\alpha}_2, \tilde{\alpha}_3 \epsilon \{0,1\}^2 & s.t. & (k+1)\left(\tilde{\alpha}_2 p_s^1 p_d^2 + \tilde{\alpha}_3 p_s^2 p_d^1\right) U \leq M \ , \tilde{\alpha}_1 = \tilde{\alpha}_4 = 0 \\ \tilde{\alpha}_2, \tilde{\alpha}_3 \epsilon \{0,1\}^2 & s.t. & \tilde{\alpha}_1, \tilde{\alpha}_4 \epsilon [0,1]^2 \ \ verify\ 2k\tilde{\alpha}_1 p_s^1 p_d^1 U + 2\tilde{\alpha}_4 p_s^2 p_d^2 V = M - (k+1)\left(\tilde{\alpha}_2 p_s^1 p_d^2 U + \tilde{\alpha}_3 p_s^2 p_d^1 U\right) \\ \tilde{\alpha}_2 \epsilon \{0,1\} & s.t. & \tilde{\alpha}_3 = \left[M - (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U\right] / \left[(k+1) p_s^2 p_d^1 U\right], \tilde{\alpha}_1 = \tilde{\alpha}_4 = 0 \\ \tilde{\alpha}_2 \epsilon \{0,1\} & s.t. & \tilde{\alpha}_3 = \left[M - 2k p_s^1 p_d^1 U - 2 p_s^2 p_d^2 V - (k+1)\tilde{\alpha}_2 p_s 1 p_d^2 U\right] / \left[(k+1) p_s^2 p_d^1 U\right], \tilde{\alpha}_1 = \tilde{\alpha}_4 = 1 \\ \tilde{\alpha}_3 \epsilon \{0,1\} & s.t. & \tilde{\alpha}_2 = \left[M - (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U\right] / \left[(k+1) p_s^1 p_d^2 U\right], \tilde{\alpha}_1 = \tilde{\alpha}_4 = 0 \\ \tilde{\alpha}_3 \epsilon \{0,1\} & s.t. & \tilde{\alpha}_2 = \left[M - 2k p_s^1 p_d^1 U - 2 p_s^2 p_d^2 V - (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U\right] / \left[(k+1) p_s^1 p_d^2 U\right], \tilde{\alpha}_1 = \tilde{\alpha}_4 = 1 \end{array} \right) \quad (3.13)$$

Proof. See Appendix-7.3.3. □

3.3.3.3 Comparison real and approximation

The exact stability region has not an explicit form because of the computational complexity of finding Π_{BS}^0 of the queue Q_{BS}. Moreover, this complexity is accentuated by the need of varying α within all the interval $[0,1]^4$ in order to elaborate the exact stability region. However, the approximated model gives the explicit expressions of the following two parameters: (i) probability that the BS queue is empty and (ii) the optimal fraction vectors that achieve the corner points of the approximated SR. Hence, the importance of this approximated model lies in the proposition of a simple analytic form of the stability region. Furthermore, we verify that the approximated stability region is an epsilon-close upper bound of the exact stability region. To do so, we demonstrate analytically that the relative error between these two regions is positive and bounded by a small ϵ.

Corner Point Γ	U	V	W	X	Y	Z	C	D
1	1	1	p_s^3	p_s^3	$p_s^3 p_d^3$	$p_s^3 p_d^3$	p_{sd}^3	p_{sd}^3
2	p_u^3	p_u^3	1	1	1	1	1	1
3	1	p_u^3	$\bar{p}_s^1$	$\bar{p}_s^1$	$\bar{p}_s^1 \bar{p}_d^1$	$\bar{p}_s^1 \bar{p}_d^1$	$\bar{p}_{sd}^1$	$\bar{p}_{sd}^1$
4	$\bar{p}_u^1$	$\bar{p}_u^1$	1	p_s^3	1	$p_s^3 p_d^3$	1	p_{sd}^3
5	1	$\bar{p}_u^1$	$\bar{p}_s^1$	p_s^3	$\bar{p}_d^1 \bar{p}_d^1$	$p_s^3 p_d^3$	$\bar{p}_{sd}^1$	p_{sd}^3
6	$\bar{p}_u^1$	p_u^3	1	$\bar{p}_s^1$	1	$\bar{p}_s^1 \bar{p}_d^1$	1	$\bar{p}_{sd}^1$

Table 3.1: Service rate parameters

Theorem 3.3.4. *For the 3-UEs cellular scenario, the approximated stability region $\tilde{\mathcal{R}}_c^{ss}$ is a close upper bound of the exact stability region $\mathcal{R}_c^{ss}$ with a maximum relative error ϵ_R^*. Therefore, $\mathcal{R}_c^{ss}$ is bounded as follows:*

$$(1-\epsilon_R^*)\,\tilde{\mathcal{R}}_c^{ss} \subseteq \mathcal{R}_c^{ss} \subseteq \tilde{\mathcal{R}}_c^{ss}$$

$$\text{with } \epsilon_R^* = \max_{\Gamma\in\Omega_\Gamma^{ss}} \frac{kb_{11}\left(\tilde{\alpha}^*\right)\sum_{i=1}^{k-1}\Pi_i\left(\tilde{\alpha}^*\right)\left(a_{02}+ka_{01}-a_{12}\left(\tilde{\alpha}^*\right)-ka_{11}\left(\tilde{\alpha}^*\right)\right)}{\left(a_{02}+ka_{01}\right)\left(b_{12}\left(\tilde{\alpha}^*\right)+kb_{11}\left(\tilde{\alpha}^*\right)\right)}$$

and the following parameters depend on the priority policy Γ:

- *$\tilde{\alpha}^*$ the optimum fraction vector computed analytically in $\mathbb{S}_\alpha$ (given by (3.13)) and that achieves a corner point of the approximated stability region.*
- *a_{01} and a_{02} are the arrival probabilities at Q_{BS} when this queue is empty with a respective rates r_1 and r_2 (see equations (7.3) and (7.4)).*
- *$a_{11}\left(\tilde{\alpha}^*\right)$ and $a_{12}\left(\tilde{\alpha}^*\right)$ are the arrival probabilities at Q_{BS} when this queue is not empty with a respective rates r_1 and r_2 (see equations (7.5) and (7.6)).*
- *$b_{11}\left(\tilde{\alpha}^*\right)$ and $b_{12}\left(\tilde{\alpha}^*\right)$ are the departure probabilities from Q_{BS} at rate r_1 and r_2 respectively. They depend on the fraction vector α (see equations (7.1) and (7.2)).*

Proof. See Appendix-7.3.6. □

Hence, for each priority policy Γ, we find $\epsilon\left(\Gamma\right)$ which correspond to the deviation between the corner point corresponding to Γ in the approximated SR and that of the exact SR. We verified that $\epsilon\left(\Gamma\right) \geq 0$ for all $\Gamma \in \Omega_\Gamma^{ss}$, hence $\tilde{\mathcal{R}}_c^{ss}$ presents an upper bound of the exact stability region $\mathcal{R}_c^{ss}$. As one can see in the numerical section, the epsilon difference between the exact stability region and its upper bound is small. Hence, we highly reduce the complexity by finding an explicit and close upper bound of the exact stability region.

3.3.4 D2D simple scenario

Proposition 1. *The stability region for the 3-UEs D2D scenario is the set of $(\lambda_s, \lambda_u) \in \mathcal{R}_d^{ss}$ such that:*

$$\mathcal{R}_d^{ss} = co\left(\bigcup_{\Gamma\in\Omega_\Gamma^{ss}} \left\{\mu_s^d\left(\Gamma\right), \mu_u^d\left(\Gamma\right)\right\}\right)$$

where the queues' service rates $\mu_s^d\left(\Gamma\right)$ and $\mu_u^d\left(\Gamma\right)$ are respectively given by (3.14) and (3.15) and the values of the parameters U, V, C and D corresponding to the six priority policies $\Gamma \in \Omega_\Gamma^{ss}$ are given in table 3.1.

$$\mu_s^d := \mu_s^d\left(\Gamma\right) = r_1 p_{sd}^1 U + r_2 p_{sd}^2 V \tag{3.14}$$

$$\mu_u^d := \mu_u^d\left(\Gamma\right) = r_1 p_u^1 C + r_2 p_u^2 D \tag{3.15}$$

Proof. See Appendix-7.3.4. □

3.3.5 Comparison between cellular and D2D 3-UEs scenarios

We provide an analytic comparison between the D2D and cellular 3-UEs scenarios. We assume that the three users are at the same distance d from the BS and we find for each distance d, the D2D distance threshold between the UE_s and UE_d under which D2D is more advantageous than cellular communication. $\mu 1(\alpha, \Gamma)$ depends on the distance d; thus this threshold d_{th} is provided as function of d. To make expressions simpler, we present the formula of the threshold pathloss PL_{th} and not that of d_{th}. However, given the path-loss model that relate the $d_{th}$to PL_{th}, the expression of d_{th} can be simply deduced from that of PL_{th}.

Theorem 3.3.5. *For a given priority policy ξ, the threshold D2D path-loss $PL_{th}(\xi)$ under which D2D communication is more advantageous than cellular communication for each distance d is given as follows;*

$$PL_{th}(\xi) = \left(\frac{V(\xi) r_2}{\mu_1(\xi)}\right)^{\frac{-1}{b}} z^{\frac{-1}{a}} y \tag{3.16}$$

where

$$a = \frac{\gamma_1^{UL} N_0}{P_{UL}}, b = \frac{\gamma_2^{UL} N_0}{P_{UL}}, z = (r_1 U(\xi) + r_2 V(\xi)) \left(\frac{V r_2}{\mu_1}\right)^{-\frac{a}{b}}$$

and

$$y = \frac{1}{a} \sum_{n=0}^{\infty} \frac{\Gamma\left(\frac{1+nb}{a}\right)}{\Gamma\left(\frac{1+nb}{a} + 1 - n\right) n!} (-1)^n z^{-\frac{b}{a}n}$$

Proof. See appendix-7.3.5. □

3.4 Multi-Users scenario

For the general scenario, we consider K UE2UE communications between pairs of nearby users ($UE_{i,s}$ and $UE_{i,d}$) and U UE2BS communications between BS and $UE_{i,u}$. In total $2K + U$ users are considered in the cell. For the multi-user case, the study is applied for a set of two bit rates $[r_1, r_2]$ with $r_2 = 0$. Unfortunately, considering a set of 3 bit rates as in section 3.3 is complex and hard to compute and to simplify. Even for the case of two bit rates, the characterization of the exact SR, given by theorem 3.4.1, remains computationally complex for the multi-user case. Therefore, we reduce the complexity of the exact SR in theorem 3.4.2). In addition, we propose in theorem 3.4.5 a simple approximation of the exact SR that is characterized by the following: (i) highly reducing the complexity of the exact SR and (ii) being an ϵ-close approximation of the exact SR (i.e. the distance between the approximated and exact SR is upper bounded by a small ϵ). The importance of this result lies on shifting a very complex problem to a simple approximated one with a low complexity

and a high precision. A trade-off exists between the precision of the stability region and the complexity of characterizing this region.

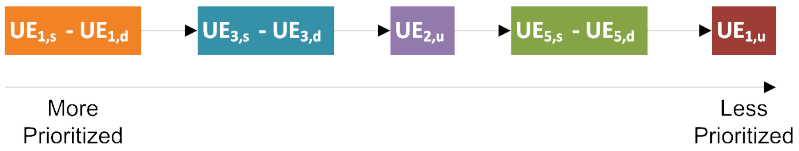


Figure 3.5: Example of one priority policy for $K = 3$ and $U = 2$.

In order to simplify the presentation of the results in the sequel, we use the following notation of the transmission probabilities at rate r_1: $p_{i,s} = p^1_{i,s}$, $\bar{p}_{i,s} = 1-p^1_{i,s}$, $p_{i,d} = p^1_{i,d}$, $\bar{p}_{i,d} = 1 - p^1_{i,d}$, $p_{i,u} = p^1_{i,u}$, $\bar{p}_{i,u} = 1 - p^1_{i,u}$, $p_{i,sd} = p^1_{i,sd}$ and $\bar{p}_{i,sd} = 1 - p^1_{i,sd}$. The vector that describes the arrival rates of the users' queues is given by:

$$\boldsymbol{\lambda} = [\lambda_{1,s} \;\; \lambda_{2,s} \;\; ... \;\; \lambda_{K,s} \;\; \lambda_{1,u} \;\; \lambda_{2,u} \;\; ... \;\; \lambda_{U,u}]$$

We denote by Γ the priority policy under which the users are arranged. The scheduling policy consists on choosing a UE if and only if all the higher-priority UEs are not able to transmit. We denote by Ω_Γ the set of all the possible sorting of the users. The number of existing priority policies consists of the number of the possible permutation of $K + U$ communications. Hence, $|\Omega_\Gamma| = (K + U)!$. We define the following two sets for each communication $i \in \{1, ..., K + U\}$:

- $\mathbf{U}_\Gamma\{i\} = \{$ UE2BS communications higher-priority than the communication i under a priority sorting $\Gamma\}$
- $\mathbf{K}_\Gamma\{i\} = \{$ UE2UE communications higher-priority than the communication i under a priority sorting $\Gamma\}$

In Fig. 3.5, we present an example that illustrate one priority policy for a scenario of $K = 3$ UE2UE communications and $U = 2$ UE2BS communications. In this example, the considered priority policies gives the highest priority to the first UE2UE communication $(UE_{1,s} - UE_{1,d})$ and the lowest priority to the first UE2BS communication $(UE_{1,u} - BS)$. Therefore, the sets $\mathbf{U}_\Gamma\{4\}$ and $\mathbf{K}_\Gamma\{4\}$ of the $UE_{1,u}$ are the following: $\mathbf{U}_\Gamma\{4\} = \{2\}$ and $\mathbf{K}_\Gamma\{4\} = \{1, 3, 5\}$. The sets $\mathbf{U}_\Gamma\{5\}$ and $\mathbf{K}_\Gamma\{5\}$ of the $UE_{5,s} - UE_{5,d}$ communication are the following: $\mathbf{U}_\Gamma\{2\} = \{2\}$ and $\mathbf{K}_\Gamma\{5\} = \{1, 3\}$.

3.4.1 Cellular scenario

Here, we derive the SR of the multi-UEs cellular scenario where only cellular communications are possible and D2D is not allowed. Devices that desire to communicate with each other must exchange their packets through the BS. No direct communication is possible between nearby devices.We consider the following notation of the service rate:

- $\mu_{i,s}(\Gamma)$: service rate of the uplink of the i^{th} UE2UE communication
- $\mu_{j,u}(\Gamma)$: service rate of the downlink of the j^{th} UE2BS communication

The traffic departure is time varying and depends on the scheduling allocation decision and the time varying channel conditions. The departure average rates from all the users queues for the cellular scenario is denoted by:

$$\boldsymbol{\mu}(\Gamma) = [\mu_{1,s}(\Gamma) \;\; \mu_{2,s}(\Gamma) \;\; ... \;\; \mu_{K,s}(\Gamma) \;\; \mu_{1,u}(\Gamma) \;\; \mu_{2,u}(\Gamma) \;\; ... \;\; \mu_{U,u}(\Gamma)]$$

We assume a user scheduling such that only one communication is possible in each time-slot. $Q_{i,BS}$ might be empty at some time-slot, therefore when the $UE_{i,s}$-$UE_{i,d}$ communication is scheduled then the choice between UL ($UE_{i,s}$ to BS) or DL (BS to $UE_{i,d}$) depends not only on the SNR states of these two links but also on the state empty or not of the corresponding queue at the BS level, i.e. $Q_{i,BS}$. In order to take that into account, we introduce in the analysis a new parameter $\alpha_i \in [0,1]$ for each UE2UE communication ($1 \leq i \leq K$) that defines α_i as the fraction of time that the resources are allocated to the UL ($UE_{i,s}$ to BS) and $1-\alpha_i$ as the fraction of time that the resources are allocated to the DL (BS to $UE_{i,d}$). We denote by $\boldsymbol{\alpha}$ the fraction vector that considers all the UE2UE communications.

$$\boldsymbol{\alpha} = [\alpha_1, \alpha_2, ..., \alpha_K]$$

3.4.2 Organization

In the sequel, we start by characterizing the exact SR of the cellular multi-UEs scenario in section 3.4.3. This region is given by considering all the feasible priority policies $\Gamma \in \Omega_\Gamma$ and all the values of $\boldsymbol{\alpha} \in [0,1]^K$ are considered. Since computing this region is highly complex. We start by reducing the complexity, in section 3.4.4, by limiting the number of fraction vector $\boldsymbol{\alpha}$ that should be considered for characterizing the SR. Furthermore, an epsilon-close approximation with a simple explicit form and a very low complexity is proposed. A trade-off is elaborated between the complexity and the precision of the SRcomputation. In addition, the SR of the D2D multi-UEs scenario is given in section 3.4.5.

3.4.3 Exact stability region

Theorem 3.4.1. *For the multi-UEs cellular scenario, the exact stability region is given by the set of $\boldsymbol{\lambda} \in \mathcal{R}_c$ such that:*

$$\mathcal{R}_c = co\left(\bigcup_{\Gamma\in\Omega_\Gamma} \bigcup_{\alpha\in[\mathbf{0},\alpha^*]} \{\boldsymbol{\mu}(\boldsymbol{\alpha},\Gamma)\}\right)$$

where $\mathbf{0}$ is the vector zero in $\mathbb{R}^K$, $\alpha^* = (\alpha_1^*, \alpha_2^*, ..., \alpha_K^*)$ with $\forall 1 \leq i \leq K : \alpha_i^* = \frac{p_{i,d}-p_{i,s}+p_{i,s}p_{i,d}}{2p_{i,s}p_{i,d}}$. Ω_Γ is the set of all the possible priority policies (with $|\Omega_\Gamma| = (K+U)!$).

The elements of $\boldsymbol{\mu}(\Gamma)$ *which are* $\mu_{i,s}(\Gamma)$ *and* $\mu_{j,u}(\Gamma)$ *(for* $1 \leq i \leq K$, $1 \leq j \leq U$*) are respectively given by (3.17) and (3.18).*

$$\mu_{i,s}(\Gamma) = \frac{1}{2} r_1 p_{i,s} \left(1 + \frac{\bar{p}_{i,s} p_{i,d}}{-2\alpha_i p_{i,s} p_{i,d} + (1 + p_{i,s}) p_{i,d}}\right) \prod_{m \in \mathbf{U}_\Gamma\{i\}} \bar{p}_{m,u} \times \prod_{n \in \mathbf{K}_\Gamma\{i\}} \bar{p}_{n,s} \left[1 + \frac{p_{n,s} p_{n,d}}{2\alpha_n p_{n,s} p_{n,d} - (1 + p_{n,s}) p_{n,d}}\right] \tag{3.17}$$

$$\mu_{j,u}(\Gamma) = r_1 p_{j,u} \prod_{m \in \mathbf{U}_\Gamma\{j+K\}} \bar{p}_{m,u} \prod_{n \in \mathbf{K}_\Gamma\{j+K\}} \bar{p}_{n,s} \left[1 + \frac{p_{n,s} p_{n,d}}{2\alpha_n p_{n,s} p_{n,d} - (1 + p_{n,s}) p_{n,d}}\right] \tag{3.18}$$

Proof. See Appendix-7.3.7.

□

The exact stability region is computed by following the instructions below:

- For all $\Gamma \in \Omega_\Gamma$ (with $|\Omega_\Gamma| = (K+U)!$)
  - For all $\boldsymbol{\alpha} \in [0, \boldsymbol{\alpha}*]$
    - Compute the service rates of the queues
    - Deduce a point of the exact SR
- Exact SR as the convex hull of all these points

Based on the steps above for the exact SR computation, we deduced the high complexity of this approach. Indeed, the complexity of the exact stability region computation comes from two factors: the high number of priority policies (depending on the number of communications) as well as the large number of the fraction vector $\boldsymbol{\alpha}$ values that should be considered for each priority policy. Having $K+U$ communications means that there exists $(K+U)!$ possible policies to sort them. Further, for each priority Γ, we should vary $\boldsymbol{\alpha} \in [\mathbf{0}, \boldsymbol{\alpha}^*]$. Suppose that for each α_i (for $1 \leq i \leq K$) we consider L different values within $[0, \alpha_i^*]$. Thus, for each priority policy we consider L^K values of the fraction vector $\boldsymbol{\alpha}$. It is clear that bigger is L higher is the precision of the SR (i.e. in numerical section $L \geq 10^3$). We deduce that even by considering only a set of two bit rates $[r_1, r_2 = 0]$, the complexity of the stability region computation remains high: $L^K (K+U)!$.

Therefore, we start reducing the complexity of the exact SR by verifying that only the border points of $\alpha_i \in \{0, \alpha_i^*\}$ (for $1 \leq i \leq K$) are sufficient for characterizing the exact SR. Furthermore, we propose an approximation of the exact SR that highly decreases the complexity with a tight loss of precision. This study shows a trade-off between the precision of the SR computation and its complexity.

3.4.4 Precision versus complexity

The main challenge is to reduce the complexity of the stability region computation while guaranteeing its precision. As we have discussed before, the complexity of the exact stability region is mainly caused by the high number of the following two parameters that should be considered for the characterization of the SR: (i) priority policies and (ii) fraction vector $\boldsymbol{\alpha}$ values. These two factors of complexity are respectively studied in this section. In the first part, we reduce the complexity by limiting the values of the fraction vector $\boldsymbol{\alpha}$ that should be considered for each priority policy. We prove that a limited set of values of the fraction vector $\boldsymbol{\alpha}$ is sufficient for characterizing the exact SR. However, the complexity remains high due to the large number of the priority policies that should be considered $(K+U)!$. In the second part we propose an epsilon-close approximation of the exact SR of the symmetric case (see definition 3.4.3). This approximation highly reduces the complexity by limiting the number of the considered priority policies while guaranteeing a high precision.

Theorem 3.4.2. *For the multi-UEs cellular scenario, the exact stability region is the set of $\boldsymbol{\lambda} \in \mathcal{R}_c$ such that $\mathcal{R}_c$ can be simplified as follows:*

$$\mathcal{R}_c = co\left(\bigcup_{\Gamma\in\Omega_\Gamma}\bigcup_{\alpha\in\mathbb{S}_\alpha}\{\boldsymbol{\mu}(\boldsymbol{\alpha},\Gamma)\}\right)$$

with

$$\mathbb{S}_\alpha = \{\boldsymbol{\alpha}\,|\,\alpha_i \in \{0,\alpha_i^*\}\ \ \forall\, 1\le i\le K\}$$

where the elements of $\boldsymbol{\mu}(\Gamma)$ which are $\mu_{i,s}(\Gamma)$ and $\mu_{j,u}(\Gamma)$ (for $1\le i\le K$, $1\le j\le U$) are respectively given by (3.17) and (3.18) and the limited set $\mathbb{S}_\alpha$ of 2^K elements reduces the complexity of the SR.

Proof. See Appendix-7.3.8. □

In theorem 3.4.2, the complexity of the exact stability region is decreased due to the fact that for each priority policy $\Gamma \in \Omega_\Gamma$, the border values of each α_i are sufficient for characterizing the corner points of the exact SR that correspond to Γ. Indeed, for a given priority policy $\Gamma \in \Omega_\Gamma$, the service rates vector $\boldsymbol{\mu}(\boldsymbol{\alpha},\boldsymbol{\Gamma})$ for any $\boldsymbol{\alpha} \in [\mathbf{0},\boldsymbol{\alpha}^*]$ can be written as a convex combination of the service rates vectors $\boldsymbol{\mu}(\boldsymbol{\alpha},\boldsymbol{\Gamma})$ with $\boldsymbol{\alpha} \in \mathbb{S}_\alpha$. Actually, the L values of each α_i that we should consider in theorem 3.4.1 for each priority policy Γ is reduced to 2 values in theorem 3.4.2 (i.e. 0 and α_i^*). The proof of the theorem 3.4.2 shows that varying $\boldsymbol{\alpha}$ within the finite set $\mathbb{S}_\alpha$ is sufficient for the characterization of the exact stability region. Hence, the computational complexity is reduced from $L^K(K+U)!$ to $2^K(K+U)!$.

However, the complexity of the exact SR remains high because of the large number of the priority policies that should be considered $(K+U)!$. Hence, for the symmetric case (i.e. defined in (3.4.3)), we reduce further the complexity by limiting the number of the priority policies that is sufficient for the characterization of the SR. An ϵ-approximation, such that the maximum distance between the approximated and exact

SR is limited to ϵ, is proposed. The main importance of this approximation is that it highly reduces the complexity (i.e. in terms of number of considered priority policies and number of fraction vector values) while guaranteeing the precision of the result.

Definition 3.4.3. *The symmetric case consists of considering all the UEs of the cell at the same distance d from the BS. Hence, this symmetric case is defined by the following values:* $p_s = p_{i,s} = p_{j,u}$ *and* $p_d = p_{i,d}$ *(*$\forall 1 \leq i \leq K$ *and* $1 \leq j \leq U$*).*

Definition 3.4.4. $\tilde{\mathcal{R}}$ *is called an* ϵ*-approximation of a stability region* $\mathcal{R}$ *(with* $0 \leq \epsilon \leq 1$*) iff the following is verified:*

$$\tilde{\mathcal{R}} \subseteq \mathcal{R} \subseteq \tilde{\mathcal{R}} + \epsilon$$

The symmetric case is defined as the case where all the UEs are at the same distance d from the base station. For this case, we prove in theorem 3.4.5 that we can additionally reduce the computation complexity by limiting the number of policies Γ that should be considered in order to characterize theSR. Thus, we avoid the computation complexity that comes from the consideration of all the policies $\Gamma \in \Omega_\Gamma$ (i.e. for a network of $K+U$ communications, there exists $|\Omega_\Gamma| = (K+U)!$ priority policies overall). Therefore, we propose an ϵ-approximation of the exact SR that highly reduces the complexity while guaranteeing a high precision.

Theorem 3.4.5. *For the symmetric case of the multi-UEs cellular scenario with a given couple* $\{p_s, p_d\}$*, for a given* ϵ*, an* ϵ*-approximation of the stability region is the set of* $\boldsymbol{\lambda} \in \tilde{\mathcal{R}}_c$ *such that:*

$$\tilde{\mathcal{R}}_c = co\left(\bigcup_{\Gamma \in \Omega_\Gamma^{K_0}} \bigcup_{\alpha \in \mathbb{S}_\alpha^{K_0}} \{\boldsymbol{\mu}(\boldsymbol{\alpha}, \Gamma)\}\right) \text{ with } K_0(\epsilon) = \left\lceil 1 + \frac{\log\left(\frac{\epsilon}{r_1 p_s}\right)}{\log(\bar{p}_s)} \right\rceil$$

Hence, the exact stability region can be bounded as follows:

$$\tilde{\mathcal{R}}_c \subseteq \mathcal{R}_c \subseteq \tilde{\mathcal{R}}_c + \epsilon$$

where $\Omega_\Gamma^{K_0}$ *is the set of the feasible priority policies of the subsets of* K_0 *communications among all the* $K+U$ *communications (with* $|\Omega_\Gamma^{K_0}| = \frac{(K+U)!}{(K+U-K_0)!}$ *),* $\mathbb{S}_\alpha^{K_0}$ *is the set of the values* $\alpha_i \in \{0, \alpha_i^*\}$ *where* i *corresponds to the UE2UE communications within these subsets of* K_0 *elements. The elements of* $\boldsymbol{\mu}(\Gamma)$ *which are* $\mu_{i,s}(\Gamma)$ *and* $\mu_{j,u}(\Gamma)$ *(for* $1 \leq i \leq K$*,* $1 \leq j \leq U$*) are respectively given by (3.17) and (3.18). Note that the value of* K_0 *is limited to* $(K+U)$ *which corresponds to the total number of communications.*

Proof. See Appendix-7.3.9.

□

The proposed approximation in theorem 3.4.5 is defined by considering all the following cases: (i) the priority policies of all the subset of K_0 elements within the $K+U$ communications, (ii) for each priority policies, the UE2UE communications within these K_0 elements have a fraction element that varies within its two border values (i.e. 0 and α_i^*). The value of K_0 depends on the precision ϵ of the approximation; it increases when ϵ decreases. Even for a tight value of ϵ, K_0 stays too much lower than $K+U$ which reduces the computational complexity to: $2^{K_0}\frac{(K+U)!}{(K+U-K_0)!}$. We can see that K_0 is constant for a given triplet $\{p_s, p_d, \epsilon\}$ and does not depend on the number of communications. Therefore, the importance of this approximation is to reduce the exponential complexity $\mathrm{O}\left(2^K\right)$ for characterizing the exact stability region to a polynomial complexity $\mathrm{O}\left(K^{K_0}\right)$.

3.4.5 D2D scenario

For the D2D scenario, both D2D and cellular communication are allowed. We assume that UE2UE communications are established directly between the nearby UEs without passing by the BS while the users of the UE2BS communications are connected to the BS. In this scenario the following links exist: $\text{link}_{i,sd}^d$: $\text{UE}_{i,s}$ - $\text{UE}_{i,d}$ and $\text{link}_{j,u}^d$: $\text{UE}_{j,u}$ - BS (with $1 \leq i \leq K$ and $1 \leq j \leq U$).

The traffic departure is time varying and depends on the scheduling allocation decision and the time varying channel conditions. The departure average rates from all the users queues for the D2D scenario is denoted by:

$$\boldsymbol{\mu^d}(\Gamma) = \left[\mu_{1,s}^d(\Gamma) \;\; \mu_{2,s}^d(\Gamma) \;\; \ldots \;\; \mu_{K,s}^d(\Gamma) \;\; \mu_{1,u}^d(\Gamma) \;\; \mu_{2,u}^d(\Gamma) \;\; \ldots \;\; \mu_{U,u}^d(\Gamma)\right]$$

Theorem 3.4.6. *For the multi-UEs D2D scenario, the stability region is the set of* $\boldsymbol{\lambda} \in \mathcal{R}_d$ *with:*

$$\mathcal{R}_d = co\left(\bigcup_{\Gamma\in\Omega_\Gamma}\{\boldsymbol{\mu}^d(\Gamma)\}\right)$$

where the elements of $\boldsymbol{\mu}^d(\Gamma)$ *which are* $\mu_{i,s}^d(\Gamma)$ *and* $\mu_{j,u}^d(\Gamma)$ *(for* $1 \leq i \leq K$, $1 \leq j \leq U$*) are respectively given by (3.19) and (3.20).*

$$\mu_{i,s}^d(\Gamma) = r_1 p_{i,sd} \prod_{m\in\mathbf{U}_\Gamma\{i\}} \bar{p}_{m,u} \prod_{n\in\mathbf{K}_\Gamma\{i\}} \bar{p}_{n,sd} \tag{3.19}$$

$$\mu_{j,u}^d(\Gamma) = r_1 p_{j,u} \prod_{m\in\mathbf{U}_\Gamma\{i\}} \bar{p}_{m,u} \prod_{n\in\mathbf{K}_\Gamma\{i\}} \bar{p}_{n,sd} \tag{3.20}$$

Proof. See Appendix-7.3.10. □

3.5 Numerical Results

Simulations are investigated in order to validate our theoretical results and compare D2D and cellular scenarios. We consider a cell of radius $R = 500\,m$. We assume a 10 MHz LTE-like TDD system. When a user is scheduled, it transmits over 50 subcarriers called Resource Block (RB)s. The considered bit rates per RB are $\{r_1 = 400,\, r_2 = 200,\, r_3 = 0\}$ $kbps/RB$. The unit of the shown results is the arrival rate per RB [kbps/RB] from which the total bit rate can be deduce by multiplying by the number of allocated RBs. The total transmission power (i.e. over all RBs) of the BS is $P_{DL} = 40\,W$ and that of the mobiles is $P_{UL} = 0.25\,W$ (see [18]). The noise power is equal to $-195\,dB/Hz$ for the DL and $-199\,dB/Hz$ for UL. We consider the pathloss model specified in [59]. The SNR thresholds are respectively given by: $\gamma_1^{UL} = 9.5\,dB$, $\gamma_2^{UL} = 2.5\,dB$, $\gamma_1^{DL} = 7.5\,dB$ and $\gamma_2^{DL} = 1.5\,dB$. These values are practical values based on a throughput-SNR mapping for a 10 MHz E-UTRA TDD network (see [60]). The minimum and maximum D2D distance between two users are $d_{min} = 3\,m$ (see [18]) and $d_{max} = 350\,m$ (see [61]).

3.5.1 Three-UEs scenario

We start by presenting the results of three-UEs scenario in order to validate the theoretical results corresponding to this case. For clarity of the presentation, we assume that the three users UE_s, UE_u and UE_d are at the same distance d from the BS. This assumption is just for simplifying the illustration of the results however the theoretical results can be applied for any distribution of the users in the cell.

3.5.1.1 Stability region

We evaluate numerically the SR for the cellular and D2D scenarios. First, we validate that $\tilde{\mathcal{R}}_c^{ss}$ (given by (3.3.3)) is a close approximation of the exact SR for the cellular scenario $\mathcal{R}_c^{ss}$ (given by (3.3.1). We present in Fig.3.6 that the SR and its approximation coincide for different distances d between UEs and BS.

Second, we illustrate the stability region of both scenarios for different cases in order to study the performance evolution as function of the distance between the UEs and the BS. In Fig. 3.7, for different distances d between UEs and BS $d = \{100, 200, 350, 500\}$, we plot in Red the SR for the cellular scenario obtained by exhaustive search (all $\Gamma \in \Omega_\Gamma$ and all $\alpha \in [0,1]^4$) and we compare it to the one plotted in Green and obtained from Theorem 3.3.3. We find that both curves coincide which verifies that the latter theorem presents the characterization of the cellular scenario SR. In addition, for each distance d we plot (in Blue) the SR for the D2D scenario by considering two different D2D distances d^d between UE_u and UE_s. This figure gives insights in the effect of the distance d on the cellular stability region as well as the effect of the D2D distance d^d on the D2D stability region. It illustrates that D2D mode is more advantageous for small d^d distances while cellular mode is

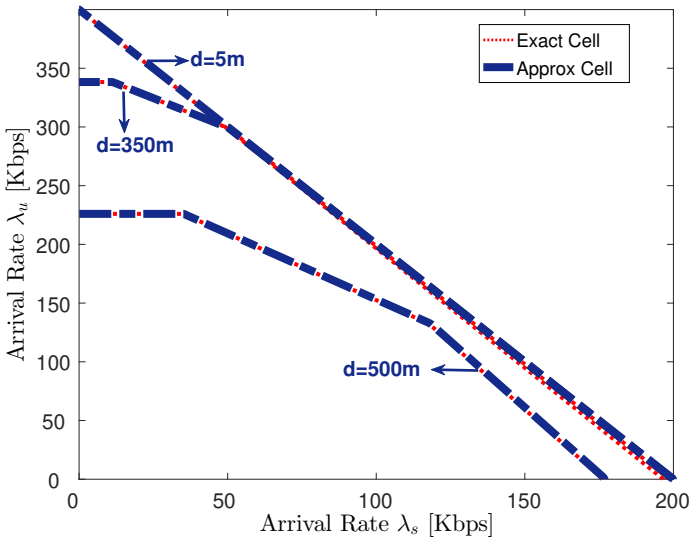


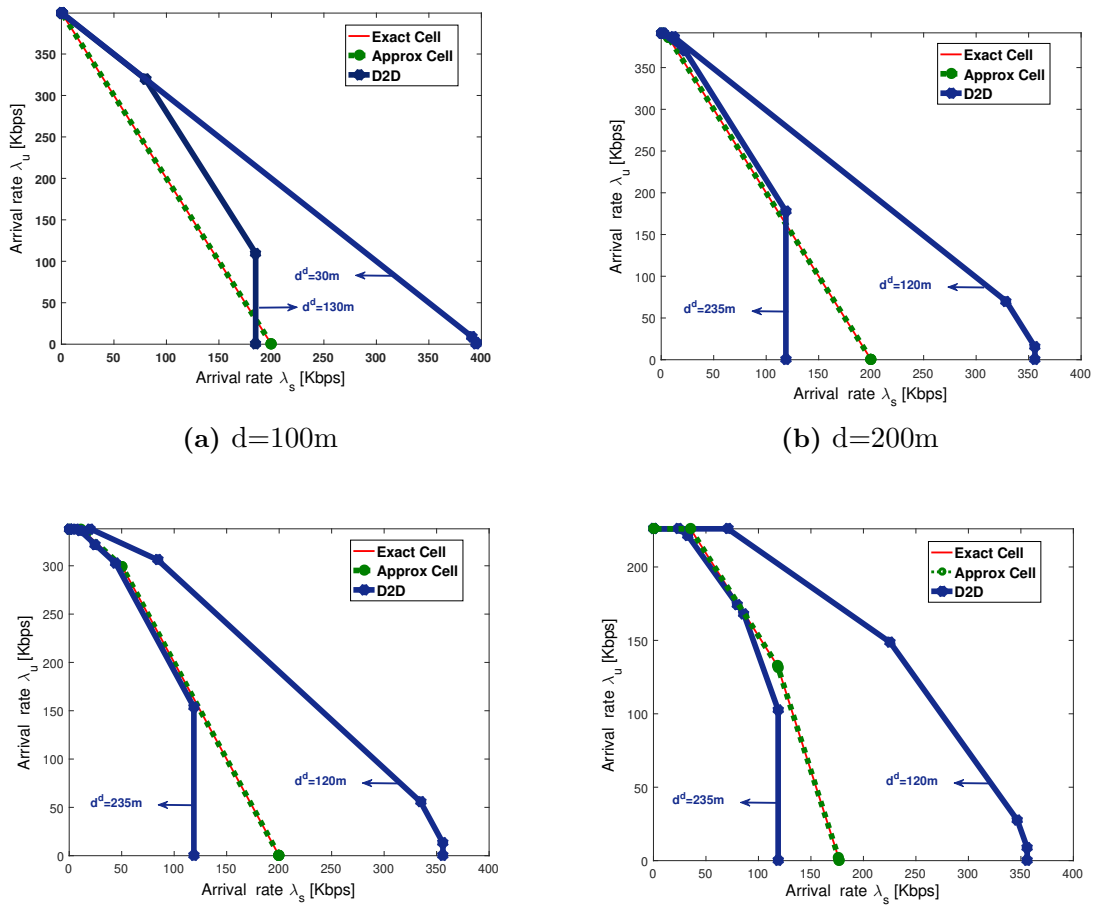
Figure 3.6: Comparison between the stability region for the cellular scenario $\mathcal{R}_c^{ss}$ and its approximation $\tilde{\mathcal{R}}_c^{ss}$

more advantageous otherwise. Fig. 3.7 shows that for each distance d, there exists a different d^d threshold (called d_{TH}) under which D2D is more advantageous than cellular communications. In order to elaborate a mode selection framework, we need to provide the value of this distance threshold d_{TH} as function of the UE to BS distance d.

3.5.1.2 Queuing impact

We start by showing the advantages of the queuing theory approach in our analysis. For this purpose, we consider a random positioning of 3 UEs in the cell and we apply the following two performance evaluation approaches applied to the cellular scenario: (1) taking into account the queuing aspects and the coupling between the queues and (2) without considering any coupling between the queues. For the first approach, we compute the stability region as the set of the arrival rate vectors to the sources that can be stably supported by the network considering all the possible policies, hence we consider the coupling between the queues (w/ coupling). For the second approach, we assume that queues have a full buffer and we compute the rate region that describes the achievable data rates depending on the channel states of the links and without any bursty traffic neither coupling between the queues (w/o coupling). Comparing these two results in Fig. 3.8 verifies that introducing the traffic pattern and the queues' coupling have an effect on the performance evaluation of the cellular scenario stability region.

Considering that the queues have a full buffer, the DL (BS-to-UE$_d$) has always packets to transmit and it will always be considered by the scheduler. This causes less scheduling for the UL thus a lower bit rate over the UL. However, when coupling between the queues is considered, the DL can be scheduled only if at least one packet



(a) d=100m **(b)** d=200m

(c) d=350m **(d)** d=500m

Figure 3.7: Stability regions for both D2D and cellular scenarios for UE to BS distance d in {100, 200, 350, 500}m and two different D2D distance d^d m

is received by the BS from UE$_s$. For this reason, when no packets are buffered at the DL, this latter will not be scheduled and the UL will be able to transmit more packets and will improve its bit rate. This explains the gain that the queuing approach offers to the performance evaluation of the cellular scenario.

3.5.1.3 Comparison D2D vs. cellular

From the results above, we note that the maximum D2D distance d_{TH} for which D2D outperforms cellular communications depends on the distance d between UEs and the BS. When the users are far from the base station then the performance of cellular communications degrades while the performance of D2D communication remains unaffected since it depends on the distance between D2D peer. Therefore, for each distance d, there exists a D2D-distance d_{TH} for which D2D peers at a distance $< d_{TH}$ are recommended to transmit in a D2D mode while otherwise cellular mode is recommended. It is interesting to determine this threshold, for each distance d, in order to elaborate a mode selection scheme between D2D and cellular mode. For this aim, we give the values of the D2D distance threshold d_{TH} for each priority policy

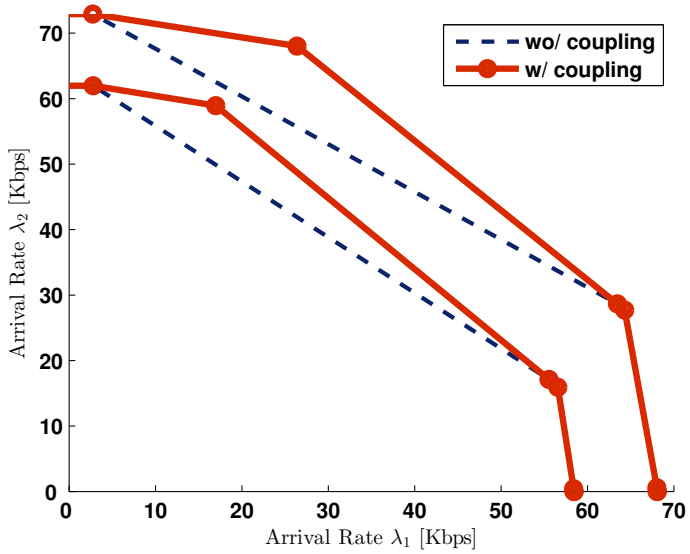


Figure 3.8: Stability region of the cellular scenario (with coupling between the users' queues and the state empty or not of Q_{BS}) vs. rate region (without coupling hence Q_{BS} is considered always not empty)

$\Gamma \in \Omega^*_\Gamma$. Fig. 3.9 shows the variation of the D2D distance threshold d_{TH} as function of the distance d between UEs and BS for all the priority policies $\Gamma \in \Omega^*_\Gamma$ (i.e. policies corresponding to the border points of the SR). D2D region represents the couples of distances (d, d^d) for which D2D outperforms cellular communications while cellular region represents the other cases.

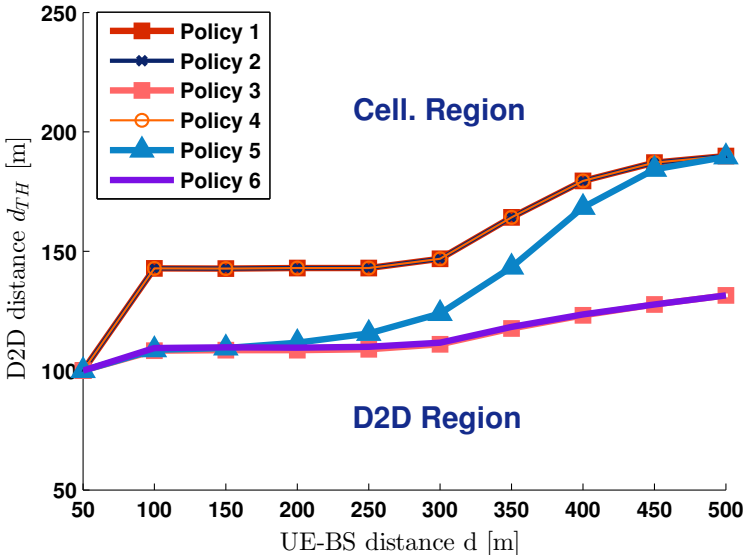


Figure 3.9: Max D2D distance for which D2D outperforms cell. comm.

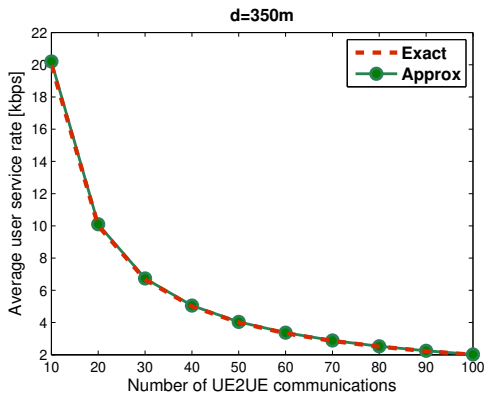
3.5.2 Multi-UEs scenario

For the multi-UEs scenario, without loss of generality and for clarity reasons, we illustrate the results for the symmetric case where all the users are at the same distance d from the BS. We consider a cell containing 50 UE2UE communications with a uniform random drop of the users in a cell of radius $R = 500m$ (see Fig. 3.11a).

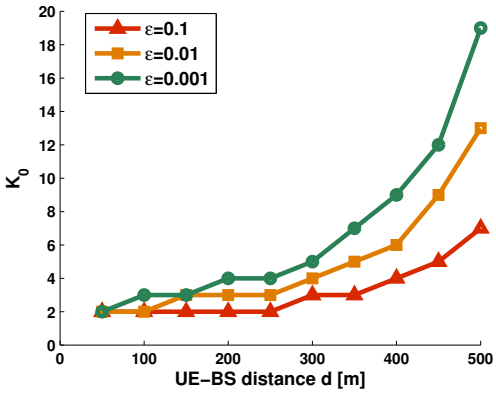
Exact vs approximated stability region: For this comparison we consider all the UEs at a distance $d = 350m$ from the BS. The performance metric that we use to compare between the SR of the exact symmetric case (i.e. from theorem 3.4.2) and the approximated symmetric case (i.e. from theorem 3.4.5) is the average service rate per user. It is equal to the sum of the service rates of all the users divided by the number of users. For $d = 350$m and $\epsilon = 0.1$ we find $K_0 = 3$, figure 3.10a illustrates that the variation of this performance metric as function of the number K of UE2UE communications in the network. It shows that the ϵ-approximated stability region is a very tight approximation of the stability region. We deduce that the complexity is reduced from $2^{50} \times 50! \simeq 3 \times 10^{79}$ for the exact SR (i.e. from theorem 3.4.2) to $2^3 \times \frac{50!}{47!} = 9 \times 10^5$ for the approximated SR (i.e. from theorem 3.4.5). This illustrates how the exponential complexity $\mathrm{O}(2^{50})$ is reduced to a polynomial one $O\left(50^3\right)$.

Trade-off complexity vs precision For the symmetric case, where all the users are considered at the same distance d from the BS, we show in theorem 3.4.5 that the value of K_0 characterizes the number of priority policies that is sufficient for defining the SR, hence the complexity of the approximated SR. For three different $\epsilon \in \{10^{-3}, 10^{-2}, 10^{-1}\}$, we show in Fig.3.10b the variation of K_0 as function of the distance d between the users and the BS. We can see that for $\epsilon = 10^{-2}$, the maximum value of K_0, achieved at the edge of the cell, is small and equal to 12. Thus, the complexity of the approximated SR can be deduced as follows:

$$2^{K_0} \frac{(K+U)!}{(K+U-K_0)!}$$

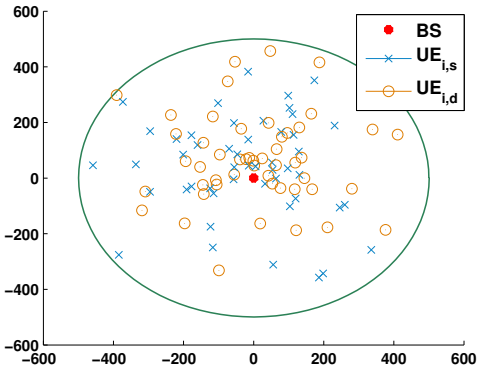


(a) Exact vs. Approximated results

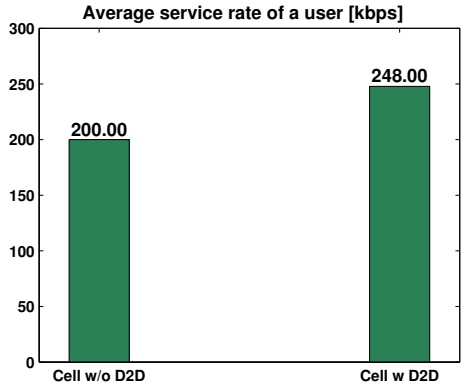


(b) Variation of the parameter K_0 as function of the distance d between BS and users

Comparison D2D vs. cellular We compare between both cases: 1. only cellular communications are allowed and 2. cellular communications are aided by the overlaid D2D communications. Considering the average service rate of a user as the performance metric, Fig. 3.11b shows how the performance can be improved by adding the use of overlaid D2D to cellular networks. An improvement of nearly 25 percent is illustrated.



(a) Uniform random drop of 50 users in a cell of radius 500m



(b) Average service rate of a user in the cellular scenario with and without adding the aid of the overlay D2D

3.6 Conclusion

In this chapter, we have carried out a queuing analysis to evaluate overlaid D2D communications. We have considered two scenarios: (i) cellular scenario where the UE2UE communications between nearby devices pass through the base station and (ii) the overlay D2D scenario where nearby users can communicate directly through an overlaid D2D link. First, we have shown the interest of our queuing theory approach by verifying that it provides a more realistic performance evaluation of the network compared to a performance analysis which is strictly based on physical layer and ignores the dynamic effect of the traffic pattern. The queuing based analysis is motivated by the coupling that exists between the UEs queues and the BS queues. Second, we give the simple analytic expressions of the stability regions for both scenarios as convex polytopes with limited number of vertices and we study the variation of these stability regions as a function of the queues and channels states. Finally, these results allow us to elaborate the cases where D2D links are more favorable than traditional cellular link. This study is used to come up with a mode selection scheme between D2D mode and cellular mode that proves an improvement in terms of network capacity.

4 | D2D Resource Allocation

Employing channel adaptive resource allocation can yield to a large enhancement in almost any performance metric of D2D communications (e.g. Energy Efficiency). We observe that D2D users have the knowledge of their local channel state which is not the case of the Base Station (BS) that requires a CSI reporting to acquire this information. Based on this observation, we compare between centralized and distributed scheduling in the aim of minimizing the D2D users transmission power while maintaining predefined throughput constraint in a limited feedback D2D-enabled cellular network. We start by proposing an energy efficient centralized scheduling that requires the CSI knowledge of the D2D links at the BS level. This CSI reporting suffers from the limited resources available for feedback transmission. Therefore, we benefit from the D2D users' knowledge of their local CSI in order to propose a distributed scheduling that improves the energy efficiency of D2D communications. The key idea is that D2D users compute their local energy efficiency metric and then use some CSI indicators to share these metrics' values between each other. Since a distributed approach is considered, collision may occur during the exchange of CSI indicators. Thus, we develop a collision reduction mechanism that allows the distributed algorithm to achieve a performance closed to that of the *ideal* scheduling (i.e. with a global CSI knowledge of all the D2D links). For instance, we describe how both the centralized and distributed algorithms can be simply integrated to existing LTE cellular networks. Numerical results show that the distributed scheduling outperforms the centralized one.

4.1 Concept and Related Work

In cellular networks, the knowledge of the channel condition at the transmitter can improve the performance of cellular communications by allowing the transmitter to dynamically adapt its transmission scheme and providing by that a better throughput (i.e. application of an Adaptive Modulation Coding (AMC) scheme). In the case of D2D communications, each D2D user has the knowledge of its local D2D channel state. For both TDD and Frequency Division Duplex (FDD) cellular networks, *feedback* is one way for keeping the BS updated with the D2D channels' measurements. Mobile users estimate their D2D link and feed it back to the BS that benefits from this knowledge to optimize the performance of D2D communications. However, this feedback is imperfect since a limited number of resources are available for the exchange of control

information. In a limited feedback network, a quantized channel measurement is reported to the BS. As discussed in [62], there is mainly two approaches for quantizing the CSI: (i) quantizing the properties of the transmitted signal (e.g. modulation, beam-forming vector) or (ii) quantizing the channel (i.e. adapt the transmitted signal to the channel). In order to overcome this limitation, previous works (e.g. [63], [64] and [65]) have proposed several strategies for reducing the number of resources consumed by the CSI feedback. Authors in [66] provide an overview of limited feedback in wireless communication systems which are particularly interesting for MIMO technology.

4.1.1 Related work

Researchers have been interested in elaborating new D2D resource allocation techniques for improving the performance of cellular network in terms of: throughput, interference management and energy saving etc. Most of the existing works assume the global CSI knowledge at the BS level and propose centralized D2D resource allocation algorithms. Several tools have been used for the study of centralized resource allocation problems: stochastic geometry modeling and resource optimization (e.g. [67]), optimization problem with objective of maximizing spatial reuse, centralized graph-theoretic approach (e.g. [68] and [69]), mixed-integer programming (e.g. [70], [71] and [72]), particle swarm optimization (e.g.[73]), non-convex optimization problem using branch-and-bound method (e.g.[74]) and coupled processors approach (e.g.[53]) etc.

In addition, several works limit the amount of CSI overhead by considering a partial CSI knowledge of the D2D-enabled cellular network: (i) in [75], authors consider that the BS has the global CSI knowledge except the interference links between UEs and (ii) in [68] the CSI knowledge is restricted to the cellular links. Authors of [76] consider statistical (and not instantaneous) CSI and propose a power allocation scheme for D2D-underlay cellular systems based on monotonic optimization. A stochastic cutting plane algorithm was proposed in [77] to achieve the cross-layer resource optimization without the knowledge of the channels' statistics.

The main challenge of centralized approaches is the need for the D2D channel state information at the BS level which suffers from a trade-off between the large amount of overhead (i.e. especially in scenarios where the channels vary rapidly with time) and the imperfect knowledge of the channels' states. Therefore, the full CSI knowledge assumption is not practical and pushes for performing distributed approaches for resource allocation of D2D communications.

Assuming the knowledge of the utility function at the D2D users' level, game theory has been the main tool used for elaborating distributed resource allocation: pricing (e.g. [78]), auctions (e.g. [26]) and coalition formation (e.g. [79]) etc. Game-theoretical approaches do not solve the overhead problem because users still need to share information (i.e. prices or bids etc). In addition, several works have evaluated the performance of both centralized and distributed approaches for D2D resource allocation. Authors in [80] proved that their distributed algorithm achieves

interesting performance gain with significant reduction of signaling overhead. In [68], both centralized and distributed resource allocation strategies were proposed for an underlay D2D communication system.

Energy efficiency consists a pertinent and important part of 5G networks at different levels: ecological side, customer satisfaction and mobile network operators' expenses. Hence, energy efficiency has been one of the main performance criteria that scheduling algorithms aim to optimize. Centralized (e.g. [57] and [81]) as well as distributed (e.g [82] and [83]) resource allocation algorithms were proposed for improving the energy efficiency of D2D-enabled cellular networks.

4.1.2 Contribution and Organization

The centralized characteristic of today's scheduling in cellular networks suffers from the ignorance of the global CSI knowledge of the network. Thus, the scheduling will always be limited by the number of resources available for CSI reporting. This weakness will be multiplied by the use of D2D technique where the D2D channels are estimated at the D2D receiver level and then reported to the BS. From that comes the idea of having a distributed scheduling that benefits for the D2D users' knowledge of their local CSI. Please note that the proposed distributed algorithm can be generalized to different performance metrics. In this work, we consider the example of energy efficiency which represents one of the relevant metrics to take into account during the scheduling of 5G networks.

Overlay D2D, i.e. dedicated resources for D2D communications, is assumed in order to avoid interference. In this work, we propose both centralized and distributed scheduling algorithm that optimizes the energy consumption of overlay-D2D networks under throughput constraints.[1] This optimization problem is studied based on Lyapunov technique. Lyapunov functions for general non-linear systems (i.e. especially stability analysis) is considered as a robust theoretical and practical tool. We start dealing with this Lyapunov optimization problem by proposing a centralized approach where the D2D resource allocation is managed by the central entity, i.e. BS. Based on channels' statistics, the BS chooses a subset of D2D pairs that will send their CSI feedback to the BS. Only the corresponding subset of CSIs is then received at the BS and the BS schedules then the optimal D2D link based on this subset of CSIs knowledge. We show that the performance of the proposed centralized algorithm achieves that of the optimal centralized scheduling in a limited feedback network.

In an ideal scenario, without limitation on the resources available for feedback transmission, centralized solution is the optimal one since the BS can acquire the instantaneous CSI knowledge of all the D2D pairs. However, in realistic context of limited feedback scenario, the proposed centralized scheduling suffers from the limited resources available for feedback transmission where only a subset of D2D pairs will be able to send its CSI to the BS. Since this subset is selected based on the

[1] This is only an example and does not limit the application of our algorithm to any other D2D performance metrics.

statistics of channel states, there will be no guarantee that the optimal D2D link will be scheduled. In this work, we show that in limited feedback networks, distributed solutions may take advantage of the local CSI knowledge of the D2D pairs to achieve higher performance.

Therefore, we propose a distributed algorithm that benefits from the users' knowledge of their local D2D channel state to intelligently manage the spectrum access and optimize the energy efficiency of D2D communications. Indeed, each user transmits a simple control indicator to reveal the value of its local D2D channel state. Based on these indicators, all the users, including the one that optimizes the energy efficiency of D2D communications, manage to report information concerning their local CSI. Therefore, this distributed approach that manages to identify the optimal D2D link, tends to achieve the performance of the *ideal* scenario (i.e. where the BS has the global CSI knowledge). However, a collision may occur while sharing these CSI indicators. The impact of this collision is discussed and some strategies for reducing its probability are proposed. Under these conditions of collision reduction, the performance of the distributed scheduling is very close to the *ideal* scheduling (i.e. where the BS knows the instantaneous CSI of all the D2D links without any cost).

We show how both centralized and distributed algorithms can be simply implemented in real cellular network, i.e. LTE. Numerical results reveal how the suggested algorithms improve the energy efficiency of D2D communications. In a limited feedback D2D networks, the proposed centralized algorithm schedules D2D communications based on the CSI statistics whereas the proposed distributed algorithm benefits from the users' knowledge of their instantaneous local CSI. Indeed, the distributed algorithm outperforms the centralized one due to the fact that, in a limited feedback D2D network, D2D users have more information about their local CSI than the BS.

This chapter will be organized as follows. In Section 4.2, the system model and the optimization problem are described. The centralized approach is detailed in Section 4.3: the algorithm is exposed and its optimistic property is proved. The distributed algorithm is detailed in section 4.4 and its performance is analyzed. Section 4.5 computes the probability of collision that may occur during the transmission of the CSI indicators. Strategies for reducing this probability are discussed. Section 4.6 reveals how the proposed algorithms can be implemented in today's cellular networks. Numerical results in Section 4.7 show the performance of the proposed algorithms and illustrate the significant energy efficiency gain provided by the distributed algorithm compared to other scheduling policies. Section 4.8 concludes the chapter whereas the proofs are provided in the appendices 7.4.

4.2 System model

We consider a set of N pairs of users that want to communicate with each other via D2D links (see figure 4.1 for $N = 6$). Users are randomly distributed in a cell of radius R_c such that the distance between D2D pair is set within the range $[d_{min}, d_{max}]$.

We denote by time-slot the time scale of the resource allocation decision. The channel between any two nodes in the network is modeled as a Rayleigh fading channel that remains constant during one time-slot and changes independently from one time-slot to another based on a complex Gaussian distribution with zero mean and unit variance.

A practical scenario with adaptive modulation is assumed. At each time slot, a device can support a bit-rate adapted to its channel conditions (i.e. SNR) and selected from a set of M bit rates $[\mathsf{R}_1, ..., \mathsf{R}_\mathsf{M}]$. These bit-rates correspond to a set of M SNR thresholds $[\mathsf{S}_1, ..., \mathsf{S}_\mathsf{M}]$. When a D2D link has a SNR that lies within the interval $[\mathsf{S}_\mathsf{m}, \mathsf{S}_{\mathsf{m}+1}[$ then its quantized value of SNR is equal to S_m and it will support a bit rate of R_m. When the n^{th} D2D link has a SNR of S_m then it can achieve a bit-rate equal to R_m by transmitting at a power $P_{n,m}$:

$$P_{n,m} := \min \left\{ \frac{\mathsf{S}_\mathsf{m} N_o}{|h_n|^2 L_n} \ , \ P_{max} \right\} \tag{4.1}$$

with h_n the fading coefficient of the n^{th} D2D pair, L_n the path-loss over the n^{th} D2D link that mainly depends on the n^{th} D2D pair distance d_n, N_o the noise power and P_{max} the maximum user's transmission power.

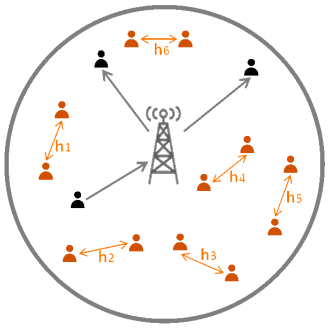


Figure 4.1: An example of D2D scenario where $N = 6$

We consider overlay D2D (e.g. see [84]) where dedicated resources are allocated to D2D links in order to mitigate interference between D2D and cellular communications (i.e. no reuse of cellular resources). Scheduling scheme is the algorithm that manages the radio resources' access of D2D communications. Assuming a user time division multiplex access scheduling, the available D2D resources are only used by one D2D communication at a given time-slot. We assume that, based on some pilot reference signals, each D2D transmitter has a channel state estimation of its corresponding D2D link. Deploying an energy efficient scheduling requires the CSI knowledge of these D2D links. Therefore, based on existing control channels, we propose that D2D links transmit a smart indicator of their local CSI information in order to enable energy efficient scheduling of D2D communications. N_{RB} represents the number of RB available each time-slot for D2D CSI reporting. We propose a new CSI reporting technique that benefits efficiently from these limited resources in order to reduce the energy consumption of the D2D communications while guaranteeing throughput constraints.

In the sequel, we use the following notations for a given time-slot t: $R_n(t)$ as the service rate of the n^{th} D2D communication, $P_n(t)$ as the transmission power of the n^{th} D2D communication and R_{th} as the threshold of the time average throughput of D2D communications with γ_{th} the corresponding SNR threshold.

The aim of this work is to design an energy efficient scheduling in a limited feedback network where D2D transmitters are limited by N_{RB} RBs for CSI exchange. We define $\Gamma(t)$ as the scheduled user at time-slot t. Hence, the Optimization Problem (OP) 4.2 consists of finding the scheduling strategy $\Gamma(t)$ that minimizes the energy consumption of D2D communications under throughput constraint and limited resources for CSI feedback transmission:

$$\begin{aligned} \underset{\Gamma}{\text{minimize}} \quad & \lim_{T\to\infty} \sup \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{\mathrm{N}} \mathbb{E}\left[P_n(t)\right] \\ \text{s.t.} \quad & \lim_{T\to\infty} \inf \frac{1}{T} \sum_{t=1}^{T} \mathbb{E}\left[R_n(t)\right] \geq R_{th} \quad \forall n, \\ & N_{RB}\text{: number of RB for CSI feedback.} \end{aligned} \tag{4.2}$$

We apply Lyapunov Optimization (see [31]) to solve the problem above. Looking at problem's constraint leads to the construction of the following virtual queues that help to meet the desired constraint:

$$Q_n(t+1) = \left[Q_n(t) - R_n(t)\right]^+ + R_{th} \tag{4.3}$$

From queuing theory [31], we known that the throughput constraint of the problem 4.2 is equivalent to the strong stability of the virtual queues 4.3. The optimization problem is transformed to a stabilization problem of the virtual queuing network while minimizing the time average of the users' transmission power. Hence, the drift-plus-penalty algorithm is used for minimizing the average power subject to network stability. When the system shifts to undesirable states, the defined Lyapunov function largely increases. Thus, scheduling actions that drift this function to the negative direction are crucial for settling the system stability. The scheduling policy Γ aims to minimize the following expression:

$$\begin{aligned} \underset{\Gamma}{\text{minimize}} \quad & \sum_{n=1}^{\mathrm{N}} V\mathbb{E}\left[P_n(t)\right] - Q_n(t)\,\mathbb{E}\left[R_n(t)\right] \\ \text{s.t.} \quad & N_{RB}\text{: number of RB for CSI feedback} \end{aligned} \tag{4.4}$$

where $Q_n(t+1) = \left[Q_n(t) - R_n(t)\right]^+ + R_{th}$ and V is a non-negative weight that is chosen in such a way that the desired performance trade-off between the power minimization and the virtual queue size is achieved. The scheduling solution of (4.4) aims to achieve a time average of the users' power consumption within a distance of at most $O\left(\frac{1}{V}\right)$ from the optimal value while ensuring a time average virtual queue backlog of $O(V)$.

4.3 Centralized approach

For the centralized approach, we suppose that the user's CSI feedback contains the following information concerning its D2D link: (i) channel quality (i.e. transmission rate) as well as (ii) transmission power. Due to the limited amount of resources available for CSI reporting, it is not possible for all the users to transmit their CSI feedback each time-slot neither to transmit the exact continuous values of their CSI. Therefore, depending on the number of resource blocks N_{RB} available for CSI reporting, a limited number of users is able to simultaneously transmit its quantized CSI feedback to the BS. The number of quantized CSI feedback (e.g. 20-22 encoded bits) that can be simultaneously supported at a given time-slot t is denote by $K^{(1)}(N_{RB})$ which depends on the number N_{RB} of resources available for CSI reporting. For clarity, we omit the variable (N_{RB}) when $K^{(1)}$ notation is used.

The centralized approach is based on the following three phases algorithm: **Phase 1** where the BS chooses, based on global statistical CSI, the subset Λ^* of users that will transmit their CSI feedback to the BS (i.e. with $|\Lambda^*| \leq K^{(1)}$); **Phase 2** where the BS receives the CSI feedback of the users within the subset Λ^* and **Phase 3** where the BS schedules the user, within the subset Λ^*, that optimizes the energy efficiency of the D2D communications based on the OP (4.4). We denote by Ω the set of all the possible subset of $K^{(1)}$ different users.

4.3.1 Centralized Algorithm

Moreover, the centralized algorithm (given by algorithm (4) for a time-slot t) will be detailed in this subsection.

Phase 1: The goal of this phase is to choose the subset Λ^* of $K^{(1)}$ D2D transmitters that will send their CSI to the BS at a given time-slot. Based on the global knowledge of the statistical CSI of D2D links, the BS computes the optimal subset Λ^* given by equation (4.5).

$$\Lambda^* := \underset{\Lambda \subset \Omega}{\operatorname{argmin}} \mathbb{E}_h \left[\min_{n \in \Lambda} \left[V P_n(t,h) - Q_n(t) R_n(t,h) \right] \right] \tag{4.5}$$

Phase 2: Each transmitter n of the subset Λ^* will proceed as follows: (i) computes the index $m^* \in \{1, .., \mathrm{M}\}$ that minimizes its utility function $V P_{n,m}(t) - Q_n(t) \mathsf{R_m}$, (ii) fixes respectively its bit rate and its transmission power as follows: $R_n(t) = \mathsf{R_{m^*}}$ and $P_n(t) = P_{n,m^*}(t)$ (iii) quantizes the transmission power $P_n(t)$ (i.e. denoted by $\tilde{P}_n(t)$) (iv) sends a CSI feedback that contains both: 1. the channel quality (i.e. which implies the chosen bit rate $R_n(t)$) and 2. the quantized transmission power $\tilde{P}_n(t)$.

Phase 3 Among the users in the subset Λ^*, the BS schedules the *optimal user* n^* which corresponds to the user that verifies equation (4.6):

$$n^* = \underset{n \in \Lambda^*}{\operatorname{argmin}} \left[V \tilde{P}_n(t) - Q_n(t) R_n(t) \right] \tag{4.6}$$

Algorithm 4 Centralized scheduling at the level of the BS

1: Finds subset Λ^* given by (4.5)
2: Sends CSI reporting request to users in subset Λ^*
3: Receives $R_n(t)$ and $\tilde{P}_n(t)$ from all users $n \in \Lambda^*$
4: Find the optimal user n^* given by (4.6)
5: Update $\boldsymbol{Q}$ based on (4.3)

4.3.2 Stability and optimistic criteria

We prove that the proposed centralized scheduling achieves a distance of at most $O\left(\frac{1}{V}\right)$ from the optimal solution of the centralized scenario while guaranteeing the stability of the system of virtual queues. We denote by P_c^* the optimal solution of the OP (4.2) when the best centralized scheduling is applied in a limited feedback network (i.e. scenario where the BS knows the global statistical but only the instantaneous CSI of a subset of $K^{(1)}$ users).

Proposition 2. *When the proposed centralized algorithm is applied, the total average backlogs of the queues is upper bounded by a finite value* $\frac{C+B}{\epsilon}$*:*

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[Q_i(t)\right] \leq \frac{C+B}{\epsilon} \tag{4.7}$$

The proposed centralized scheduling policy ensures the strong stability of the virtual queuing network with an average queue backlog of $O(V)$. Hence, the throughput constraint of the optimization problem (4.2) is satisfied.

Proposition 3. *For the centralized approach, the time average of power consumption verifies the following:*

$$P_c^* \leq \lim_{T\to\infty} \sup \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{\mathrm{N}} \mathbb{E}\left[P_n\left(\Gamma^{cent}(t)\right)\right] \leq P_c^* + \frac{C}{V} \tag{4.8}$$

Where C and V are finite and the value of V is tuned in such a way that the time average power is as close as possible to the solution of the optimal centralized limited-feedback scenario P_c^* with a corresponding queue size trade-off.

Proof. Proofs of propositions 2 and 3 are based on Lyapunov technique and are detailed in Appendix-7.4.2. □

We deduce that for a large finite value of V, the proposed centralized algorithm achieves the optimal solution of the centralized scenario which has P_c^* as the optimal time average power .

Algorithm 5 Distributed scheduling at the level of each user n

Receives, from BS, constants R_{th}, T_p and V (given by 4.18)
for $1 \leq t \leq T_p$ **do**
    Estimates channel state $h_n(t)$
    Computes performance metric $v_n(t)$ from (4.9)
    Computes v_{min} and v_{max} from (4.10) and (4.11)
    Finds $\{\tilde{v}_n, \tilde{k}_n\}$ given by (4.13) and (4.14)
    Shares CSI indicator at the $\tilde{k}_n^{th}$ RE
    if Collision **then**
        Detects collision index c
        Updates the values of r and f from (4.19)
    end if
    if $n == n^*$ given by (4.15) **then**
        Transmits data to its D2D pair
    end if
    Update $Q_n(t)$ from (4.3)
end for

4.4 Distributed approach

Finding the optimal solution of problem (4.4) requires the global CSI knowledge of D2D links. This knowledge is limited by the restricted number N_{RB} of resources available for CSI reporting. However, one can profit from the local CSI knowledge at the D2D users' level in order to propose a new way for handling the resources available for CSI feedback. This enables all the users to feedback some indicators concerning their CSI at each time-slot. After sharing CSI indicators, the user that optimizes (4.4) is identified. This study shows that the proposed distributed approach largely improves the energy efficiency of the D2D network. The proposed distributed energy efficient scheduling (given by Algo.5) can be summarized by the following three phases:

Phase 1: each D2D pair estimates its channel state in order to compute its energy efficiency metric. Then, each D2D user shares a simple CSI indicator (e.g. 1 or 2 encoded bits per time-slot) in such a way that the Resource Emplacement (RE) used for this CSI indicator transmission indicates the value of the energy efficiency metric.

Phase 2: The D2D user that has transmitted its CSI indicator at the RE of the lowest index and which corresponds to the user that maximizes the energy efficiency of D2D communications is scheduled. Thus, D2D link that optimizes the energy efficiency metric is chosen for data transmission.

Phase 3: This phase aims to reduce the collision that may occur during the transmission of CSI indicators.

4.4.1 Distributed algorithm

The different steps of the distributed algorithm (given in algorithm 5) are detailed in the sequel. The number of CSI indicators (e.g. 1 or 2 encoded bits) that can be simultaneously supported at a given time-slot t is denote by $K^{(2)}(N_{RB})$ which depends on the number N_{RB} of resources available for CSI reporting. For clarity, we omit the variable (N_{RB}) when $K^{(2)}$ notation is used.

Phase 1: Based on pilot reference signals, at each time-slot t, the n^{th} D2D pair estimates its D2D channel state $h_n(t)$ and deduces its energy efficiency metric given by equation (4.9).

$$v_n(t) = \min_{m \in \{1,\dots,\mathrm{M}\}} VP_{n,m}(t) - Q_n(t) R_m(t) \tag{4.9}$$

We can verify that the values of the utility function $v_n(t)$ fit within the range $[v_{min}(t), v_{max}(t)]$ given by equations (4.10) and (4.11) where $r = 1$ and $f = 0$.

$$v_{min}(t) = -tR_{th}\mathsf{R_M} \tag{4.10}$$

$$v_{max}(t) = v_{min}(t) + r\frac{VP_{max} - R_{th}\mathsf{R_1} - v_{min}(t)}{(K^{(2)})^f} \tag{4.11}$$

These border variables are identically computed by each device in such a way that all the devices will have the same values of $v_{min}(t)$ and $v_{max}(t)$. The interval $[v_{min}(t), v_{max}(t)]$ serves for the discretization of the utility function $v_n(t)$ into $K^{(2)}$ equal intervals. The values of $v_n(t)$ from the continuous set $[v_{min}(t), v_{max}(t)]$ are mapped to a finite set $\mathbb{S}_v$ of $K^{(2)}$ elements (see equation (4.12)). The simple way to quantize the utility function v_n is to choose the closest element to v_n within $\mathbb{S}_v$.

$$\mathbb{S}_v = \bigcup_{j=1,\dots,K^{(2)}} a_j = \bigcup_{j=1,\dots,K^{(2)}} \left\{ v_{min} + (j-1)\frac{v_{max} - v_{min}}{K^{(2)} - 1} \right\} \tag{4.12}$$

In practice, cellular networks contain limited resources for CSI reporting (i.e. here denoted by N_{RB} resource blocks). Thus, instead of transmitting the quantized CSI value, each D2D pair limits its feedback to a simple CSI indicator (e.g. 1 or 2 encoded bit) that is sufficient for describing its utility function. This mechanism, i.e. called ***Channel Indexing Feedback***, consists of introducing a mapping between the quantized value of the utility function (i.e. within the $\mathbb{S}_v$ set) and the $K^{(2)}$ REs available for the transmission of the CSI indicators. In other terms, the index of the REs used for the CSI indicator's transmission is sufficient for identifying the corresponding quantized value of the energy efficiency metric. Figure 4.2 illustrates an example where the proposed Channel Indexing Feedback technique is applied on 4 different users with $n = 1, 2, 3$ and 4. For the n^{th} D2D pair, the expressions of the quantized utility function (i.e. denoted by $\tilde{v}_n$) and its corresponding mapped RE used for the transmission of the CSI indicator (i.e. denoted by $\tilde{k}_n$) are formally given by:

$$\tilde{v}_n(t) = \underset{x_i \in \mathbb{S}_v}{\operatorname{argmin}} \left(v_n(t) - x_i\right) \mathbb{1}_{\{v_n(t) > x_i\}} \tag{4.13}$$

$$\tilde{k}_n(t) = \underset{i \in \left\{1,\ldots,K^{(2)}\right\}}{\operatorname{argmin}} \left(v_n(t) - x_i\right) \mathbb{1}_{\{v_n(t) > x_i\}} \tag{4.14}$$

where $\mathbb{1}$ represents the indicator function.

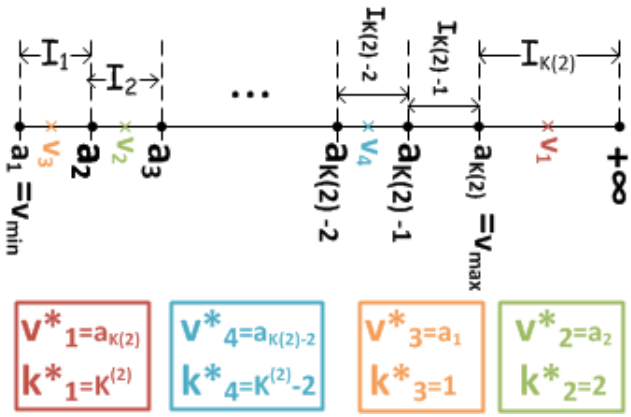


Figure 4.2: Example of channel indexing feedback for 4 users

In phase 1, at each time-slot t, each transmitter n computes its performance metric $v_n(t)$ given by (4.9) and then applies *Channel Indexing Feedback* technique to compute the couple $\left\{\tilde{v}_n(t), \tilde{k}_n(t)\right\}$ given respectively by (4.13) and (4.14) and which correspond to the quantized value of $v_n(t)$ and its corresponding index within the set $\mathbb{S}_v$. The transmitter n sends its CSI indicator at the $\tilde{k}_n^{th}$ RE among the $K^{(2)}$ available REs.

Phase 2: Depending on the shared CSI indicators by all the D2D pairs, the scheduled user is the one that exclusively transmits its CSI indicator on the RE with the lowest index $\tilde{k}_n(t)$. The chosen user corresponds to the D2D pair that minimizes the utility function $v_n(t)$ at time-slot t and by that maximizes the energy efficiency of D2D network. The index n^* of the scheduled D2D pair is formally given by:

$$n^* = \underset{n \in \{1,\ldots,N\}}{\operatorname{argmin}} \{\tilde{v}_n \mid \tilde{v}_m \neq \tilde{v}_n \ \ \forall\, m \in \{1,\ldots,N\}/n\} \tag{4.15}$$

The proposed algorithm is distributed in the sense that D2D users are responsible of computing their utility function as well as transmitting the corresponding CSI indicators based on their estimated CSI. The computation efforts are highly reduced at the BS level. However, recognizing the optimal D2D pair remains an open question since it can be done in a fully distributed manner or in a BS-assisted way. In the sequel, we discuss these two approaches:

- **BS-assistance:** BS is responsible of listening to the transmitted CSI indicators by all the D2D pairs in order to detect the optimal user whose CSI indicator has been transmitted at the RE of the lowest index. Then, the BS announces the identified optimal user as the scheduled user. Such network assistance encounters the business challenge of distributed algorithm for mobile network operators that generally prefer to support centralized administrated solutions for controlling the network and guaranteeing the performance of cellular communications.

- **Autonomous:** Supposing that devices have full duplex capacities, they can simultaneously transmit and receive the CSI indicators. Hence, each user will be able to autonomously recognize whether it corresponds to the optimal user that should be scheduled or not. If a D2D pair send its CSI indicator on RE of index c and does not receive any CSI indicators on the REs of index $< c$ then this D2D pair will recognize that it will be scheduled. This algorithm has important benefit for autonomous networks (i.e. without any centralized entity) but faces important security issues that need to be solved.

Phase 3: During the transmission of CSI indicators in phase 1, a collision may occur when at least two users transmit their CSI indicators at the same RE (i.e. at least two users have the same quantized performance metric $\tilde{v}_n$). Phase 3 consists of applying some strategies that reduce the occurrence probability of such collisions. These procedure are detailed in section 4.5.

4.4.2 Performance analysis

We denote by P_{id}^* the time average of the power when an *ideal* scheduling is considered in the sense that the network has a global CSI knowledge of the D2D communications. In this case, the *ideal* scheduling achieves the optimum of (4.2) without any constraints on the number of resources available for CSI feedback. When a collision free scenario is considered, we prove that the proposed distributed scheduling achieves a distance of at most $O\left(\frac{1}{V}\right)$ from the *ideal* solution while ensuring the strong stability of the virtual queuing network with an average queue backlog of $O(V)$.

Proposition 4. *Assuming that no collision occurs; the distributed scheduling* $\Gamma^{dist}(t)$ *guarantees a time average power consumption that verifies the following:*

$$P_{id}^* \leq \lim_{T\to\infty} \sup \frac{1}{T} \sum_{t=1}^{T} \sum_{n=1}^{\mathrm{N}} \mathbb{E}\left[P_n\left(\Gamma^{dist}(t)\right)\right] \leq P_{id}^* + \frac{C}{V} \tag{4.16}$$

Where C and V are finite and the value of V is tuned to make the time average power as close as desired to the ideal solution with a corresponding virtual queue size trade-off.

Proof. See Appendix-7.4.2 based on Lyapunov technique. □

We deduce that for a tuned finite value of V, the proposed distributed algorithm achieves the performance of the *ideal* solution P_{id}^{*} of the optimization problem (4.2). with an error of $O\left(\frac{1}{V}\right)$. For high finite values of V, the distributed scheduling reaches the *ideal* performance when no collision occurs at the transmission level of the CSI indicators. The process of reducing this collision are discussed in the coming section.

4.5 Probability of collision

During the phase 2 of the distributed algorithm, the transmission of the CSI indicators may suffer from a collision. It is crucial to note that this collision occurs at the level of the CSI indicators' transmission and not at the level of data transmission. A collision takes place when at least two users have the same quantized utility function $\tilde{v}_n$ and thus transmit their CSI indicators at the same RE. In particular, when a collision occurs at the level of the RE of index k^*, identifying the optimal user to schedule is not possible anymore. We define the overall collision as the scenario where each user collides at least with another one; hence none of the users is scheduled at this time-slot. In order to avoid such scenario, two precautions detailed in the sequel, were adopted:

- The Lyapunov constant V is chosen based on equation (4.18) in order to minimize the collision probability.
- The mapping, based on which the D2D users match their discrete values of the Energy Efficiency (EE) metric with the $K^{(2)}$ available RE, is updated as in equation (4.19) in the aim of avoiding future collisions.

4.5.1 Choosing Lyapunov constant

We describe how to limit the probability of feedback collision by choosing the appropriate value of the Lyapunov constant V. We consider that the value of the Lyapunov constant V is updated within a periodicity of T_p time-slots. We start by giving the analytic expression of the collision probability. A collision occurs at a given RE when at least two users transmit their CSI indicators at the same RE. We call **probability of collision** P_c as the probability of occurring an overall collision event where each CSI indicator transmission collides with at least another one in such a way that none of the users is scheduled (i.e. none of the users has exclusively transmitted its CSI indicator on one of the available REs). In this part, we limit the analytic result to the single bit-rate case $\mathrm{M} = 1$ (i.e. corresponding to bit-rate R and SNR S).

Proposition 5. *The probability of collision P_c is given by:*

$$P_c = 1 - \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{K^{(2)}} \bar{p}_c\{i,j\} \prod_{k=1}^{j-1} \left(1 - \sum_{l=1 \neq i}^{\mathrm{N}} \bar{p}_c\{l,k\}\right) \tag{4.17}$$

$$\textit{where } \bar{p}_c\{i,j\} = 2\left[\exp\left(c_{i,j-1}\right) - \exp\left(c_{i,j}\right)\right] \times \prod_{k=1\neq i}^{\mathrm{N}} \left[1 - 2\exp\left(c_{k,j-1}\right) + 2\exp\left(c_{k,j}\right)\right]$$

$$\textit{and } c_{i,j} = -\frac{V\mathsf{S}N_o}{\left(a_j + Q_i\mathsf{R}\right)L_i}$$

Proof. See Appendix-7.4.3. □

Based on the expression (4.17), we can tune the value of the Lyapunov constant V in order to limit the overall collision probability to a small given ϵ.

Theorem 4.5.1. *The probability P_c is bounded by a given ϵ (with $0 \leq \epsilon \leq 1$) when the value of the Lyapunov constant V is given by:*

$$V(\epsilon) = -\frac{R_{th}\mathsf{R}\ln(\epsilon')T_p}{P_{max}\ln(\epsilon') + \mathsf{S}N_0 L_{min}^{-1}} \tag{4.18}$$

where

$$\epsilon' := \frac{1}{2\mathrm{N}}\left[1 - \left(\frac{1-\epsilon}{\mathrm{N}K^{(2)}}\right)^{\frac{1}{\mathrm{N}+K^{(2)}}}\right]$$

and L_{min} is the path-loss over a D2D link with d_{min} as the smallest distance allowed between a D2D pair.

Proof. See Appendix-7.4.4. □

4.5.2 Updating the mapping

If a collision occurs then the granularity of the mapping set $\mathbb{S}_v$ is not sufficient for proposing different CSI indicators to describe the EE metric of the different D2D pairs. Hence, an accuracy improvement of the mapping set $\mathbb{S}_v$ is done in order to avoid future collisions. We denote by c the smallest RE's index where a collision has occurred (with $1 \leq c \leq K^{(2)}$). The mapping, described by the set $\mathbb{S}_v$, between the quantized value of the EE metric and the $K^{(2)}$ available REs is updated by modifying the parameters r and f of the v_{max} formula (4.11) as follows:

$$r = c \text{ ; if } \left(r < K^{(2)}\right) f = f + 1 \quad \text{else } \left(r == K^{(2)}\right) f = 0 \tag{4.19}$$

When collision occurs at the RE of index c with $c < K^{(2)}$ then the mapping update aims to reduce the probability of collision by taking the following actions: (i) $v_{max} = a_c$ to reduce the interval $[v_{min}, v_{max}]$ and (ii) $f = f + 1$ to increase the granularity of the intervals within the new subset $\mathbb{S}_v$. However, when collision occurs at the RE with the highest index ($c = K^{(2)}$), this means that all the D2D users have a utility function higher than the current v_{max}. Thus, the mapping update aims to reduce the probability of collision by enlarging the interval $[v_{min}, v_{max}]$ (with $r = K^{(2)}$ and $f = 0$).

4.6 Implementation

We show how the proposed centralized and distributed approaches can be implemented in real networks. To do so, we consider the example of LTE and beyond networks. We focus on how the existing PUCCH formats can be modified in order to support the proposed algorithms.

4.6.1 Existing feedback Standardization

In this work, we benefit from the existing DL feedback standards in LTE specifications developed by 3GPP (i.e. one can refer to [85] for more details). The PUCCH appears mainly in two formats depending on the type of the handled information: (i) the formats 1,1a,1b of 1-2 encoded bits and which are dedicated for the ACK/NACK feedback and (ii) the formats 2,2a,2b of 20-22 encoded bits and which are mainly used for CSI feedback. The CSI feedback consists of three component: (i) the Rank Index (RI), the Precoding Matrix Indicator (PMI) and the Channel Quality Index (CQI). The preferred triplet CSI (RI/PMI/CQI) is computed by each user based on its instantaneous channel estimations obtained from DL pilot. In the aim of reducing the complexity of such computation, several algorithms have been proposed in the literature (e.g. [86]).

As mentioned before, the BS may improve the performance of D2D communications by acquiring the CSI of D2D links. The reporting of these CSI can be performed in two different ways: (i) periodic CSI report (summarized version and economical in terms of radio resources) and (ii) aperiodic CSI report (detailed version and costly in terms of radio resources). By default, the periodic CSI reporting is done on the PUCCH. However, when the user's data are planned to be transmitted on the Physical Uplink Shared Channel (PUSCH) then the CSI reporting is multiplexed with the data and sent on the PUSCH. Meanwhile, the aperiodic CSI reporting is exclusively transmitted on the PUSCH after the reception of a BS request via a Physical Downlink Control Channel (PDCCH) that carries a DL Downlink Control Information (DCI) of format 0.

The periodic CSI feedback, sent on PUCCH resources, are configured based on a semi-static scheduling. This configuration is specifically assigned to each UE via radio resource control Radio Resource Control (RRC) in order to avoid the need of increasing the size of PDCCH. When semi-static scheduling is deployed, the BS pre-configures each user with a given resource allocation identifier and periodicity. The limited amount of PUCCH resources prohibits the transmission of all the users' PUCCH at each time-slot (i.e. refers to Time Transmission Interval (TTI) in LTE) and obliges the possible CSI reporting to include only the necessary information and not the detailed one. This limitation motivates us to propose a new management of these critical PUCCH resources that enables an energy efficient scheduling.

The performance of DL communications depends on the precision of the CSI feedback which is function of the frequency granularity. However, the more accurate

the CSI feedback is the more the UL feedback load is important. Different CSI reporting mode on PUCCH resources are defined as function of the trade-off existing between the DL performance and the UL load: (i) mode 1: wideband CQI report where a single CQI value corresponds to the entire system bandwidth and (ii) mode 2: subband CQI report where the system bandwidth is divided into multiple subbands with different CQI value for each subband. In this work, we limit the implementation section to the case of wideband CQI report. Please note that considering the case of subband CSI reporting is a straightforward process (i.e. in this case, channel indexing feedback technique is applied respectively for each subband CSI by the use of the resources blocks available for this subband feedback).

Cellular networks are designed based on a centralized approach where the BS presents the entity that controls the operations and guarantees the quality of cellular communications. The BS will start by configuring the sub-frames corresponding to the CSI feedback via the identification of the bandwidth of the PUCCH region, the period of feedback and the cyclic shifting (i.e. permitting the time multiplexing between the CSI reports of different UEs and/or between the CSI reports of the same UE). Depending on these parameters, each user transmits periodically its CSI feedback. However, due to the restricted number of PUCCH resources, only a subset of users, i.e. not the totality of the users, will transmit their CSI feedback at a given TTI. Depending on the received CSI feedback, the BS runs its scheduling algorithm in order to allocate D2D resources. This is the baseline protocol to which we will compare the proposed scheduling algorithms.

4.6.2 Centralized algorithm implementation

Recall that we consider user scheduling in such a way that only the optimal user is scheduled at a given TTI to transmit its data overall the available D2D resources. For the centralized approach, we suppose that the users send their CSI by the use of PUCCH format 2b. The control information contained in this PUCCH format (CQI and 2-bits for ACK or NACK) will be modified as follows: CQI will remain intact however the 2-bits of ACK-NACK will be used for indicating the D2D users' transmission. Thus, users' transmission powers will be quantized by mapping their continuous values to a countable small set of 4 elements: $\left\{\tilde{P}_1, \tilde{P}_2, \tilde{P}_3, \tilde{P}_4\right\}$. Based on these modifications, the three phases centralized algorithm (detailed in algorithm 4) can be implemented.

When the PUCCH format 2b is adopted for CSI feedback, then based on [87] we can deduce the number $K^{(1)}$ of CSI feedback that can be simultaneously supported. As shown in figure 4.3, $K^{(1)}$ is equal to the product of the two following identifiers (i.e. communicated by the BS via RRC):

- N_{RB} that indicates the number of RBs that are available per TTI for CSI feedback of a PUCCH 2/2a/2b format. It can be a configurable parameter such that the BS can control the UL bandwidth and can eventually dimension the size of these resources depending on the need.

- C_{MUX} that indicates the multiplexing capacity per RB. This corresponds to the number of users that can send their CSI feedback on the same RB. Giving that this parameter depends only on the cyclic shifting of the base sequences (which is a fix value in this case) then $C_{MUX} = 12$.

We deduce that $K^{(1)} = N_{RB} \times C_{MUX} = 12N_{RB}$.

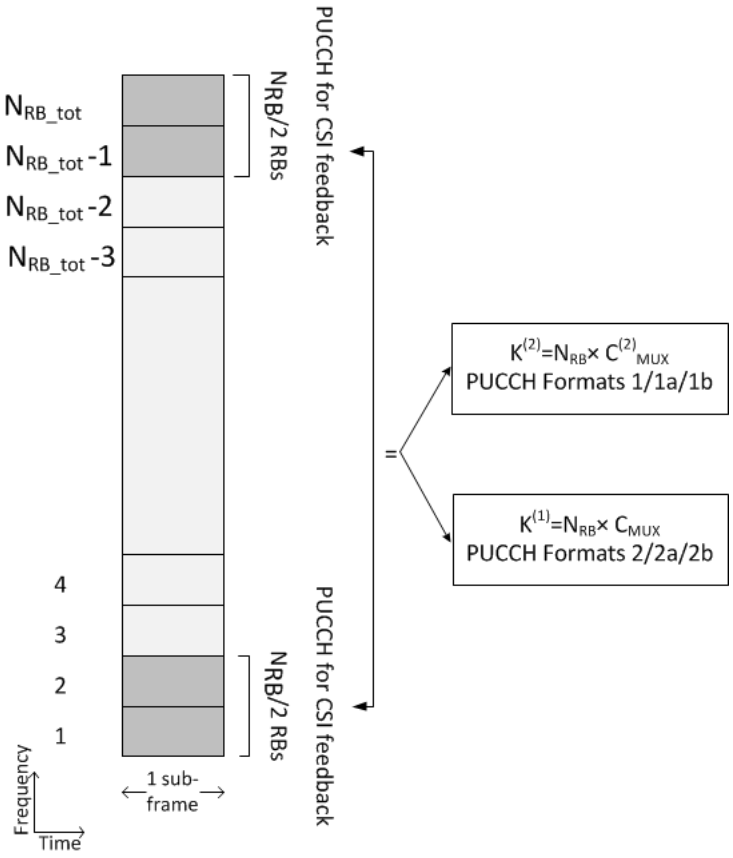


Figure 4.3: The resource blocks allocated for users' feedback (PUCCH Format 1 and 2)

4.6.3 Distributed algorithm implementation

In the distributed approach, the resources available for CSI feedback are handled in a new distinct way in order to guarantee that the optimal user, at each TTI, sends its CSI indicator. Based on existing standards, we consider that the CSI indicators consists of PUCCH format 1/1a/1b which are commonly used for ACK/NACK messages or scheduling request. These control information correspond to 1 or 2 encoded bits per TTI. Indeed, users transmit their CSI indicators as PUCCH format 1/1a/1b on the RE that corresponds to their quantized EE metric.

Based on [87], we can deduce the number $K^{(2)}$ of PUCCH format 1/1a/1b that can be simultaneously supported at a given time-slot and thus can be used for the *Channel Indexing Feedback* technique. As shown in figure 4.3, $K^{(2)}$ is equal to the product of the two following identifiers communicated by the BS via RRC (i.e. $K^{(2)} = N_{RB} \times C^{(2)}_{MUX}$):

- N_{RB}: indicates the number of RBs available per TTI for the transmission of PUCCH format 1/1a/1b. It can be a configurable parameter such that the BS can control the UL bandwidth and eventually dimension the size of these resources depending on the need.

- $C_{MUX}^{(2)}$ as the multiplexing capacity per RB, which means the number of PUCCH format 1/1a/1b that can be transmitted on the same RB. This parameter depends on both parameters: (i) the number of possible orthogonal codes N_{OC} (e.g. three for normal cyclic prefix and two for extended cyclic prefix) and (ii) the difference Δ_{shift}^{PUCCH} between two consecutive cyclic shifting for resources using the same orthogonal code. Then, $C_{MUX}^{(2)} = 12N_{OC}/\Delta_{shift}^{PUCCH}$.

Using PUCCH format 1/1a/1b, the procedure of the distributed scheduling is described as is follows (see Algo.5). The BS initiates the scheduling by announcing the constants used by the users to compute their EE metric (e.g. the Lyapunov constant V, update duration T_p, throughput threshold R_{th}). At each time-slot t, each user n computes the couple $P_n(t)$ and $R_n(t)$ that minimizes its utility function $v_n(t)$ based on its instantaneous estimated D2D channel $h_n(t)$. Applying equations (4.10) and (4.11), the users can locally compute the values of $v_{min}(t)$ and $v_{max}(t)$ in order to deduce the mapping set $\mathbb{S}_v$. Each D2D user deduces the quantized value of its utility function $\tilde{v}_n$ from equation (4.13) as well as its corresponding RE index $\tilde{k}_n$ from equation (4.14). Each D2D user n sends a CSI indicator, as a PUCCH format 1b, on the $\tilde{k}_n^{th}$ RE. Considering BS-assistance approach of the distributed algorithm, BS decides to schedule the user that exclusively transmits its CSI indicator at the RE of the lowest index. This is equivalent to scheduling the user that minimizes the utility function $v_n(t)$ at time-slot t and maximizes by that the EE of D2D network. If a feedback collision occurs at some level of the REs, then the BS transmits the smallest index c of REs where collision has occurred. The broadcasting of these scheduling information can be done via PDCCH that carries a DCI of formats 1A or 1C. Depending on the collision index received from the BS, users update their parameters r and f as in equation (4.19) in order to increase the granularity of the intervals within the set $\mathbb{S}_v$ and reduce by that the probability of future collision.

4.7 Numerical Results

4.7.1 Numerical settings

We summarize the numerical settings of this section in the table 4.1. The performance are evaluated by averaging over 100 different UEs realizations. We estimate the value of the Lyapunov constant V based on equation (4.18) and we find $V = 10^{15}$ for $\epsilon = 0.1$, $\mathsf{S} = 80$ dB, $\mathsf{R} = 700$ kbps/RB and $R_{th} = 500$ kbps/RB for $\gamma_{th} = 14$dB.

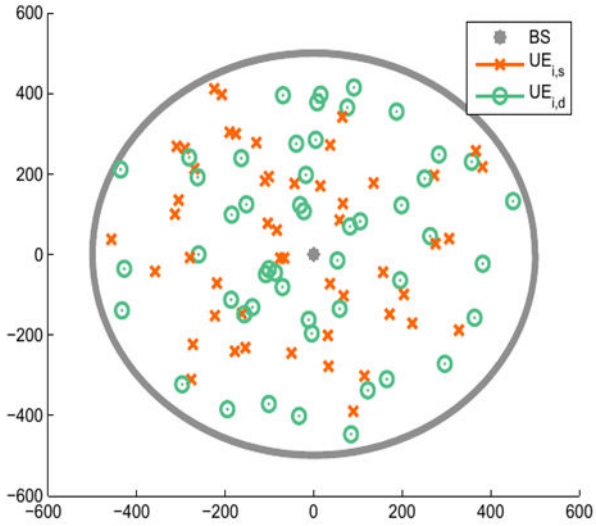


Figure 4.4: Uniform Random Localization of $N = 50$ pairs of D2D

In LTE, there are 15 different values for CQI (i.e. mapping between CQI and modulation etc...). Hence, we suppose the existence of 15 different bit-rates that could be applied for link adaptation model (i.e. M = 15). From an internal link-level simulator we deduce a throughput-SNR mapping for a 10 MHz Evolved Universal Mobile Telecommunications System Terrestrial Radio Access, (E-UTRA) TDD network. This mapping gives practical values of the bit-rates $\{\mathsf{R}_1, ..., \mathsf{R}_{15}\}$ as well as their corresponding SNR values $\{\mathsf{S}_1, ..., \mathsf{S}_{15}\} = \{0, 1, ..., 14\}$ dB.

Parameter	**Value**
Cell Radius R_c	500m
Bandwidth	10MHz equivalent to 50 RB
UE drop	$N = 50$ UEs Random Uniform drop with $d_{min} = 3$ m , $d_{max} = 350$ m from [18] and [61] e.g. Fig. 4.4
Feedback Parameters	$N_{RB} = 2$, $\Delta_{shift}^{PUCCH} = 1$, $N_{OC} = 3$ $\Rightarrow K^{(1)} = 24$ and $K^{(2)} = 72$
P_{max}	250 mW
Quantized Powers	$\tilde{P}_1 = 50$, $\tilde{P}_2 = 100$ mW $\tilde{P}_3 = 150$, $\tilde{P}_4 = 200$ mW
Path-loss	outdoor-to-outdoor path-loss in Channel Models section of [18]
Simu. Settings	100 realizations of $T_p = 10^6$ ms each
Noise density	-174 dBm/Hz
D2D Noise Figure	9 dB

Table 4.1: Numerical Settings of Chapter 4

4.7.2 Numerical Evaluation

N D2D pairs are uniformly distributed in a cell of radius R_c. The scheduling scheme determines how these D2D communications access the D2D resources in the network. The performance of the proposed algorithms is evaluated by comparing the time average of the users' energy consumption (Energy Consumption (EC)) as well as their EE) between the following different algorithms:

- **Centralized-limited feedback** scheduling: proposed in section 4.3.
- **Distributed** scheduling: proposed in section 4.4.
- **Ideal** scheduling: BS has the global knowledge of the instantaneous channel states of all the D2D links.
- **Round-Robbin** scheduling: each subset Λ of users is chosen in equal portions of time and in a circular order for the transmission of their CSI feedback using PUCCH format 2/2a/2b. The number of users that can send simultaneously their CSI feedback depends on the number of resource blocks available for feedback transmission. For $N_{RB} = 2$ resources available for feedback and a multiplexing capacity $C^{(1)}_{MUX} = 12$ of PUCCH format 2/2a/2b per RB; the number of users in each subset Λ is $|\Lambda| = N_{RB} \times C^{(1)}_{MUX} = 2 \times 12 = 24$.

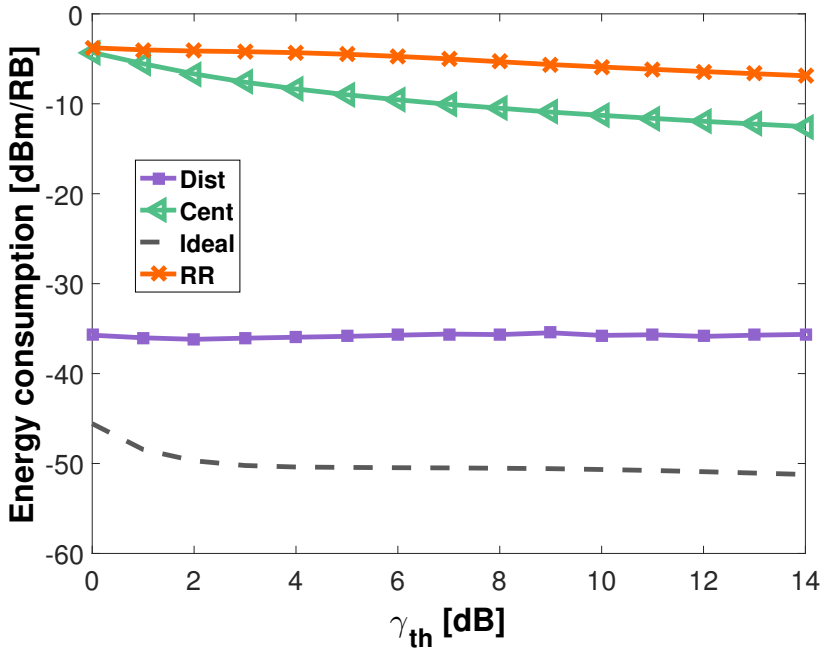


Figure 4.5: EC as function of the SNR threshold γ_{th}

Fig. 4.5 shows how the proposed distributed and centralized approaches reduce the time average of users' transmitted power compared to the Round-Robbin approach

for different SNR thresholds $0 \leq \gamma_{th} \leq 14$ dB. The reduction is up to 70% for the centralized approach and up to 98% for the distributed approach. The proposed distributed algorithm outperforms the centralized one that suffers from a limited number of users that simultaneously transmit their CSI feedback to the BS, hence scheduling the optimal user is not guaranteed. The distributed algorithm proposes a new way to manage the limited resources available for feedback transmission. Thus, all the D2D users benefit from their local CSI knowledge and limit their feedback transmission to a small CSI indicator. This new channel indexing feedback technique guarantees the scheduling of the optimal user. Nevertheless, the distributed scheme does not achieve the ideal one as collision may occur. Even though a collision probability of 0.1 occurs, the distributed algorithm highly reduces the users' transmission power.

In the aim of verifying that the throughput constraint is ensured while minimizing the users' transmission power, we study the EE of the proposed algorithms. The EE metric is defined as the ratio of the total throughput to the total transmitted power over all the simulation duration (see [88]). Figure 4.6 represents the evolution of the EE of the proposed algorithms as function of the SNR threshold γ_{th}. These results show an important enhancement of the network EE and underline the performance of the distributed algorithm compared to the other non *ideal* scheduling.

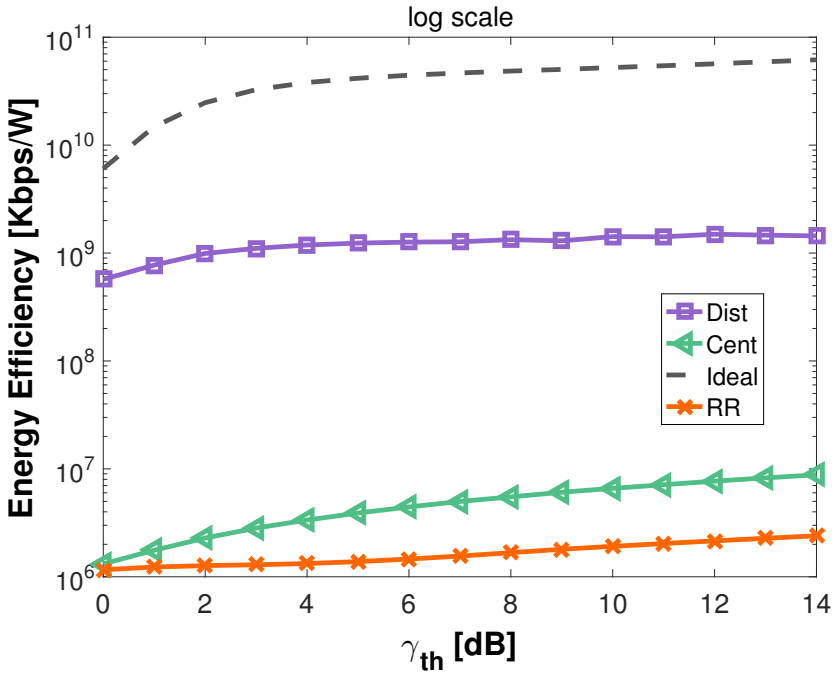


Figure 4.6: EE as function of the SNR threshold γ_{th}

4.8 Conclusion

In this chapter, we have proposed a scheduling scheme that minimizes D2D transmission power under throughput constraints in a limited feedback cellular systems. Lyapunov technique is used for solving this optimization problem. Both centralized and distributed algorithms were proposed based on new CSI reporting. The main idea of the centralized approach is to integrate information concerning the user's transmission power in their CSI reporting message. However, centralized approach suffers from the restricted number of resources available for CSI reporting. Hence, we avoid this limitation by considering a distributed resource allocation scheme that benefits from the users' knowledge of their local CSI and proposes a new way for reporting CSI. The distributed algorithm is based on the idea that all the D2D users compute their EE metric and then share a simple CSI indicator (e.g. PUCCH format 1 for LTE) that is sufficient for describing the value of this EE metric. In this way, we guarantee the scheduling of the D2D pair that optimizes the EE of the network. However, collision may occur during the transmission of CSI indicators; thus collision reduction mechanisms are developed. Furthermore, we give a detailed description on how the centralized and distributed algorithms can be simply implemented in practice, i.e. in a LTE cellular network. Numerical results show an energy consumption reduction between 70% to 98% and a high EE enhancement compared to a Round-Robbin CSI reporting. These results illustrate that the distributed algorithm outperforms the centralized one.

5 | Mobile D2D Relay Selection

In D2D enabled cellular networks, user to network relaying can be handled for improving the performance of cellular networks. Under realistic assumption that relays are in mobility, it is crucial to define a strategy for designating the relays that will respectively serve each Master User Equipment (MU) in the network. In this chapter, we propose a relay selection policy that maximizes the performance of cellular networks (e.g. throughput, reliability, coverage) under cost constraints (e.g. transmission power, power budget). We assume that the relays' dynamics are represented by a Markov Decision Process (MDP) where the MU cannot directly observe the locations of all the potential relays that have been discovered. Therefore, the sequential relay decision process is modeled by a Constrained Partially Observed Markov Decision Process (CPOMDP). Since the exact solutions of such framework are computationally intractable to find, we have developed an approximated solution as well as discussed the existing trade-off between its complexity and its preciseness. Moreover, proving the submodularity property of the reward and cost functions leads us to propose a greedy form of the approximated solution. Numerical results are presented to endorse our relay selection policies and to show how introducing D2D relaying can highly improve the performance of cellular networks. Furthermore, a system-level simulator is developed in order to implement our strategy of relay selection and to test its performance in a nearly realistic cellular network.

5.1 Concept and related work

In traditional cellular networks, single hop communications are deployed between the users and the BS. However, introducing relays to cellular networks has become one of the major concern of cellular network planners that aim to improve the capacity and the coverage of their networks. The emergence of D2D communications encourages the deployment of UE-to-Network relaying functionality. D2D communications will take place between the MU and the relays however cellular communications are maintained between the relays and the BS. The main advantages of such relaying feature is: (i) improving the performance of the network (e.g. capacity enhancement, coverage extension, transmission power reduction, load balancing, network offloading etc.) and enabling new services (e.g. data on demand).

D2D relaying technique has been the subject of interest of significant research in both academia and industry. Several tools have been used for evaluating the

performance of this technique. Stochastic geometry is used for analytically modeling and analyzing the performance of cellular network with fix D2D relays in [89] and mobile D2D relays in [90]. Monte-Carlo simulations in [91] and [92] show how enabling D2D communications to carry relayed traffic can enhance the capacity and coverage of cellular networks. A system-level simulator in [93] was developed to evaluate the extension of the cellular coverage due to D2D relaying.

Despite the performance gain that UE-to-Network relaying has promised, several challenging issues require further investigation. One can ask how the relays should be strategically positioned in order to optimize the performance of the network.Sharing the spectrum between D2D and cellular communciations is one of the existing challenges. Using stochastic geometry modeling, authors in [94] studied underlay D2D enabled cellular networks. They analytically derived the tradeoff generated by the spectrum partition between D2D and cellular communications and they deduce the optimal spectrum partition that guarantees the fairness in the network. Several solutions of power and/or resource allocation in D2D relayed cellular networks have been studied in the recent literature. The work [95], that proposes a distributed resource allocation for D2D relay-aided cellular networks using game theory tool, exposes a summary of the different existing centralized and distributed resource allocation schemes. In addition to resource and power allocation, a mode and path selection algorithm was developed and simulated in [96]. Furthermore, Constrained Markov Decision Process (CMDP) problems were formulated in [97] and [98] to obtain the optimal decision of packet scheduling that mobile relays should take.

One of the main challenges of D2D relay-aided cellular networks is to decide how relays should be selected in order to achieve the performance of cooperative relaying. Due to relay mobility, a dynamic relay selection policy is unavoidable. Indeed, a relay selected at one position can be no longer helpful at another position. In this work, we address this question by proposing a dynamic relay selection strategy that MU may take for optimizing a certain performance metric (e.g. throughput, coverage, reliability etc.). Since the relaying functionality is costly for the relays (e.g. in terms of energy, data consumption etc), we propose to charge MUs for using their selected relays. This cost aims to encourage the users to behave as potential relays. Based on the fact that the reward as well as the cost of each relay depends on its location, we propose a dynamic relay selection policy that maximizes some QoS metric of cellular networks while satisfying some cost constraints.

5.1.1 Related Work

Relay selection is crucial for improving the performance of D2D relay-aided cellular networks. The rich literature on relay selection problems can be divided into two categories: mobile relay selection and fix relay selection schemes. Most of the previous works considers fix relay selection scenarios. Some of these schemes are briefly presented in the following. Based on a stochastic formulation, authors in [99] propose a fully distributed single relay association scheme that aims to increase the spectral efficiency of the network. An energy efficient relay selection was proposed in [100]

based on a Discrete Time Markov Chain (DTMC) modeling of the relay node with Discontinuous Reception (DRX) mechanism. Based on an iterative technique, [101] proposes a joint relay selection and power allocation problem scheme for relay-aided D2D underlying cellular networks. The work in [102] uses a queuing theory model to propose a single relay selection scheme that optimizes the network in terms of relay remaining battery life, end-to-end data rate and end-to-end delay criteria. For underlay D2D enabled cellular networks, interference is mitigated between cellular and D2Dcommunications by considering a distributed relay selection algorithm in [28].

However, the aforementioned works do not take user mobility into account but consider fix relays which limits their applicability in cellular mobile networks. Considering the mobility of the relays seems to be a challenging scenario. Mobile relays have been the subject of the work in [103]. Based on the knowledge of the relay mobility pattern, a dynamic relay selection scheme aiming to minimize the cost of relaying under QoS requirements was proposed. The studied optimization model is based on CMDP. Since relay selection appears as decision making process, several MDP based formulation has been studied in order to propose optimal relay selection scheme. For example, while predicting the channel states of the available relays, authors in [104] address a relay selection policy that maximizes the long term transmission rate. This decision strategy is obtained by solving a POMDP with dynamic programming-based algorithm.

5.1.2 Contribution and Organization

The use cases of D2D communications that will be studied in this chapter is user to network relaying in cellular networks. The mobility of the relays is the main challenge for relay selection decision. We propose a dynamic relay selection policy that maximizes a certain performance metric of the network (e.g. throughput, reliability, coverage etc.) under cost constraints (e.g. energy consumption, data consumption etc.). The main contribution in this work can be summarized as follows:

- CPOMDP formulation of the sequential relay selection process that a MU will decide and discussion concerning the complexity of this problem.
- Proposition of a dynamic policy of relay selection that approximately optimizes the formulated problem based on a greedy point based value iteration.
- Extending the results to multiple users scenario.
- Numerical results as well as system-level simulations show the performance that cellular networks may gain by implementing the proposed relay selection policy.

The main particularity of this work compared to other relay selection schemes previously proposed in the literature is the following:

- The consideration of a realistic scenario where relays are in mobility, thus dynamic relay selection scheme is necessary.
- The cellular network performance is optimized under cost constraints.
- The sequential relay decision process is modeled by a Constrained Partially Observed Markov Decision Process (CPOMDP) (and not a CMDP) since the states of the potential relays cannot be observed until these relays are selected by the MU.
- An approximated solution is proposed for avoiding the intractability of exact solutions. A trade-off between the preciseness of the approximated solution and its complexity is derived.
- The relay selection is not limited to one relay, thus the performance gain is improved due to the increasing in the cooperative diversity order. Actually, the proposed relay selection policy aims to maximize the performance of the network under cost constraints without any constraints on the number of chosen relays.
- The relays' mobility pattern is assumed communicated to the agent decision maker (MU or the BS).

The rest of this chapter is organized as follows. Section 5.2 describes the system model for a single MU scenario. Section 5.3 formulates the optimization problem as a CPOMDP. Since the exact solutions of such problem are intractable to find, a low-complexity dynamic relay selection, called CPBVI, is proposed in section 5.4. The submodularity property is verified for this problem, thus a greedy form of this approximation, called GCPBVI, is deduced. These results are extended to a multi-MU scenario in section 5.5. For this scenario, a distributed approach is exposed for reducing the complexity of centralized solutions. Numerical results in section 5.6 corroborate our claims. The performance enhancement of the cellular networks is shown in section 5.7 by implementing the proposed relay selection scheme in a system-level simulator. Section 5.8 concludes the paper whereas the proofs are provided in the appendices 7.5.

5.2 System Model

For the sake of clarity, we describe the system model for a single MU scenario. We show that extending the result to multi-MU scenario is a straight forward process.

5.2.1 Network model

We consider a single cell scenario with one MU and a set of K potential mobile relay. MU is allowed to use D2D communications to access the network via mobile relays and by that improving the performance of its cellular communications. MU discovers nearby potential relays by launching a discovery process of periodicity T. The time between two discovery processes is partitioned into decision epochs t (e.g. TTI in LTE networks) of constant duration (with $t \in \{1, 2, ..., T\}$). The goal of this work is to determine the relay selection policy that should be applied by the MU at each epoch t in order to maximize its cumulative reward under cost constraints.

$\mathcal{K} = \{0, 1, 2, ..., K\}$ denotes the set of K existing potential relays (from index 1 to K) as well as the direct cellular link (index 0). The relay selection policy consists in deciding whether the MU will have direct communication with the network or will pass by some mobile relays. In the latter case, this policy chooses the subset of relays that will be used for attaining the network. Please note that D2D relaying can be applied to both DL and UL communications. Due to the practical consideration that mobile terminals do not support simultaneous signal transmission and reception, a two-phases transmission scheme is assumed. Therefore, for UL (resp. DL) relayed communication, the transmission protocol is divided into two phases: (i) in the first phase the relay receives the data from the MU (resp. BS) and (ii) in the second phase the relay transmits the received data to the BS (resp. MU).

5.2.2 Mobility Model

Relays' locations in a service coverage area are quantized and represented by a set of regions $\mathcal{S} = \{S_1, S_2, ..., S_{|\mathcal{S}|}\}$. We assume that the relays remain in the same region during a decision epoch. We denote by $s_i(t)$ the location of relay i at epoch t. In the next epoch, the location of each relay is changed (i.e. by either staying in the same region or moving to another neighboring region). Thus, the mobility of relay i is modeled by a transition matrix $\boldsymbol{P_i}$ where each element $P_i(S_n, S_{n'})$ of this matrix denotes the probability that relay i moves from region $S_n \in \mathcal{S}$ to region $S_{n'} \in \mathcal{S}$ in the next decision epoch. In this work, we assume that the MU is aware of the mobility pattern of each discovered relay (i.e. equivalent to the relay's transition matrix). The vector $\boldsymbol{s}_t = (s_1(t), s_2(t), ..., s_K(t))$ denotes the location of the K potential relays at epoch t.

We underline that the location of each relay is changing over time epochs based on its transition probability matrix and independently from the fact that this relay

has been or not selected. We assume that the MU will be able to know the current localization region of only the chosen relays. On the other hand, the MU will just have a certain belief of the locations of the non selected relays. It is a practical assumption because when a relay is selected by the MU some signaling is exchanged between these two equipments in order to allow D2D communications. This exchange may include some information that indicates the region of the selected relay. However, when the relay is not selected then no need for any signaling between the MU and the relay since no communication will take place between these two equipments. In this case, recognizing the region of the relays requires an overhead of signaling which can be avoided.

5.2.3 Cost and Reward Model

We propose a relay selection policy that aims to optimize a cellular reward under cost constraints. Here, cost and reward models are described. The reward function, which depends on the relay's location, represents the benefit of choosing a relay in terms of throughput and/or reliability and/or coverage etc. The cost function, which depends on the relay's location, defines the charge that the MU owes to each selected relay (i.e. in terms of energy and/or incentive budget etc.). We respectively denote by $r_i(s_i)$ and $c_i(s_i)$ the reward and the cost of the i^{th} relay in location $s_i \in \mathcal{S}$. We assume that when the MU chooses a set of relays then the total reward (resp. cost) is the sum of the reward (resp. cost) of each relay in the chosen set. We denote by $\boldsymbol{a} = (a_1, a_2, ..., a_K)$ the vector such that $a_i = 1$ if relay i is selected and 0 otherwise. Thus, the total reward and cost at a given state $\boldsymbol{s}$ are given by:

$$R(\boldsymbol{s}, \boldsymbol{a}) = \sum_{i \in \mathcal{K}} r_i(s_i) \, \mathbb{1}_{\{a_i = 1\}} \tag{5.1}$$

$$C(\boldsymbol{s}, \boldsymbol{a}) = \sum_{i \in \mathcal{K}} c_i(s_i) \, \mathbb{1}_{\{a_i = 1\}} \tag{5.2}$$

5.3 Problem formulation

Enabling user to network relaying functionality based on D2D communications leads to the following question: which mobile relays should be chosen by the MU for ensuring an enhancement in cellular communications. This decision depends essentially on the reward and cost parameters of the discovered relays. The main challenge of such relay selection policy remains in the mobility of these potential relays. The goal is to make the relay selection decision $(\boldsymbol{a}_0, ..., \boldsymbol{a}_T)$ that optimize the following:

$$\max \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t R(\boldsymbol{s}_t, \boldsymbol{a}_t)\right] \text{ s.t. } \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t C(\boldsymbol{s}_t, \boldsymbol{a}_t)\right] \leq C_{th} \tag{5.3}$$

where $\gamma \in [0, 1]$ is a discount factor that represents the difference in importance between future rewards (resp. costs) and present rewards (resp. costs).

The considered reward model includes a large scope of reward metrics. In the following we give few examples that indicate how the reward mode 5.1 can be applied:

- Throughput criteria: different packets are transmitted to the selected relays, thus the total throughput is the sum of the throughput of each selected link. For example, in the case of considering Shannon capacity over a bandwidth of WHz, one can write the reward for each selected relay i at state s_i as follows:

$$r_i\left(s_i\right) = W \log_2\left[1 + \min\left\{\mathrm{SINR}_{\mathrm{MU-Relay}_i}\left(s_i\right), \mathrm{SINR}_{\mathrm{Relay}_i-\mathrm{BS}}\left(s_i\right)\right\}\right]$$

  Therefore, the total reward is given by:

$$R\left(\boldsymbol{s}, \boldsymbol{a}\right) = W \sum_{i \in \mathcal{K} | a_i = 1} W \log_2\left[1 + \min\left\{\mathrm{SINR}_{\mathrm{MU-Relay}_i}\left(s_i\right), \mathrm{SINR}_{\mathrm{Relay}_i-\mathrm{BS}}\left(s_i\right)\right\}\right]$$

- Reliability criteria: same packets are transmitted to the selected relays, thus the error probability q is equal to the product of the error probability of each selected link q_i. For example, in the case of considering the bit error ratio in the case of Quadrature Phase-Shift Keying (QPSK) modulation and Additive White Gaussian Noise (AWGN) channel, the error probability of relay i at state s_i can be written as follows:

$$q_i\left(s_i\right) = \frac{1}{2} erfc\left(\sqrt{\mathrm{SINR}_{\mathrm{MU-Relay}_i}\left(s_i\right)}\right) \times \frac{1}{2} erfc\left(\sqrt{\mathrm{SINR}_{\mathrm{Relay}_i-\mathrm{BS}}\left(s_i\right)}\right)$$

  When action $\boldsymbol{a}$ is taken, then $q\left(\boldsymbol{a}, \boldsymbol{s}\right) = \prod_{i \in \mathcal{K} | a_i = 1} q_i\left(s_i\right)$. Thus, minimizing this error probability corresponds to maximizing the $-\log$ of the probability error of each selected link

$$\min_{\boldsymbol{a}} q\left(\boldsymbol{a}, \boldsymbol{s}\right) = \min_{\boldsymbol{a}} \prod_{i \in \mathcal{K} | a_i = 1} q_i\left(s_i\right) = \max_{\boldsymbol{a}} - \sum_{i \in \mathcal{K} | a_i = 1} \log\left[q_i\left(s_i\right)\right]$$

  Therefore, we can express the reliability reward as $-\log$ of the error probability in order to have the overall reward as the sum of the reward of each selected relay:

$$R\left(\boldsymbol{s}, \boldsymbol{a}\right) = - \sum_{i \in \mathcal{K} | a_i = 1} \log\left[\frac{1}{4} erfc\left(\sqrt{\mathrm{SINR}_{\mathrm{MU-Relay}_i}\left(s_i\right)}\right) erfc\left(\sqrt{\mathrm{SINR}_{\mathrm{Relay}_i-\mathrm{BS}}\left(s_i\right)}\right)\right]$$

  The process above can be applied to any other error probability function (e.g. [105]).

- Coverage probability criteria: same packets are transmitted to the selected relays, thus the overall outage probability is equal to the product of the outage probability of each selected link. Similarly to reliability criteria, we can express the coverage reward as $-\log$ of the outage probability in order to have the overall reward as the sum of the reward of each selected relay.

The considered cost model can be applied to a wide range of cost metrics. In the following, we give some examples to illustrate how the cost model given by 5.2 can be applied:

- Energy criteria: the total consumed energy is the sum of the energy consumed by each selected relay. For example, if we denote by $P_i(s_i)$ the transmission power of relay i when it is in state s_i, thus:

$$C(\boldsymbol{s}, \boldsymbol{a}) = \sum_{i \in \mathcal{K} | a_i = 1} P_i(s_i)$$

- Incentive criteria: the total charged cost is equal to the sum of the incentive budget required by each selected relay. For example, one can denote by $L_i(s_i)$ the tokens used by relay i in state s_i, thus:

$$C(\boldsymbol{s}, \boldsymbol{a}) = \sum_{i \in \mathcal{K} | a_i = 1} L_i(s_i)$$

5.3.1 Restless Markov Multi-armed Bandit (RMAB) representation

The MU relay selection procedure consists of choosing a limited set of mobile relays that optimizes its subjective function under cost constraints. However, the positions of the potential relays are partially known at the epoch of decision and turn to be more observable as time passes. This problem can be modeled by a RMAB due to the three following reasons:

- It is a multi-armed bandit problem because the set of potential relays represents the arms among which the MU will make its selection.
- It is a Markov multi-armed bandit problem because the location of the i^{th} arm is modeled by a discrete, irreducible and aperiodic Markov chain with finite space $\mathcal{S}$ and transition probability matrix $\boldsymbol{P}_i$ (i.e. relays' locations are the states of this RMAB)
- It is restless because the potential relays change their location/state from a decision epoch to another independently from the current decision (whether they've been selected or not).

The objective of the relay selection strategy is to optimize problem 5.3 through a sequence of selection $1 \leq t \leq T$. For each selection decision, there exists a trade-off between exploitation phase that tends to achieve the highest expected reward and exploration phase that tends to get more information concerning the positions of other relays. In the sequel, we describe the CPOMDP formulation equivalent to the proposed RMAB problem.

5.3.2 CPOMDP formulation

In order to study the RMAB defined above, we formulate its equivalent finite-horizon CPOMDP which is characterized by the tuple $\langle \mathcal{S}, \mathcal{A}, T(), \mathcal{Z}, O(), R, C, \boldsymbol{b}_0, T, C_{th}, \gamma \rangle$ defined below:

- **State** $\boldsymbol{s} = (s_1, s_2, ..., s_K)$ denotes the state vector or the location vector of the K potential relays (where $s_i \in \mathcal{S}$ for all $1 \leq i \leq K$). $\boldsymbol{\mathcal{S}}$ represents the set of all possible state vectors with $|\boldsymbol{\mathcal{S}}| = |\mathcal{S}|^K$.

- **Action** $\boldsymbol{a} = (a_1, a_2, ..., a_K) \in \{0,1\}^K$ denotes the vector of the K binary actions such that $a_i \in \{0,1\}$ specifies whether relay i is selected ($a_i = 1$) or not ($a_i = 0$) for all $1 \leq i \leq K$. $\hat{a} = \{i : a_i = 1\}$ represents the set of the indexes of the selected relays. $\mathcal{A}$ denotes the set of all action vectors, it contains $|\mathcal{A}| = \sum_{i=1}^{K} \binom{K}{i} = 2^K$ elements.

- **Transition function** $T(\boldsymbol{s}, \boldsymbol{s}') : \boldsymbol{\mathcal{S}} \times \boldsymbol{\mathcal{S}} \rightarrow [0,1]$ represents the probability of transiting between states. $T(\boldsymbol{s}, \boldsymbol{s}')$ characterizes the probability of passing to state $\boldsymbol{s}'$ in the next decision epoch knowing that the current state is $\boldsymbol{s}$. Assuming that the relays' mobility are independent and described by their transition probability matrix, then $T(\boldsymbol{s}, \boldsymbol{s}') = \prod_{i=1}^{K} P_i(s_i, s_i')$. This probability is independent from the action $\boldsymbol{a}$ since the selection decision has a purely observational role (i.e. the relays change their locations independently from the fact that they have been selected or not). $\boldsymbol{T}$ represents the transition matrix of $|\boldsymbol{\mathcal{S}}| \times |\boldsymbol{\mathcal{S}}|$ elements.

- **Observation** $\boldsymbol{z} = (z_1, z_2, ..., z_K)$ denotes the observation vector of the K potential relays. Selecting a relays leads to the observation of its state, hence when relay i is not selected then $z_i = \emptyset$ otherwise $z_i = s_i$. The set of all observations is denoted by $\mathcal{Z}$.

- **Conditional observation probability** $O(\boldsymbol{z}', \boldsymbol{s}', \boldsymbol{a}), : \mathcal{Z} \times \boldsymbol{\mathcal{S}} \times \mathcal{A} \rightarrow [0,1]$ represents the probability of receiving an observation $\boldsymbol{z}' \in Z$ knowing that the decision policy takes action $\boldsymbol{a} \in \mathcal{A}$ and by that transits to state $\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}$, we define:

$$O\left(\boldsymbol{z}', \boldsymbol{s}', \boldsymbol{a}\right) = O\left(\boldsymbol{z}' | \boldsymbol{s}', \boldsymbol{a}\right) = Pr\left(\boldsymbol{z}_{t+1} = \boldsymbol{z}' | \boldsymbol{s}_{t+1} = \boldsymbol{s}', \boldsymbol{a}_t = \boldsymbol{a}\right)$$

- **Reward** $R(\boldsymbol{s}, \boldsymbol{a})$ is the reward achieved by taking action $\boldsymbol{a}$ when the K potential relays are in state $\boldsymbol{s}$. We suppose $R_{min} \leq R(\boldsymbol{s}, \boldsymbol{a}) \leq R_{max}$ for all $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$ and $\boldsymbol{a} \in \mathcal{A}$.

- **Cost** $C(\boldsymbol{s}, \boldsymbol{a})$ is the cost charged by taking action $\boldsymbol{a}$ when the K potential relays are in state $\boldsymbol{s}$. We suppose $C_{min} \leq C(\boldsymbol{s}, \boldsymbol{a}) \leq C_{max}$ for all $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$ and $\boldsymbol{a} \in \mathcal{A}$.

- **Initial belief** $\boldsymbol{b}_0$ is a vector of $\boldsymbol{\mathcal{S}}$ elements that denotes the initial distribution probability of being at each state $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$.
- **Horizon** T of the CPOMDP represents the number of epochs of the relay selection policy.
- C_{th} the cost threshold.
- $\gamma \in [0, 1]$ as a discount factor.

The MU chooses its action as function of the history of observations and actions that have been executed in the past. The result of [33] demonstrates that using the belief states for defining the optimal policy provides as much information as using the entire history of actions taken and observations received. Indeed, there is no need to explicitly save this history but having the current belief state is sufficient for deciding the upcoming actions.

Therefore, a POMDP can be represented as a belief MDP where every belief is a state. Since there is infinite belief states, this MDP is defined over a continuous set of belief states. At a given epoch t, the MU gathers all the information concerning the past decisions in the belief state defined as follows:

$$b_t(\boldsymbol{s}) := P(\boldsymbol{s}_t = \boldsymbol{s} | \boldsymbol{z}_t, \boldsymbol{a}_{t-1}, \boldsymbol{z}_{t-1}, ..., \boldsymbol{a}_0) \, \forall \boldsymbol{s} \in \boldsymbol{\mathcal{S}} \tag{5.4}$$

The belief vector of $|\boldsymbol{\mathcal{S}}|$ elements is denoted by $\boldsymbol{b}_t = \left(b_t(\boldsymbol{s}_1), b_t(\boldsymbol{s}_2), ..., b_t\left(\boldsymbol{s}_{|\boldsymbol{\mathcal{S}}|}\right)\right)$. By analogy, we define $b_t^i(s) := P(s_i(t) = s | \boldsymbol{z}_t, \boldsymbol{a}_{t-1}, \boldsymbol{z}_{t-1}, ..., \boldsymbol{a}_0)$ as the belief state of the mobile relay i and its corresponding belief vector:

$$\boldsymbol{b}_t^i = \left(b_t^i(S_1), b_t^i(S_2), ..., b_t^i\left(S_{|\mathcal{S}|}\right)\right)$$

Given the fact that the mobile relays move independently, the belief state $b^t(\boldsymbol{s})$, at a given state $\boldsymbol{s} = (s_1, s_2, ..., s_K) \in \boldsymbol{\mathcal{S}}$ and epoch t, can be deduced from the belief state of each relay as follows:

$$b_t(\boldsymbol{s}) = b_t(s_1, s_2, ..., s_K) = \prod_{i \in \mathcal{K}} b_t^i(s_i)$$

The belief state $\boldsymbol{b}_t^i$ of relay i at epoch t is recursively computed based on the previous belief state $\boldsymbol{b}_{t-1}^i$, previous action $\boldsymbol{a}_{t-1}$ and current observation $\boldsymbol{z}_t = (z_1(t), z_2(t), ..., z_K(t))$:

$$\boldsymbol{b}_t^i = \begin{cases} \boldsymbol{b}_{t-1}^i \boldsymbol{P}_i & \text{if } a_i(t-1) = 0 \\ \boldsymbol{P}_i(z_i(t), :) & \text{if } a_i(t-1) = 1 \end{cases} \tag{5.5}$$

where $\boldsymbol{P}_i(S_j, :) = \left(P_i(S_j, S_1), P_i(S_j, S_2), ..., P_i\left(S_j, S_{|\mathcal{S}|}\right)\right)$. Therefore, when the action $\boldsymbol{a}$ (with equivalent set $\hat{a}$) is taken and the observation $\boldsymbol{z}$ is detected then the belief of being in state $\boldsymbol{s} = (s_1, s_2, ..., s_K) \in \boldsymbol{\mathcal{S}}$ is updated as follows:

$$b_{\boldsymbol{z}}^{\boldsymbol{a}}(\boldsymbol{s}) = \prod_{i \in \hat{a}_t} P_i(z_i, s_i) \prod_{j \in \mathcal{K} | \hat{a}_t} \sum_{s'_j \in \mathcal{S}} b_j\left(s'_j\right) P_j\left(s'_j, s_j\right) \tag{5.6}$$

The belief states pursue a Markov Process where the belief state at a given epoch depends on the belief, action and observation of the previous epoch with the following transition function:

$$\tau\left(\boldsymbol{b},\boldsymbol{a},\boldsymbol{b}'\right)=\sum_{\boldsymbol{z}\in Z}Pr\left(\boldsymbol{b}'|\boldsymbol{b},\boldsymbol{a},\boldsymbol{z}\right)Pr\left(\boldsymbol{z}|\boldsymbol{b},\boldsymbol{a}\right) \tag{5.7}$$

$$\text{with } Pr\left(\boldsymbol{z}|\boldsymbol{b},\boldsymbol{a}\right)=\sum_{\boldsymbol{s}'\in\mathcal{S}}O\left(\boldsymbol{s}',\boldsymbol{a},\boldsymbol{z}\right)\sum_{\boldsymbol{s}\in\mathcal{S}}T\left(\boldsymbol{s},\boldsymbol{s}'\right)b\left(\boldsymbol{s}\right) \tag{5.8}$$

$$\text{and } P\left(\boldsymbol{b}'|\boldsymbol{b},\boldsymbol{a},\boldsymbol{z}\right)=\begin{cases}1 & \text{iff } \boldsymbol{b}'=\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}\\ 0 & \text{otherwise}\end{cases}$$

A policy $\pi\left(\boldsymbol{b}\right):\boldsymbol{b}\rightarrow\boldsymbol{a}$ is a function that determines the action $\boldsymbol{a}$ to take at each belief state $\boldsymbol{b}$. For a given initial policy $\boldsymbol{b}_0$, a policy π is characterized by a value function $V^{r,\pi}\left(\boldsymbol{b}\right)$ for the reward evaluation and $V^{c,\pi}\left(\boldsymbol{b}\right)$ for the cost evaluation:

$$V^{r,\pi}\left(\boldsymbol{b}_0\right)=\sum_{t=1}^{T}\gamma^t\rho_r\left(\boldsymbol{b}_t,\pi\left(\boldsymbol{b}_t\right)\right)=\mathbb{E}\left[\sum_{t=1}^{T}\gamma^t R\left(\boldsymbol{s}_t,\boldsymbol{a}_t\right)|\boldsymbol{b}_0,\pi\right] \tag{5.9}$$

$$V^{c,\pi}\left(\boldsymbol{b}_0\right)=\sum_{t=1}^{T}\gamma^t\rho_c\left(\boldsymbol{b}_t,\pi\left(\boldsymbol{b}_t\right)\right)=\mathbb{E}\left[\sum_{t=1}^{T}\gamma^t C\left(\boldsymbol{s}_t,\boldsymbol{a}_t\right)|\boldsymbol{b}_0,\pi\right] \tag{5.10}$$

where the belief-based reward $\rho_r\left(\boldsymbol{b},\boldsymbol{a}\right)$ and the belief-based cost $\rho_c\left(\boldsymbol{b},\boldsymbol{a}\right)$ are:

$$\rho_r(\boldsymbol{b},\boldsymbol{a})=\sum_{\boldsymbol{s}\in\mathcal{S}}b(\boldsymbol{s})R(\boldsymbol{s},\boldsymbol{a}) \tag{5.11}$$

$$\rho_c(\boldsymbol{b},\boldsymbol{a})=\sum_{\boldsymbol{s}\in\mathcal{S}}b(\boldsymbol{s})C(\boldsymbol{s},\boldsymbol{a}) \tag{5.12}$$

The value functions $V_t^{r,\pi}$ and $V_t^{c,\pi}$ at epoch t and under policy π are given by the *Bellman* equation:

$$V_t^{r,\pi}\left(\boldsymbol{b}\right)=\rho_r\left(\boldsymbol{b},\boldsymbol{a}_\pi\right)+\gamma\sum_{\boldsymbol{z}\in Z}Pr\left(\boldsymbol{z}|\boldsymbol{a}_\pi,\boldsymbol{b}\right)V_{t+1}^{r,\pi}\left(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}_\pi}\right) \tag{5.13}$$

$$V_t^{c,\pi}\left(\boldsymbol{b}\right)=\rho_c\left(\boldsymbol{b},\boldsymbol{a}_\pi\right)+\gamma\sum_{\boldsymbol{z}\in Z}Pr\left(\boldsymbol{z}|\boldsymbol{a}_\pi,\boldsymbol{b}\right)V_{t+1}^{c,\pi}\left(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}_\pi}\right) \tag{5.14}$$

where $Pr\left(\boldsymbol{z}|\boldsymbol{a}_\pi,\boldsymbol{b}\right)$ is given by equation (5.8).

The action-value function $Q_t^{r,\pi}\left(\boldsymbol{b},\boldsymbol{a}\right)$ is the reward of taking action $\boldsymbol{a}$ at epoch t and following policy π thereafter:

$$Q_t^{r,\pi}\left(\boldsymbol{b},\boldsymbol{a}\right)=\rho_r\left(\boldsymbol{b},\boldsymbol{a}\right)+\gamma\sum_{\boldsymbol{z}\in Z}Pr\left(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}\right)V_{t+1}^{r,\pi}\left(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}\right) \tag{5.15}$$

The action-value function $Q_t^{c,\pi}\left(\boldsymbol{b},\boldsymbol{a}\right)$ is the cost of taking action $\boldsymbol{a}$ at epoch t and following policy π thereafter:

$$Q_t^{c,\pi}\left(\boldsymbol{b},\boldsymbol{a}\right)=\rho_c\left(\boldsymbol{b},\boldsymbol{a}\right)+\gamma\sum_{\boldsymbol{z}\in Z}Pr\left(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}\right)V_{t+1}^{c,\pi}\left(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}\right) \tag{5.16}$$

The objective of the optimal relay selection policy is to achieve the highest reward under cost constraints. The optimal policy π^* maximizes the value function $V^{r,\pi}$ under constraints on $V^{c,\pi}$. There exists an optimal Markov policy which defines the action to be executed for each belief state supposing that the upcoming actions will be chosen in an optimal manner. The value of the optimal policy π^* function $V_t^*(\boldsymbol{b})$ satisfies the Bellman optimal equation:

$$V_t^{r,*}(\boldsymbol{b}) = \max_{\boldsymbol{a}} Q_t^{r,\pi}(\boldsymbol{b},\boldsymbol{a}) = \max_{\boldsymbol{a}} \left[\rho_r(\boldsymbol{b},\boldsymbol{a}) + \gamma \sum_{\boldsymbol{z}\in Z} Pr(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}) V_{t+1}^{r,*}(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}) \right] \tag{5.17}$$

s.t.

$$V_t^{c,*}(\boldsymbol{b}) = \left[\rho_c(\boldsymbol{b},\boldsymbol{a}) + \gamma \sum_{\boldsymbol{z}\in Z} Pr(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}) V_{t+1}^{c,*}(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}) \right] \leq C_{th}$$

We define $\mathcal{L}$ the Bellman dynamic operator: $V_t^*(\boldsymbol{b}) = \mathcal{L}V_{t+1}^*(\boldsymbol{b})$.

5.4 Relay Selection Policies

5.4.1 Exact Solution of CPOMDP

[33] shows that the value function of POMDP is a PWLC function over the infinite belief simplex (denoted by Δ). Thus, the value function can be characterized by a finite set of value function vectors (hyperplans) [37]. Each vector represents the optimal value of the value function in a given region of the belief space. The exact solution of the CPOMDP, proposed in [106], is inspired from the value iteration of POMDP. This technique presents the value functions $V_t^{r,\pi}$ and $V_t^{c,\pi}$ as PWLC functions over the infinite belief simplex Δ. Therefore, the reward and cost value functions $V_t^{r,\pi}$ and $V_t^{c,\pi}$ are represented as a finite set of α-vectors pairs (α_r^i, α_c^i).

Considering $\mathcal{V}_t$ the set of α-vectors pairs at epoch t, then the set $\mathcal{V}_{t+1}$ at the following epoch is constructed by applying the following dynamic programming operator $\mathcal{L}$ over all the action-observation pairs: [1]

$$\Gamma_r^{\boldsymbol{a},*} \leftarrow \alpha_r^{\boldsymbol{a},*}(\boldsymbol{s}) = R(\boldsymbol{s},\boldsymbol{a}) \text{ and } \Gamma_c^{\boldsymbol{a},*} \leftarrow \alpha_c^{\boldsymbol{a},*}(\boldsymbol{s}) = C(\boldsymbol{s},\boldsymbol{a})$$

$$\Gamma_r^{\boldsymbol{a},\boldsymbol{z}} \leftarrow \alpha_{r,i}^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s}'\in\boldsymbol{\mathcal{S}}} T(\boldsymbol{s},\boldsymbol{s}') O(\boldsymbol{z},\boldsymbol{s}',\boldsymbol{a}) \alpha_r^i(\boldsymbol{s}')$$

$$\Gamma_c^{\boldsymbol{a},\boldsymbol{z}} \leftarrow \alpha_{c,i}^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s}'\in\boldsymbol{\mathcal{S}}} T(\boldsymbol{s},\boldsymbol{s}') O(\boldsymbol{z},\boldsymbol{s}',\boldsymbol{a}) \alpha_c^i(\boldsymbol{s}'), \forall \left(\alpha_r^i, \alpha_c^i\right) \in \mathcal{V}_t \tag{5.18}$$

$$\Gamma_r^{\boldsymbol{a}} = \Gamma_r^{\boldsymbol{a},*} + \oplus_{\boldsymbol{z}\in\mathcal{Z}} \Gamma_r^{\boldsymbol{a},\boldsymbol{z}} \text{ and } \Gamma_c^{\boldsymbol{a}} = \Gamma_c^{\boldsymbol{a},*} + \oplus_{\boldsymbol{z}\in\mathcal{Z}} \Gamma_c^{\boldsymbol{a},\boldsymbol{z}}$$

This dynamic programming update produces $\mathcal{V}_{t+1}$:

$$\mathcal{V}_{t+1} \leftarrow \mathcal{L}\mathcal{V}_t = \bigcup_{\boldsymbol{a}\in A} (\Gamma_c^{\boldsymbol{a}}, \Gamma_r^{\boldsymbol{a}})$$

[1] The cross sum operator $\oplus$ of two sets $\mathcal{A}$ and $\mathcal{B}$ is given by the set : $\mathcal{A}\oplus\mathcal{B} = \{a+b | a \in A, b \in B\}$.

At the worst case, each dynamic programming update will generate an exponentially increasing $|\mathcal{V}_{t+1}| = |\mathcal{A}||\mathcal{V}_t|^{|\mathcal{Z}|}$ pairs of α-vectors. However, some pairs of α-vectors are never the optimal one in any region of the belief simplex (i.e. called useless vectors). Therefore, for mitigating this exponential explosion, different pruning algorithms were developed in order to exclude these useless vectors. For POMDP problem, [37] have proposed an algorithm that identifies these useless vectors in order to ignore them. For CPOMDP a pruning operation for the value functions were proposed in [106] in order to generate the minimal set of pairs of α-vectors at each iteration. The pruning function *Prune* that we adopt for implementing this solution consists of keeping each pair $(\alpha_r^i, \alpha_c^i) \in \mathcal{V}_t$ that satisfies the cumulative cost constraint $\alpha_c^i.\boldsymbol{b} \leq C_{th}$ and that have the higher cumulative reward $\alpha_r^i.\boldsymbol{b}$ in some region of the belief simplex Δ. This can be decided by solving a Mixed-Integer Linear Program (MILP) for each pair of α-vector (one can refer to equation (3) in [107]). This pruning method eliminates each pair of vectors (α_r^i, α_c^i) that violates the cumulative cost constraint at a given iteration t. Note that this operation may lead to a suboptimal policy. Therefore, randomized policies are the subject of further study to achieve the optimal solution of CPOMDP.

As presented in algorithm 6, the iteration t of the exact solution of CPOMDP follows this procedure: (i) exact dynamic programming to generate the pairs α-vectors, (ii) pruning operation, based on a mixed integer linear program, that produces the minimal set of α-vectors, (iii) deducing the optimal value function.

Algorithm 6 Iteration of CPOMDP Exact Solution

Input: α-vector set $\mathcal{V}_t$, Actions $\mathcal{A}$, States $\boldsymbol{\mathcal{S}}$, Observations $\mathcal{Z}$, Reward function $R(\boldsymbol{s}, \boldsymbol{a})$, cost function $C(\boldsymbol{s}, \boldsymbol{a})$, Cost threshold C_{th}
for $\boldsymbol{a} \in \mathcal{A}$ **do**
  $(\alpha_r^{\boldsymbol{a},*}, \alpha_c^{\boldsymbol{a},*}) \leftarrow (R(., \boldsymbol{a}), C(., \boldsymbol{a}))$
  for $\boldsymbol{z} \in \mathcal{Z}$ **do**
    for $(\alpha_r^i, \alpha_c^i) \in \mathcal{V}_t$ **do**
      $\alpha_{r,i}^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s}') O(\boldsymbol{z}, \boldsymbol{s}', \boldsymbol{a}) \alpha_r^i(\boldsymbol{s}')$
      $\alpha_{c,i}^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s}') O(\boldsymbol{z}, \boldsymbol{s}', \boldsymbol{a}) \alpha_c^i(\boldsymbol{s}')$
      $\Gamma_r^{\boldsymbol{a},\boldsymbol{z}} = \Gamma_r^{\boldsymbol{a},\boldsymbol{z}} \cup \alpha_{r,i}^{\boldsymbol{a},\boldsymbol{z}}$ and $\Gamma_c^{\boldsymbol{a},\boldsymbol{z}} = \Gamma_c^{\boldsymbol{a},\boldsymbol{z}} \cup \alpha_{c,i}^{\boldsymbol{a},\boldsymbol{z}}$
    end for
  end for
  $\Gamma_r^{\boldsymbol{a}} = \alpha_r^{\boldsymbol{a},*} + \bigoplus_{\boldsymbol{z} \in \mathcal{Z}} \Gamma_r^{\boldsymbol{a},\boldsymbol{z}}$ and $\Gamma_c^{\boldsymbol{a}} = \alpha_c^{\boldsymbol{a},*} + \bigoplus_{\boldsymbol{z} \in \mathcal{Z}} \Gamma_c^{\boldsymbol{a},\boldsymbol{z}}$
end for
$\mathcal{V}_{t+1} = \mathcal{V}_{t+1} \bigcup Prune\left(\bigcup_{\boldsymbol{a}} (\Gamma_r^{\boldsymbol{a}}, \Gamma_c^{\boldsymbol{a}})\right)$
Output: $\mathcal{V}_{t+1}$

Due to PWLC propriety of the value function, the value iteration algorithm of CPOMDP is limited to find the set of α-vector pairs $\mathcal{V}_{t+1}$ that represents the value functions V_{t+1}^r and V_{t+1}^c given the previous set $\mathcal{V}_t$. Constructing the set of hyperplans $\mathcal{V}_{t+1}$ by considering all the possible pairs of observations and actions over the previous set $\mathcal{V}_t$ has an exponential complexity of $O\left(|\mathcal{A}||\mathcal{V}_t|^{|\mathcal{Z}|}\right)$ (i.e. considering the states it

gives $O\Big(|\boldsymbol{\mathcal{S}}|^2|A||V_t|^{|\mathcal{Z}|}\Big)$). Since many pairs of vectors in $\mathcal{V}_t$ are dominated by others, pruning algorithms are developed in order to eliminate useless vectors and find the smallest subset sufficient for representing the value functions. Given the exponential complexity increase at each iteration, the necessity of pruning operations is obvious. However, pruning techniques consists on resolving a Mixed-Integer Linear Program MILP for each pair of α-vector. Therefore, value iteration algorithm for CPOMDP remains computationally demanding to solve as the size of the problem increases. This requires the exploration of approximated solutions for finding the optimal solution of CPOMDP.

5.4.2 CPBVI

Since the value iteration algorithms for POMDP do not scale to highly sized real problems, an approximate POMDP planning solution called Point-Based Value Iteration (PBVI) was introduced in [36]. Most POMDP problems unlikely reach most of the points in the belief simplex Δ. Thus, it is preferable to focus the planning on the most probable belief points without considering all the possible belief points as exact algorithms do. Instead of considering the entire belief simplex, PBVI limits the value update to a representative small set of belief points $\mathcal{B} = \{\boldsymbol{b}_0, \boldsymbol{b}_1, ..., \boldsymbol{b}_q\}$. An α-vector is initialized for each belief point and then the value of this vector is iteratively updated. The PBVI algorithm can be simply adapted to solve CPOMDP problem (e.g. [107]).

5.4.2.1 CPBVI algorithm

We call CPBVI the proposed suboptimal algorithm of relay selection inspired from PBVI. Algorithm 6 is modified in such a way that the value function update is restrictively done over a finite belief set $\mathcal{B}$ and then the pruning algorithm chooses the dominated pair of vectors for each belief state $\boldsymbol{b} \in \mathcal{B}$. This allows the CPBVI algorithm to achieve much better scalability. At each iteration, CPBVI follows the following steps for computing the set of $\alpha-$vectors $\mathcal{V}_{t+1}^B$ at epoch $t+1$ given the previous one $\mathcal{V}_t^B$:

1. The first step is to generate the sets $\Gamma_r^{\boldsymbol{a},\boldsymbol{z}}$ and $\Gamma_c^{\boldsymbol{a},\boldsymbol{z}}$ for all $\boldsymbol{a} \in \mathcal{A}$, all $\boldsymbol{z} \in \mathcal{Z}$ and all pairs of α-vectors $(\alpha_r^i, \alpha_c^i) \in \mathcal{V}_t^B$:

$$\Gamma_r^{\boldsymbol{a},\boldsymbol{z}} \leftarrow \alpha_r^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s}')\, O(\boldsymbol{z}, \boldsymbol{s}', \boldsymbol{a})\, \alpha_r^i(\boldsymbol{s}')$$

$$\Gamma_c^{\boldsymbol{a},\boldsymbol{z}} \leftarrow \alpha_c^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s}')\, O(\boldsymbol{z}, \boldsymbol{s}', \boldsymbol{a})\, \alpha_c^i(\boldsymbol{s}')$$

2. The next step is to generate the sets $\Gamma_r^{\boldsymbol{a}}$ and $\Gamma_c^{\boldsymbol{a}}$ for all $\boldsymbol{a} \in \mathcal{A}$:

$$\Gamma_r^{\boldsymbol{a}} = R(., \boldsymbol{a}) + \bigoplus_{\boldsymbol{z} \in \mathcal{Z}} \Gamma_r^{\boldsymbol{a},\boldsymbol{z}} \text{ and } \Gamma_c^{\boldsymbol{a}} = C(., \boldsymbol{a}) + \bigoplus_{\boldsymbol{z} \in \mathcal{Z}} \Gamma_c^{\boldsymbol{a},\boldsymbol{z}}$$

Contrarily to PBVI algorithm used for approximately solving POMDP, we need the cross-summation overall the possible observations to find the sets $\Gamma_r^{\boldsymbol{a}}$ and $\Gamma_c^{\boldsymbol{a}}$. This cross-summation is mandatory in order to not impose the cost constraint on each action and observation pair while computing the best pair of α-vector for each belief state $\boldsymbol{b}$. A local combinatorial explosion of $|\mathcal{A}||\mathcal{B}|^{|\mathcal{Z}|}$ is required.

3. The final step is the pruning operation. In our algorithm we propose to find one optimal pair of α-vectors $\left(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\right)$ for each belief point $\boldsymbol{b} \in \mathcal{B}$ as follows:

$$\left(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\right) = \underset{\alpha_r \in \Gamma_r^{\boldsymbol{a}}, \alpha_c \in \Gamma_c^{\boldsymbol{a}}}{\operatorname{argmax}} \{\alpha_r.\boldsymbol{b} : \alpha_c.\boldsymbol{b} \leq C_{th}\}$$

Note that the proposed deterministic policy ensures the satisfaction of the cumulative cost constraint 5.3 at each epoch. Thus, such deterministic policies can be sub-optimal for CPOMDP (by analogy to CMDP). Ideally, randomized policies that consider convex combination of α-vectors during the pruning operation can be applied to guarantee optimality.

4. Finally $\mathcal{V}_{t-1}^B = \bigcup\limits_{\boldsymbol{b} \in \mathcal{B}} \left(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\right)$.

Algorithm 7 CPBVI Iteration

Input: Pair of α-vectors $\mathcal{V}_t^B$, Actions $\mathcal{A}$, States $\boldsymbol{\mathcal{S}}$, Observations $\mathcal{Z}$, Rewards $R(\boldsymbol{s}, \boldsymbol{a})$, costs $C(\boldsymbol{s}, \boldsymbol{a})$, Belief subset $\mathcal{B}$, thresholds C_{th}
for $\boldsymbol{a} \in \mathcal{A}$ **do**
  $(\alpha_r^{\boldsymbol{a},*}, \alpha_c^{\boldsymbol{a},*}) \leftarrow (R(., \boldsymbol{a}), C(., \boldsymbol{a}))$
  for $\boldsymbol{z} \in \mathcal{Z}$ **do**
    for $(\alpha_r^i, \alpha_c^i) \in \mathcal{V}_t^B$ **do**
      $\alpha_{r,i}^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum\limits_{\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s}') O(\boldsymbol{s}', \boldsymbol{a}, \boldsymbol{z}) \alpha_r^i(\boldsymbol{s}')$
      $\alpha_{c,i}^{\boldsymbol{a},\boldsymbol{z}}(\boldsymbol{s}) = \gamma \sum\limits_{\boldsymbol{s}' \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s}') O(\boldsymbol{s}', \boldsymbol{a}, \boldsymbol{z}) \alpha_c^i(\boldsymbol{s}')$
      $\Gamma_r^{\boldsymbol{a},\boldsymbol{z}} = \Gamma_r^{\boldsymbol{a},\boldsymbol{z}} \cup \alpha_{r,i}^{\boldsymbol{a},\boldsymbol{z}}$ and $\Gamma_c^{\boldsymbol{a},\boldsymbol{z}} = \Gamma_c^{\boldsymbol{a},\boldsymbol{z}} \cup \alpha_{c,i}^{\boldsymbol{a},\boldsymbol{z}}$
    end for
  end for
  $\Gamma_r^{\boldsymbol{a}} = \alpha_r^{\boldsymbol{a},*} + \bigoplus\limits_{\boldsymbol{z} \in \mathcal{Z}} \Gamma_r^{\boldsymbol{a},\boldsymbol{z}}$ and $\Gamma_c^{\boldsymbol{a}} = \alpha_c^{\boldsymbol{a},*} + \bigoplus\limits_{\boldsymbol{z} \in \mathcal{Z}} \Gamma_c^{\boldsymbol{a},\boldsymbol{z}}$
end for
for $\boldsymbol{b} \in \mathcal{B}$ **do**
  $\left(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\right) = \underset{\alpha_r \in \Gamma_r^{\boldsymbol{a}}; \alpha_c \in \Gamma_c^{\boldsymbol{a}}}{\operatorname{argmax}} \sum\limits_{\boldsymbol{s} \in \boldsymbol{\mathcal{S}}} \alpha_r(\boldsymbol{s}) b(\boldsymbol{s})$ s.t. $\sum\limits_{\boldsymbol{s} \in \boldsymbol{\mathcal{S}}} \alpha_c(\boldsymbol{s}) b(\boldsymbol{s}) \leq C_{th}$
  $\mathcal{V}_{t+1}^B = \mathcal{V}_{t+1}^B \bigcup \left(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\right)$
end for
Output: $\mathcal{V}_{t+1}^B$

In order to study the complexity of each iteration of the CPBVI algorithm, we detail the complexity of each step. Step 1 creates $|\mathcal{A}||\mathcal{Z}||\mathcal{B}|$ pair of vectors because the previous set of α-vector is limited to $|\mathcal{B}|$ components. The cross sum in the second step generates $O\left(|\mathcal{A}||\mathcal{B}|^{|\mathcal{Z}|}\right)$ operations. Since the size of the α-vector set

remains constant (equals to $|\mathcal{B}|$), each CPBVI update takes only polynomial time to be executed. This complexity is linear with the number of possible actions $|\mathcal{A}|$. However, in our settings, $|\mathcal{A}| = 2^K$ leads to a poor scalability of our algorithm in the number of potential relays K. Therefore, we propose, in subsection 5.4.3, a greedy version of the CPBVI algorithm.

5.4.2.2 Performance Evaluation

A belief set $\mathcal{B}$ is characterized by a density $\epsilon^{\mathcal{B}}$ which is the maximum distance from any point in the belief simplex Δ to the set $\mathcal{B}$.

Definition 5.4.1. *The density $\epsilon^{\mathcal{B}}$ of a belief set $\mathcal{B}$ is defined as:*

$$\epsilon^{\mathcal{B}} := \max_{\boldsymbol{b}' \in \Delta} \min_{\boldsymbol{b} \in \mathcal{B}} ||\boldsymbol{b} - \boldsymbol{b}'||_1 \tag{5.19}$$

We denote by $V_t^{r,B}$ and $V_t^{c,B}$ the reward and cost value functions produced by the proposed CPBVI algorithm. From theorem 1 of [36], we know that for a belief set $\mathcal{B}$ of density $\epsilon^{\mathcal{B}}$, the errors η_t^r and η_t^c of the CPBVI algorithm at horizon t are bounded as follows:

$$\begin{aligned} \eta_t^r &:= ||V_t^{r,B} - V_t^{r,*}||_\infty \leq \frac{(R_{max} - R_{min})\,\epsilon^{\mathcal{B}}}{(1-\gamma)^2} \\ \eta_t^c &:= ||V_t^{c,B} - V_t^{c,*}||_\infty \leq \frac{(C_{max} - C_{min})\,\epsilon^{B}}{(1-\gamma)^2} \end{aligned} \tag{5.20}$$

One can remark that this result does not take into account the case where the discount factor is equal to 1. Hence, we extend the result to the case where $\gamma = 1$.

Theorem 5.4.2. *At horizon h and discount factor $\gamma = 1$, the errors η_h^r and η_h^c of applying the CPBVI algorithm over a belief set $\mathcal{B}$ of density $\epsilon^{\mathcal{B}}$ are bounded as follows:*

$$\begin{aligned} \eta_h^r &\leq \sum_{t=1}^{h} t\,(R_{max} - R_{min})\,\epsilon^{\mathcal{B}} = \frac{h\,(h+1)}{2}\,(R_{max} - R_{min})\,\epsilon^{\mathcal{B}} \\ \eta_h^c &\leq \sum_{t=1}^{h} t\,(C_{max} - C_{min})\,\epsilon^{\mathcal{B}} = \frac{h\,(h+1)}{2}\,(C_{max} - C_{min})\,\epsilon^{\mathcal{B}} \end{aligned} \tag{5.21}$$

5.4.2.3 Belief set $\mathcal{B}$ Selection

The bound calculation of the error that occurs with the application of the CPBVI algorithm shows the importance of the selected belief set $\mathcal{B}$. Indeed, these bounds are proportional to the density $\epsilon^{\mathcal{B}}$ of the chosen belief set $\mathcal{B}$. For this reason we show how to find the belief set, denoted by $\mathcal{B}^\epsilon$, that limits the upper bounds η_h^r and η_h^c to a small ϵ. It is obvious that a trade-off exists between the size of the belief set $\mathcal{B}$ and the precision of the value functions ϵ.

We denote by $\bar{\Delta}^h$ the set of reachable belief states in our settings with an horizon h. Therefore, for a relay i, the corresponding set of reachable beliefs $\bar{\Delta}_i^h$ is given by:

$$\bar{\Delta}_i^h = \{\boldsymbol{P}_i^t\left(S_j,:\right) : t = \{1,...,h\} \text{ and } S_j \in \mathcal{S}\} \tag{5.22}$$

where $\boldsymbol{P}_i^t$ correspond to the transition matrix $\boldsymbol{P}_i$ to the power of t (i.e. the transition matrix from a state to another after t epochs). Therefore, the set $\bar{\Delta}_i^h$ contains $h \times |\mathcal{S}|$ vectors of $|\mathcal{S}|$ elements. Thus, considering all the candidate relays K, the size of the overall reachable belief points for an horizon h is given by $|\bar{\Delta}^h| = h|\mathcal{S}|^K$.

It is clear that, for an horizon h, considering the set of reachable belief points $\bar{\Delta}^h$ as the belief set $\mathcal{B}$ of the CPBVI algorithm minimizes the errors η_h^r and η_h^c (i.e. because all the reachable belief points were taken into account in $\mathcal{B}$). However, due to the exponentially large size of the set $\bar{\Delta}^h$, we propose to construct a smaller set of belief points $\mathcal{B}^\epsilon$ that is sufficient for upper bounding the performance errors η_h^r and η_h^c by a small $\epsilon > 0$. We denote by $\mathcal{N}_t\left(\boldsymbol{s}\right)$ the set of reachable states after passing t epochs and knowing that the initial state is $\boldsymbol{s}$ (i.e. $\mathcal{N}_1\left(\boldsymbol{s}\right)$ represents the neighbor states of $\boldsymbol{s}$). For constructing the belief set $\mathcal{B}^\epsilon$, we give the following definition.

Definition 5.4.3. *We define the h-belief set $\mathcal{B}\left(\boldsymbol{s},h\right)$ for an initial state $\boldsymbol{s}_0$ and an horizon h as follows:*

$$\mathcal{B}\left(\boldsymbol{s},h\right) = \{\bigcup_{t=1}^{h} \boldsymbol{T}^n\left(\mathcal{N}_t\left(\boldsymbol{s}\right),:\right) : n = 1,...,h-t\} \tag{5.23}$$

with $\boldsymbol{T}^n$ the transition matrix $\boldsymbol{T}$ power to n.

We emphasize that the size of a set $|\mathcal{B}\left(\boldsymbol{s}_0,h\right)|$ is bounded by $|\mathcal{B}\left(\boldsymbol{s}_0,h\right)| \leq \sum_{t=1}^{h}\left(h-t\right)\mathcal{N}_1\left(\boldsymbol{s}_0\right)^{t-1}$.

Theorem 5.4.4. *The density $\epsilon^{\mathcal{B}}$ of an h-belief set $\mathcal{B}\left(\boldsymbol{s},h\right)$ is bounded by:*

$$\epsilon^{\mathcal{B}}\left(\mathcal{B}\left(\boldsymbol{s},h\right)\right) \leq 2\sum_{i=1}^{K}\frac{\lambda_i^{*h}}{\pi_{i,min}} \tag{5.24}$$

where λ_i^* is the highest eigenvalue of the transition matrix $\boldsymbol{P}_i$ of relay i and $\pi_{i,min} = \min_{s\in\mathcal{S}}\pi_i\left(s\right)$ with π_i the stationary distribution corresponding to $\boldsymbol{P}_i$.

Proof. Please refer to the proofs section 7.5.1. □

Theorem 5.4.5. *For an initial state $\boldsymbol{s}_0$ and a problem of horizon T and discount factor $\gamma < 1$, the belief set $\mathcal{B}^\epsilon$ that should be selected in order to achieve an ϵ performance (i.e. $\eta_T^r \leq \epsilon$ and $\eta_T^c \leq \epsilon$) is given by:*

$$\mathcal{B}^\epsilon = \mathcal{B}\left(\boldsymbol{s}_0, \left\lceil \min\left(\frac{f_r\left(\epsilon\right)}{\log\left(\lambda^*\right)}, \frac{f_c\left(\epsilon\right)}{\log\left(\lambda^*\right)}\right)\right\rceil\right) \tag{5.25}$$

with

$$f_r(\epsilon) = \log\left(\frac{\epsilon\pi_{min}(1-\gamma)^2}{2K(R_{max}-R_{min})}\right); f_c(\epsilon) = \log\left(\frac{\epsilon\pi_{min}(1-\gamma)^2}{2K(C_{max}-C_{min})}\right)$$

$$\lambda^* = \max_{i\in\mathcal{K}} \lambda_i^* \ ; \ \pi_{min}^* = \min_{i\in\mathcal{K}} \pi_{i,min}^*$$

Proof. Please refer to the section 7.5.2 for the proof of the theorem as well as its extension to $\gamma = 1$. □

5.4.3 Greedy CPBVI

The computational complexity of each iteration of the CPBVI algorithm is proportional to the size of the set of actions $|\mathcal{A}| = 2^K$ which increases exponentially with the number of potential relays K. Therefore, we propose a GCPBVI algorithm that exploits greedy maximization which consists of iteratively choosing the relays that should be selected in the aim of optimizing the value functions of the problem. This can be done by replacing argmax in line 14 of algorithm 7 by $greedy - argmax$ (see section 2.4 for more details). Greedy maximization is the main motivation for considering CPBVI algorithms to solve CPOMDP. In fact, CPBVI algorithms perform argmax operations over a finite set of belief points which is essential for the application of the greedy approach (i.e. contrarily to exact methods that compute the value function over all the continuous belief simplex). The GCPBVI algorithm is basically deduced from the submodularity property of the Q-function.

5.4.3.1 Submodularity of Q-function

Since the upcoming result can be applied on both reward and cost Q-functions (given respectively by equations (5.15) and (5.16)), we omit the index r (for reward) and c (for cost) and use Q_t^π notation to generalize the results on both $Q_t^{r,\pi}$ and $Q_t^{c,\pi}$. Recall that:

$$Q_t^\pi(\boldsymbol{b},\boldsymbol{a}) = \rho(\boldsymbol{b},\boldsymbol{a}) + \sum_{\boldsymbol{z}\in Z} Pr(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}) V_{t+1}^\pi(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}) = \rho(\boldsymbol{b},\boldsymbol{a}) + \sum_{k=t+1}^{T} G_k^\pi\left(\boldsymbol{b}^t,\boldsymbol{a}^t\right) \quad (5.26)$$

where $G_k^\pi(\boldsymbol{b}^t,\boldsymbol{a}^t)$ is the expected immediate value (reward or cost) under policy π at epoch k conditioned on the belief $\boldsymbol{b}_t$ and the action $\boldsymbol{a}_t$ at epoch t, thus:

$$G_k^\pi\left(\boldsymbol{b}^t,\boldsymbol{a}^t\right) = \gamma^k \sum_{\boldsymbol{z}^{t:k}} P\left(\boldsymbol{z}^{t:k}|\boldsymbol{b}^t,\boldsymbol{a}^t,\pi\right)\rho\left(\boldsymbol{b}^k,\boldsymbol{a}^k|\pi\right) \quad (5.27)$$

where $\boldsymbol{z}^{t:k}$ is a vector of observations received in the interval t to k epochs.

Theorem 5.4.6. *For all policies π, $Q_t^{r,\pi}(\boldsymbol{b},\boldsymbol{a})$ and $Q_t^{c,\pi}(\boldsymbol{b},\boldsymbol{a})$ are non-negatives, monotones and submodulars in $\boldsymbol{a}$.*

Proof. Please refer to section 7.5.3. □

5.4.3.2 GCPBVI algorithm

Since the Q-functions in terms of reward and cost are submodulars, then based on the algorithm 2, a greedy version of the CPBVI algorithm is deduced (see algorithm 8). The objective of this greedy algorithm is to avoid the iteration over all the possible actions $\mathcal{A}$ of size 2^K. The iteration of the GCPBVI algorithm shows that we limit the computation on considering each relay $i \in \mathcal{K}$ aside in such a way that the set $\mathcal{A}$ of all possible action is not introduced at any level of the algorithm (i.e. avoiding by that the exponential complexity in K). Indeed, at each iteration, the GCPBVI follows the steps below for computing the set of $\alpha-$vectors $\mathcal{V}_{t+1}^G$ at epoch $t+1$ given the previous set $\mathcal{V}_t^G$:

1. The first step is to generate the sets Γ_r^{k,z_k} and Γ_c^{k,z_k} for all the relays $k \in \mathcal{K}$, all possible observations of each relay $z_k \in \mathcal{S}$ and all pairs of α-vectors $(\alpha_r^i, \alpha_c^i) \in \mathcal{V}_t^G$:
$$\Gamma_r^{k,z_k} \leftarrow \alpha_r^{k,z_k}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s'} \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s'}) O(z_k, \boldsymbol{s'}, k) \alpha_r^i(\boldsymbol{s'})$$
$$\Gamma_c^{k,z_k} \leftarrow \alpha_c^{k,z_k}(\boldsymbol{s}) = \gamma \sum_{\boldsymbol{s'} \in \boldsymbol{\mathcal{S}}} T(\boldsymbol{s}, \boldsymbol{s'}) O(z_k, \boldsymbol{s'}, k) \alpha_c^i(\boldsymbol{s'})$$
2. The next step is to generate the sets Γ_r^k and Γ_c^k for each relay $k \in \mathcal{K}$:
$$\Gamma_r^k = R(., \boldsymbol{a}) + \bigoplus_{z_k \in \mathcal{S}} \Gamma_r^{k,z_k} \text{ and } \Gamma_c^k = C(., \boldsymbol{a}) + \bigoplus_{z_k \in \mathcal{S}} \Gamma_c^{k,z_k}$$
The computation of the sets Γ_r^k and Γ_c^k of each relay $k \in \mathcal{K}$ has a complexity of $O\Big(|\mathcal{S}||\mathcal{B}|^{|\mathcal{S}|}\Big)$.
3. The final step aims to find the optimal action to take for each belief state $\boldsymbol{b} \in \mathcal{B}$. The *greedymax* optimization given in algorithm 2 is applied:
$$\Big(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\Big) = greedy \operatorname*{argmax}_{\alpha_r, \alpha_c} \{\alpha_r.\boldsymbol{b} : \alpha_c.\boldsymbol{b} \leq C_{th}\}$$
The greedy optimization complexity is limited to $O(K^2)$.
4. Finally $\mathcal{V}_{t+1}^G = \bigcup_{\boldsymbol{b} \in \mathcal{B}} \Big(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\Big)$.

Thus, the complexity of one GCPBVI iteration $O\Big(K^2|\mathcal{S}||\mathcal{B}|^{|\mathcal{S}|}\Big)$ enables much better scalability in the number of potential relays K. In the sequel, we denote by $V_t^{r,G}$ the reward value function generated by the GCPBVI algorithm.

5.4.3.3 Performance Evaluation

Theorem 5.4.7. *At a given epoch t, the error in the reward value function due to greedy optimization is bounded by:*

$$V_t^{r,G}(\boldsymbol{b}) \geq \left(1 - \frac{1}{e}\right)^{2t} V_t^{r,B}(\boldsymbol{b}) \tag{5.28}$$

Proof. Please refer to proofs section 7.5.4. □

Algorithm 8 GCPBVI Iteration

Input: Pair of α-vectors $\mathcal{V}_t^G$, Relays $\mathcal{K}$, States $\mathcal{S}$, Observations $\mathcal{Z}$, Rewards $r_i(s)$, Costs $c_i(s)$ for relay i, Belief subset $\mathcal{B}^\epsilon$, Threshold C_{th}

for $k = 1 : K$ **do**
  $\left(\alpha_r^{k,*}, \alpha_c^{k,*}\right) \leftarrow (r_i(.), c_i(.))$
  for $z_k \in \mathcal{S}$ **do**
    for $(\alpha_r^i, \alpha_c^i) \in \mathcal{V}_t^G$ **do**
      $\alpha_{r,i}^{k,z_k}(\boldsymbol{s}) = \sum_{\boldsymbol{s}' \in \mathcal{S}} T(\boldsymbol{s}, \boldsymbol{s}') O(z_k, \boldsymbol{s}', k) \alpha_r^i(\boldsymbol{s}')$
      $\alpha_{c,i}^{k,z_k}(\boldsymbol{s}) = \sum_{\boldsymbol{s}' \in \mathcal{S}} T(\boldsymbol{s}, \boldsymbol{s}') O(z_k, \boldsymbol{s}', k) \alpha_c^i(\boldsymbol{s}')$
      $\Gamma_r^{k,z_k} = \Gamma_r^{k,z_k} \cup \alpha_{r,i}^{k,z_k}$ and $\Gamma_c^{k,z_k} = \Gamma_c^{k,z_k} \cup \alpha_{c,i}^{k,z_k}$
    end for
  end for
  $\Gamma_r^k = \alpha_r^{k,*} + \bigoplus_{z_k \in \mathcal{S}} \Gamma_r^{k,z_k}$ and $\Gamma_c^k = \alpha_c^{k,*} + \bigoplus_{z_k \in \mathcal{S}} \Gamma_c^{k,z_k}$
end for
for $\boldsymbol{b} \in \mathcal{B}^\epsilon$ **do**
  $v_{sum} = 0$, $\alpha_r^{\boldsymbol{b}} = 0$, $\alpha_c^{\boldsymbol{b}} = 0$, $K_{left} = \mathcal{K}$, $\Gamma^{\boldsymbol{a},\boldsymbol{b}} = \emptyset$
  while $|K_{left}| > 0$ **do**
    $(k^*, \alpha_r^*, \alpha_c^*) = \underset{k \in K_{left}; \alpha_r \in \Gamma_r^k; \alpha_c \in \Gamma_c^k}{\operatorname{argmax}} \frac{\sum_{\boldsymbol{s} \in \boldsymbol{\mathcal{S}}} \alpha_r(\boldsymbol{s}) b(\boldsymbol{s})}{\sum_{\boldsymbol{s} \in \boldsymbol{\mathcal{S}}} \alpha_c(\boldsymbol{s}) b(\boldsymbol{s})}$
    if $v_{sum} + \alpha_c^* < C_{th}$ **then**
      $\alpha_r^{\boldsymbol{b}} = \alpha_r^{\boldsymbol{b}} + \alpha_r^*$, $\alpha_c^{\boldsymbol{b}} = \alpha_c^{\boldsymbol{b}} + \alpha_c^*$ and $v_{sum} = v_{sum} + \alpha_c^* \boldsymbol{b}$
      $\Gamma^{\boldsymbol{a},\boldsymbol{b}} = \Gamma^{\boldsymbol{a},\boldsymbol{b}} \cup k^*$
    end if
    $K_{left} = K_{left} | k^*$
  end while
  $\mathcal{V}_{t+1}^G = \mathcal{V}_{t+1}^G \bigcup \left(\alpha_r^{\boldsymbol{b}}, \alpha_c^{\boldsymbol{b}}\right)$ and $\Gamma_{t+1}^{\boldsymbol{a}} = \Gamma_{t+1}^{\boldsymbol{a}} \bigcup \Gamma^{\boldsymbol{a},\boldsymbol{b}}$
end for
Output: $\mathcal{V}_{t+1}^G$ and Γ_{t+1}^a

5.5 Extension to Multi player scenario

The system model presented in section 5.2 is extended to the multi player scenario where N MU aim to use D2D aided relaying in order to improve the performance of their cellular communications. For the multi-player scenario, we generalized the single player notation as follows:

- $\mathcal{K}_i$ of size K_i elements denotes the set of candidate relays discovered by each MU i with $i \in \{1, ..., N\}$. Thus, the set of potential relays is given by $\mathcal{K} = \bigcap_{i=1}^{N} \mathcal{K}_i$ with $K = |\mathcal{K}|$.
- $\overrightarrow{\boldsymbol{a}} = (\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_N)$ denotes the matrix of all the actions $\boldsymbol{a}_i$ taken by each MU i with $i \in \{1, ..., N\}$. Thus, the size of the set of possible actions $|\mathcal{A}| = 2^K$. The

size of matrix $\overrightarrow{\boldsymbol{a}}$ is $K \times N$ with $a(i,j) = 1$ if relay i is selected by MU j and 0 otherwise.

- $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, ..., \boldsymbol{s}_K)$ denotes the matrix of all the states $\boldsymbol{s}_i$ of each relay i with $i \in \{1, ..., K\}$. Thus, the size of the set of possible states $|\boldsymbol{\mathcal{S}}| = \mathcal{S}^K$. We denote by $\boldsymbol{s}_{\mathcal{K}_i}$ the state vector of the candidate relays of MU i.
- The reward and cost model for a given action $\overrightarrow{\boldsymbol{a}} = (\boldsymbol{a}_1, \boldsymbol{a}_2, ..., \boldsymbol{a}_N)$ and a given state $\boldsymbol{s} = (\boldsymbol{s}_1, \boldsymbol{s}_2, ..., \boldsymbol{s}_K)$ is the following:

$$R\left(\boldsymbol{s}, \overrightarrow{\boldsymbol{a}}\right) = \sum_{i=1}^{N} R_i\left(\boldsymbol{s}_{\mathcal{K}_i}, \boldsymbol{a}_i\right) \tag{5.29}$$

where $R_i\left(\boldsymbol{s}_{\mathcal{K}_i}, \boldsymbol{a}_i\right)$ of each MU i is given by equation (5.1).

$$C\left(\boldsymbol{s}, \overrightarrow{\boldsymbol{a}}\right) = \sum_{i=1}^{N} C_i\left(\boldsymbol{s}_{\mathcal{K}_i}, \boldsymbol{a}_i\right) \tag{5.30}$$

where $C_i\left(\boldsymbol{s}_{\mathcal{K}_i}, \boldsymbol{a}_i\right)$ of each MU i is given by equation (5.2).

For this multi player scenario, the relay selection scheme aims to find the decision policy $\left(\overrightarrow{\boldsymbol{a}}_1, ..., \overrightarrow{\boldsymbol{a}}_T\right)$ that optimizes the following problem:

$$\max \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t R\left(\boldsymbol{s}_t, \overrightarrow{\boldsymbol{a}}_t\right)\right]$$

$$\text{s.t. } \mathbb{E}\left[\sum_{t=1}^{T} \gamma^t C_i\left(\boldsymbol{s}_t, \overrightarrow{\boldsymbol{a}}_t\right)\right] \leq C_{th} \ \forall i \in \{1, ..., N\} \tag{5.31}$$

5.5.1 Centralized Relay Selection Policy

The first strategy of relay selection is to make the decision by the BS in a centralized manner. Thus, the problem 5.31 can be modeled, based on the single player scenario given in subsection 5.3.2, as a CPOMDP with the following characteristics: $\boldsymbol{s}$ represents the state of all the potential relays; $\overrightarrow{\boldsymbol{a}}$ represents the matrix of the taken action (with $\overrightarrow{\boldsymbol{a}}(i,j) = 1$ if relay i is selected by MU j); $\boldsymbol{T}$ represents the transition matrix with $\boldsymbol{T}(\boldsymbol{s}; \boldsymbol{s}')$ the probability of passing from a state $\boldsymbol{s}$ to another $\boldsymbol{s}'$; reward $R\left(\boldsymbol{s}, \overrightarrow{\boldsymbol{a}}\right)$ and cost $C\left(\boldsymbol{s}, \overrightarrow{\boldsymbol{a}}\right)$ models are respectively given by equations (5.29) and (5.30); $\boldsymbol{b}_0$ represents the initial belief of being in each state $\boldsymbol{s} \in \boldsymbol{\mathcal{S}}$; T is the horizon of the problem and C_{th} the cost threshold that should not be exceed by any MU. Considering a centralized approach leads to the assumption that the state of each relay i is reported to the BS each time this relay i is selected by one of the MUs. Therefore, the observation $\boldsymbol{z} = (z_1, z_2, ..., z_K)$ corresponds to $z_i = s_i$ if relay i is selected by any MU ($\overrightarrow{\boldsymbol{a}}(i,j) = 1$ for any $j \in \mathcal{K}$) and $z_i = \emptyset$ otherwise ($\overrightarrow{\boldsymbol{a}}(i,j) = 0$ for all $j \in \mathcal{K}$). $O\left(\boldsymbol{z}', \boldsymbol{s}', \overrightarrow{\boldsymbol{a}}\right)$ represents the probability of receiving an observation $\boldsymbol{z}'$ knowing that action $\overrightarrow{\boldsymbol{a}}$ is taken and that the relays transit to state $\boldsymbol{s}'$.

The motivation of such centralized approach lies on the observation process described above. Since the BS is informed by the state of each selected relay, BS will have a global knowledge of all the selected relays' state contrarily to the local observation of each MU which is limited by its own decision. Therefore, in this centralized approach, the BS profits from having a global state information for inducing a more efficient relay selection decision.

This CPOMDP is a straight-forward generalization of the single player CPOMDP formulation. Thus, the results provided in the previous sections remain intact. This includes the GCPBVI algorithm that is still applicable for the case of multi-player scenario since the submodularity property of the Q-functions remains valid. However, this centralized approach suffers from two challenges:

1. **Overhead**: This centralized approach requires that each MU reports the states of its selected relays. This procedure generates an overhead of signaling.

2. **Complexity**: The state space for multi-MUs scenario will blowup because it exponentially increase with the total number of MUs N as well as the total number of candidates relays K. Indeed, the total number of states is $|\boldsymbol{\mathcal{S}}| = |\mathcal{S}|^{|\bigcap_{n=1}^{N} \mathcal{K}_n|}$.

5.5.2 Distributed Relay Selection Policy

To address the exponentially increasing in the size of the state space for multi-player scenario and to avoid the overhead of signaling, we propose a distributed variant for resolving the generalized CPOMDP. We divide the multi-player problem into N single-player problems (given by 5.3), one for each MU, and solve them independently. For such distributed approach, each MU will not take advantage of the observation of other MUs because the states selected by each MU are not shared between each others. However, distributed designs remain interesting for escaping the large amount of signaling that is required to be exchanged as well as reducing the computational complexity at the BS level.

The distributed approach is equivalent to considering N parallel and independent single MU problems. Each MU n (with $n = \{1, ..., N\}$) formulates its own relay selection problem as a CPOMDP as shown in section 5.7. Then, each MU launches its relay selection strategy based on the GCPBVI algorithm proposed in 5.4.3. Numerical results 5.6 show that this distributed approach reduces the complexity of the problem while achieving performance close to that of the centralized approach.

5.6 Numerical Results

5.6.1 Numerical Settings

We evaluate our claims by considering the case where the reward that we aim to maximize is the throughput of the MU to BS communications under energy consumption constraints. We consider that the relays are moving within different rectangular regions (i.e. we used to refer to these regions as states in the POMDP formulation). An example of partitioning the horizontal plane as well as the vertical plane into 5 parts is illustrated in figure 5.1 with $N = 3$ MUs and $K = 6$ relays. This example of partitioning generates 25 different locations for mobile relays. It is obvious that the preciseness and the performance of the results are improved by considering smaller partitioning granularity. However, this may lead to increase the complexity of the solution.

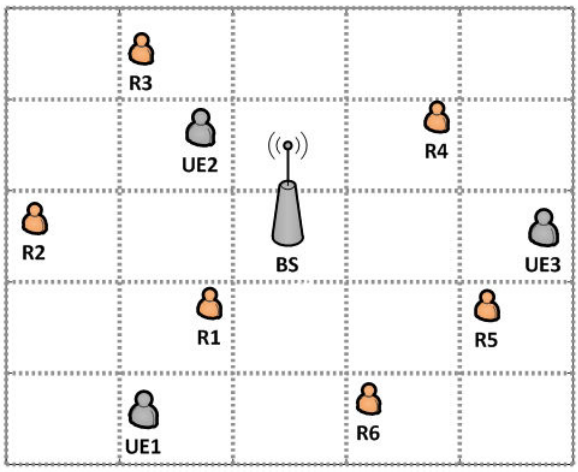


Figure 5.1: D2D relaying scenario for numerical results

We call S_x (resp. S_y) the number of divisions that is applied at the horizontal (resp. vertical) axis. Each state $s \in \mathcal{S}$ is defined by a x and y coordinates. We denote by ϵ_{fix} as the probability that the relay stays in the same region. This parameter is essential for defining the mobility pattern that is assumed in this numerical section. The transition matrix in the horizontal plane $\boldsymbol{P}_x$ is constructed in such a way that each relay move left or right with an equal probability of $\frac{1}{2}\left(1-\sqrt{\epsilon_{\text{fix}}}\right)$. The transition matrix in the vertical plane $\boldsymbol{P}_y$ is constructed in such a way that each relay move up or down with an equal probability of $\frac{1}{2}\left(1-\sqrt{\epsilon_{\text{fix}}}\right)$. Considering that the horizontal and vertical travels are independent, the transition matrix of each relay i $\boldsymbol{P}_i$ is deduced as follows: $\boldsymbol{P}_i\left[(x,y),(x',y')\right] = \boldsymbol{P}_x\left(x,x'\right)\boldsymbol{P}_y\left(y,y'\right)$. One can remark that this structure of mobility matrix leads to a probability ϵ_{fix} of staying in the same location. The value of ϵ_{fix} is given in table 5.1.

Settings Parameter	Value
Immobile Probability ϵ_{fix}	0.7
Max. Reward R_{max}	500 Kbps/RB
Max. Cost C_{max}	250 mW
Cost Threshold C_{th}	1000 mW
Reward and Cost Model	Given by 5.32 and 5.33
Horizon T	5
Discount Factor γ	1
Number of realizations	100
ϵ^B of CPBVI algorithm	0.01

Table 5.1: Numerical Settings

The cost and reward of a relay depends on its position $s = (x, y)$ in the network. Closer the source and the destination nodes are higher is the reward and smaller is the cost of the corresponding source to destination link. In this section, we assume that the throughput is the reward criteria and the energy consumption is the cost metric. The maximal achieved throughput is $R_{max} = 500$ kbps/RB and the maximum transmitted power is $C_{max} = 250$ mW. The reward (resp. cost) value of a given link between two nodes is inversely proportional (resp. proportional) to the values of the horizontal and vertical divisions d_x and d_y. For the numerical results, we consider the following reward and cost models:

$$r\left(d_x, d_y\right) = \frac{R_{max}}{d_x \times d_y} \tag{5.32}$$

$$c\left(d_x, d_y\right) = \frac{C_{max}}{\left(S_x - d_x + 1\right) + \left(S_y - d_y + 1\right)} \tag{5.33}$$

The cost threshold C_{th} that the average cumulative average cost should not exceed is given by table 5.1. Since throughput is the performance criteria, the reward of the MU-BS link passing through relay i is equal to the half of the min of the throughput of both links MU-relay i and relay i-BS. However, the cost of such link is equal to the transmission power of relay i. Beside computing the cumulative average reward and cost, we evaluate the average cumulative Energy Efficiency (EE) of the proposed algorithms. The average cumulative EE metric of MU n is given by:

$$EE_n = \mathbb{E}\left[\sum_{t=1}^{T} \gamma^{T-t} \frac{R_n\left(\boldsymbol{s}_t, \boldsymbol{a}_t\right)}{C_n\left(\boldsymbol{s}_t, \boldsymbol{a}_t\right)}\right] \tag{5.34}$$

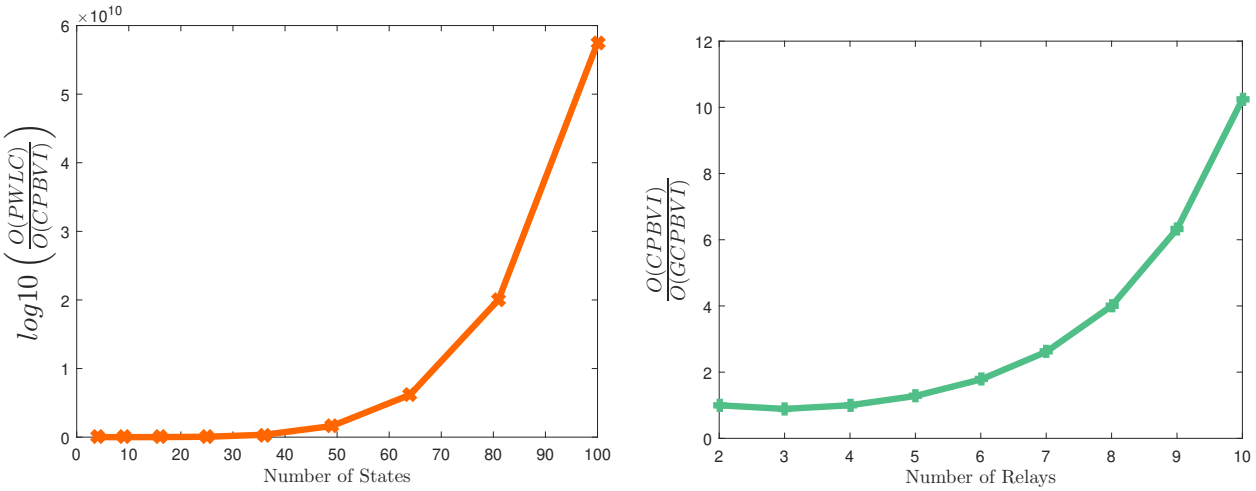
The numerical settings that will commonly be used in the sequel are summarized in the table 5.1.

5.6.2 Single MU scenario

5.6.2.1 GCPBVI performance

Before evaluating the performance of our relay selection policies, we show the motivation behind these approximations. In fact, we illustrate in figure 5.2a and 5.2b how the suggested algorithms highly reduce the computation complexity of the exact solution of the CPOMDP problem. For a single MU scenario and $K = 5$ relays, we plot in figure 5.2a the $\log 10$ of the complexity ratio between the exact solution and the proposed CPBVI solution (i.e. given in 5.4.2) as function of the number of possible regions (i.e. called states). This shows that exact solution for of realistic sizes CPOMDP is unfeasible .

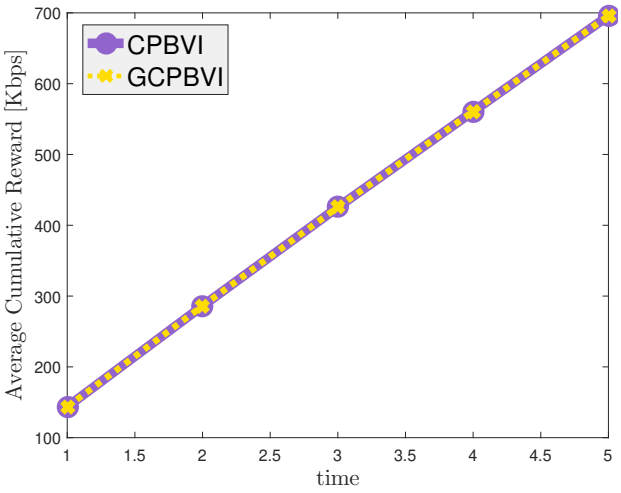
In addition, the exponential number of actions 2^K motivated us to propose a greedy design of the CPBVI algorithm that avoids the study of all the possible actions. We consider $\mathcal{S} = 25$ possible states and we plot in figure 5.2b the complexity ratio between the CPBVI solution and its greedy alternative (i.e. given in 5.4.3) as function of the number of the potential relays K. This figure shows that, even for a small number of potential relays $K = 10$, we can reduce the complexity of a factor of 12.



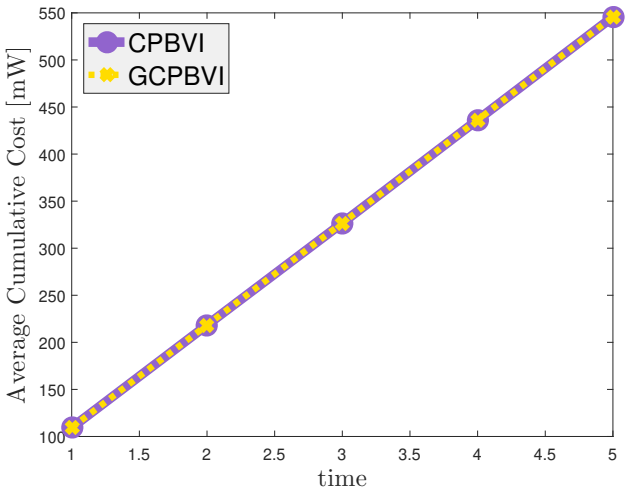
(a) Complexity of CPBVI vs exact solution **(b)** Complexity of GCPBVI vs CPBVI

Figure 5.2: Complexity study of the proposed relay selection policies

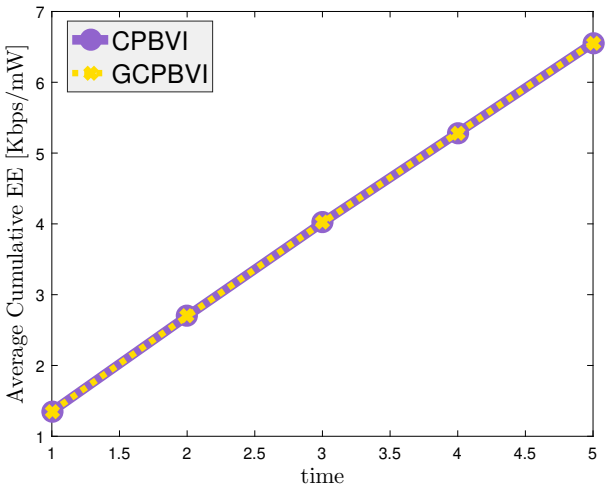
Moreover, we evaluate the performance of the greedy GCPBVI algorithm (given in 5.4.3) compared to the CPBVI scheme (given by 5.4.2). Indeed, the comparison is done in terms of the average cumulative reward, cost and EE of both CPBVI and GCPBVI relay selection solutions. For this comparison, we consider the following simple scenario: single MU with $K = 2$ potential relays and $S_x = 3$ times $S_y = 3$ rectangular regions (i.e. 9 states). Figure 5.3 verifies that the GCPBVI almost return the same average cumulative reward as the CPBVI algorithm. Both solutions have an average cumulative cost that is lower than the cost threshold C_{th} given in 5.1. Therefore, similar average cumulative energy efficiency is deduced for both algorithms.



(a) Average Cumulative Reward



(b) Average Cumulative Cost



(c) Average Cumulative EE

Figure 5.3: Comparison between the performance of following relay selection policies: (i) the CPBVI policy (see 5.4.2) and (ii) the GCPBVI policy (see 5.4.3)

5.6.2.2 D2D relaying performance

Recall that the aim of this work is to enhance the performance of cellular networks. Thus, we consider the throughput as the performance metric to study and show how implementing the proposed relay selection policy can improve the throughput of cellular networks. We consider the single MU scenario with $K = 3$ relays and $|\mathcal{S}| = 16$ regions and we show the results of both scenarios: with and without D2D relaying. For different speeds of state changing v, we plot in figure 5.4 the histogram of the average cumulative reward of both scenarios with and without D2D relaying. In the scenario where D2D is enabled we apply the proposed GCPBVI based relay selection algorithm. The speed v in this figure illustrates the velocity of the relay in moving between the regions. Hence, a speed v is modeled by considering a transition matrix of $\boldsymbol{P}^v$. Figure 5.4 shows that we can gain up to 55% percent in terms of throughput by deploying our policy of relay selection in a cellular network. We note a slow decreasing in the throughput when the speed of the relays increases.

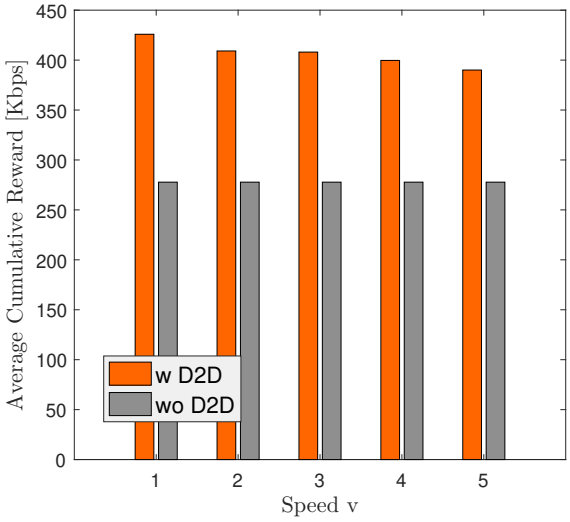


Figure 5.4: For single MU, performance comparison between both scenarios: with and without D2D relaying

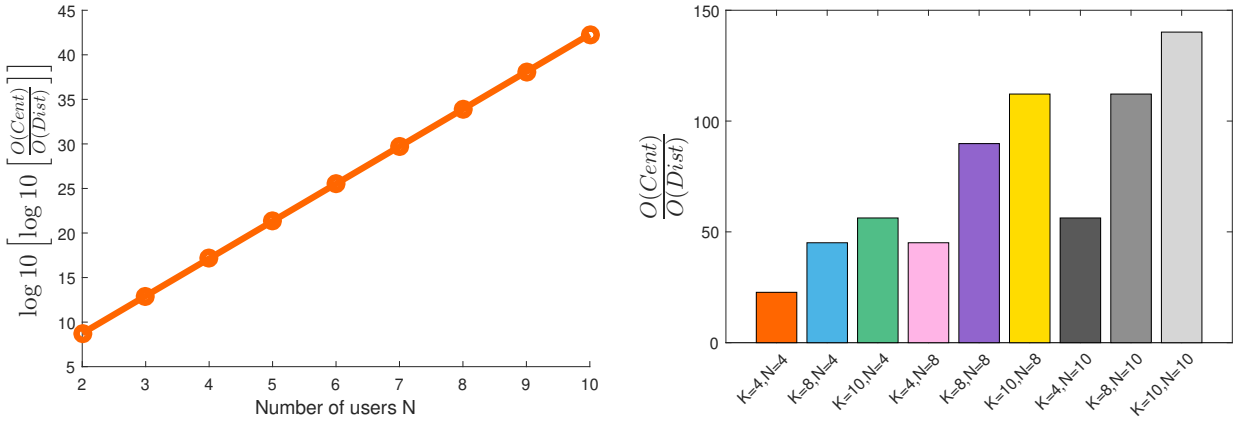
5.6.3 Multi-MU scenario

We consider the case of multiple MUs in the network. Both centralized and distributed relay selection policies were proposed. For evaluating these algorithms we consider $|\mathcal{S}| = 16$ locations in the network and we study the complexity as well as the performance of both distributed and centralized relay selection solutions.

5.6.3.1 The Performance of the Distributed Approach

In figure 5.5, the complexity of the centralized solution for multi-MU scenario (given in 5.5.1) is compared to that of the distributed relay selection solution (given in

5.5.2). The choice of developing a low complexity distributed approach for the multi-MU scenario is validated in figure 5.5. Indeed, the complexity reduction that the distributed approach offers is proportional to both the number of MU N and the number of relays K. As an example, the distributed approach reduces the complexity up to 150 times compared to the centralized one in a realistic scenarios of $K = 10$ relays and $N = 10$ users.



(a) Complexity of centralized vs distributed solutions as function of N

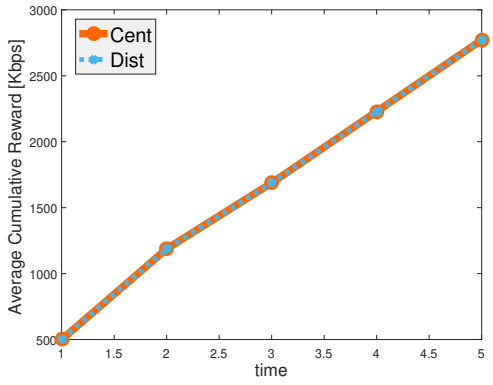
(b) Complexity of centralized vs distributed solutions for different configurations

Figure 5.5: Complexity study of the proposed relay selection policies

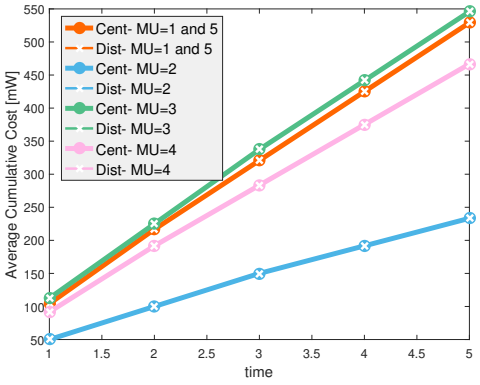
The distributed approach reduces the computational complexity while satisfying a performance close to the centralized one. In figure 5.6, the performance of the distributed approach is compared to the centralized one. Similarly to the single MU case, the comparison is done in terms of the average cumulative reward, cost and EE of both distributed and centralized solutions. For this comparison, a simple scenario of $N = 5$ and $K = 4$ is considered. Figure 5.6 verifies that the low complexity distributed approach almost return the same average cumulative reward as the centralized one. Moreover, both centralized and distributed solutions verify the required cost constraints for each MU (i.e. average cumulative cost lower than the cost threshold C_{th} given in 5.1). In result, similar average cumulative energy efficiency is deduced for both algorithms. We deduce, that applying distributed relay selection induces interesting performance enhancement of the network with a low computational complexity compared to the centralized approach.

5.6.3.2 D2D Relaying Performance

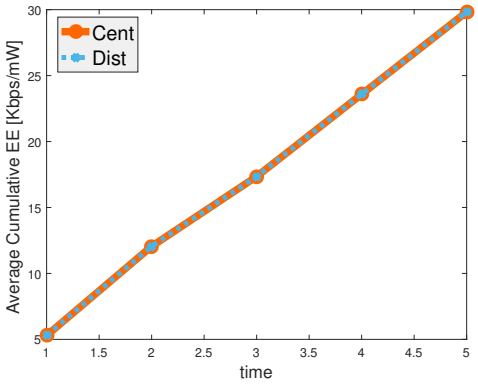
We show how implementing the proposed relay selection policy can improve the throughput of cellular networks. We consider the multiple MUs scenario with $N = 5$, $K = 4$ discovered relays and $|\mathcal{S}| = 16$ regions. For different speeds of state changing v, we plot in figure 5.6d the histogram of the average cumulative reward per MU for both scenarios with and without D2D relaying. In the scenario where D2D is enabled we apply the proposed GCPBVI based relay selection algorithm. The speed v in this



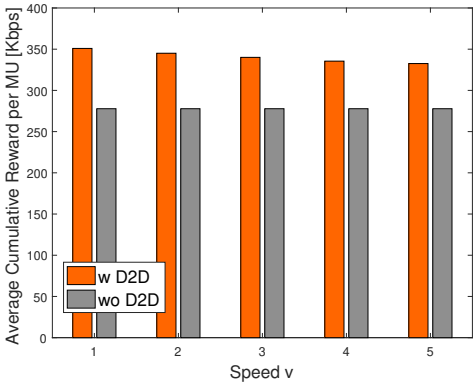
(a) Average Cumulative Reward



(b) Average Cumulative Cost



(c) Average Cumulative EE



(d) For multi MU scenario, performance comparison between both scenarios: with and without D2D relaying

Figure 5.6: Comparison between the performance of the centralized approach (see 5.5.1) and the distributed approach (see 5.5.2) of a multi-MU scenario

figure illustrates the velocity of the relays in moving between the regions. Hence, a speed v is modeled by considering a transition matrix of $\boldsymbol{P}^v$. Each MU applies, in a distributed manner, the relay selection policy proposed in 5.4.3. Figure 5.6d shows that, in average, a MU can gain up to 30% percent of throughput by deploying our policy of relay selection in a cellular network. We note a slow decreasing in the throughput when the speed of the relays increases.

5.7 Simulation Results

During this thesis, we have developed a 3GPP compliant system level simulator that supports DL, UL and D2D communications. We describe the details of the simulations' settings in appendix 7.6. In this section we describe the D2D relaying scenario that has been implemented in the simulator and we evaluate the gain that our proposed relay selection policy achieves in a simulated cellular network.

5.7.1 Scenario

To complete the results of this chapter, we simulate the scenario presented in the flow chart 5.7. The performance of D2D relaying is evaluated separately for UL and DL scenarios. In this report, we will only present the results corresponding to the DL case for lack of space; however, we would like to note that similar results were evaluated for the UL scenario. The initialization phase of the simulations consists of filling the simulation's parameters described in 7.6 and then generating the cellular network (BS, UEs, received power map, traffic generation etc.). We denote by N the number of users and K the number of relays generated in each macro cell of the network. The mobility model of the relays is the one considered in the numerical section 5.6. We use v to denote the speed of the relays. (i.e. transition matrix $\boldsymbol{P}^v$).

The three main blocks of the simulated scenario are the following: (i) **relays discovery** where each user discovers nearby relays, (ii) **relays selection** where each user selects the relays that will ensure the transmission of its traffic (none of the relays is selected if the user found it more beneficial to have traditional cellular communications) and (iii) **transmission** where a user is scheduled in order to transmit its data (i.e. either in a direct way or through relays). With a periodicity of 10 s, each user runs its discovery process from which it deduces nearby devices. The relays move from a region to another (among the 16 possible regions of each cell) with a time-scale of $T_r = 2$ s. The relay selection decision is taken with the same periodicity T_r. Each TTI, the fast fading is updated for all cellular and D2D communications. The allocation of cellular and D2D resources is done at the TTI scale. After data transmission, if the whole user's file is transmitted then a new file is generated according to the FTP2 model described in 7.6.

We denote by T_{simu} the simulation duration of 30000 TTIs and $W = 10$ MHz the bandwidth of the cellular network. The three metrics that are considered for studying the performance of the proposed relay selection policy are the following:

- **CDF of the SINR** which reflexes the coverage area enhancement due to D2D relaying in cellular network.
- If V_{data} [bits] is the amount of data transmitted by all the users then the **first throughput** criteria (i.e. from the network point of view) is computed as follows:
$$Th_1 = \frac{V_{data}}{\bar{N}_{waiting} \times T_{simu} \times W} \tag{5.35}$$
where $\bar{N}_{waiting}$ is the average number of users that are waiting to be served at a given TTI.
- If F is the number of files that have been transmitted with R_i the throughput at which the i^{th} was transmitted, then **the second throughput** criteria (i.e. from the user's point of view) is computed as follows:
$$Th_2 = \frac{\sum_{i=1}^{F} R_i}{F} \tag{5.36}$$

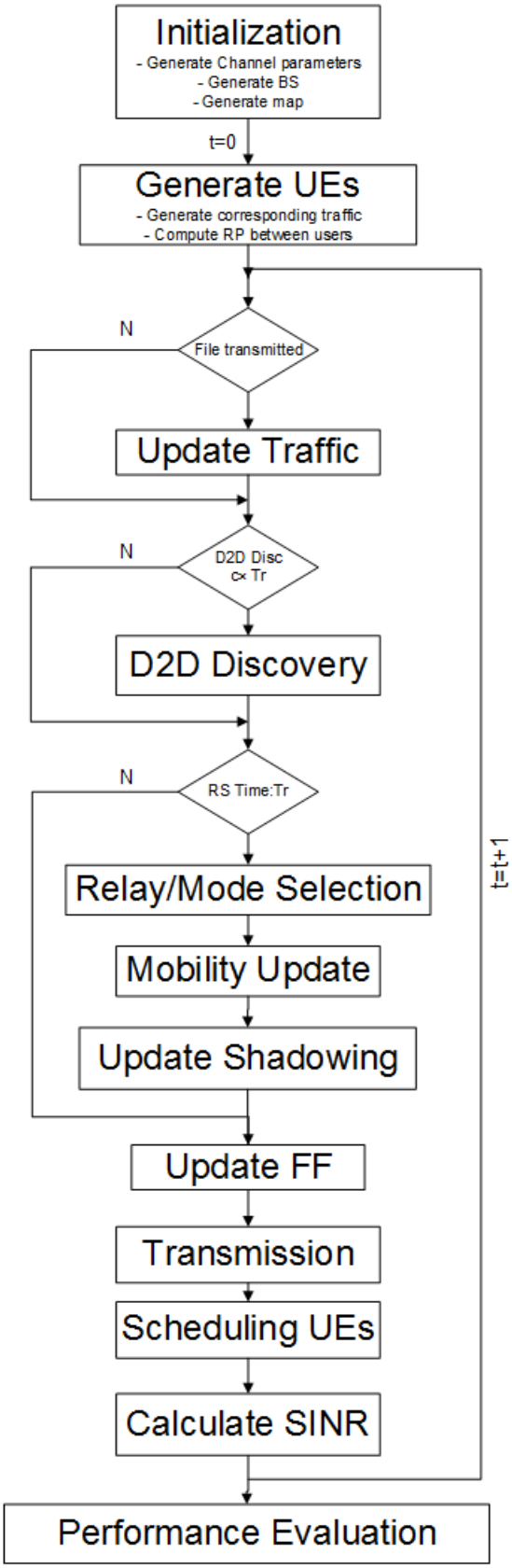
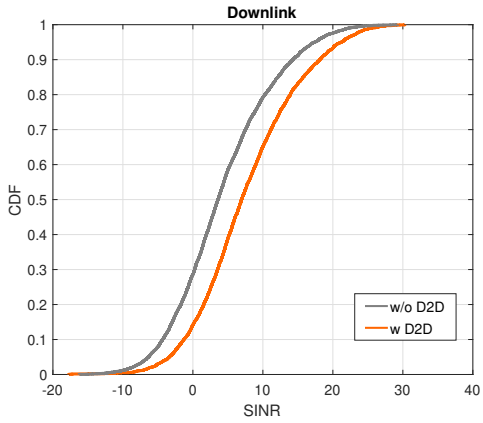
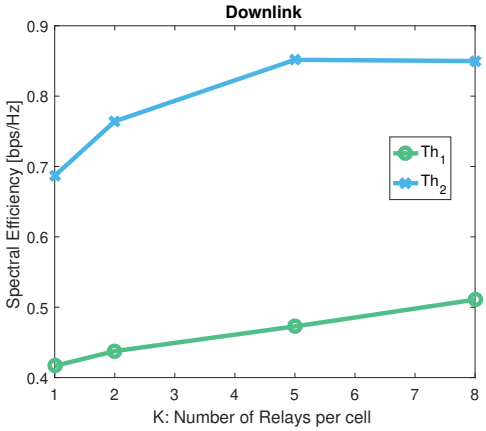


Figure 5.7: Flow Chart of D2D relaying procedures in the system level simulator



(a) CDF of the DL SINR for both scenarios (with and without D2D relaying)



(b) The spectral efficiency of DL communications as function of K

In the following results, the considered reward criteria is the bit-rate which is deduced from the SINR based on the SINR-throughput mapping described in 7.6. In addition, the cost criteria is given by the relays' transmission power. Since a power control mechanism is applied to UL and D2D communications, the relays' transmission power will depend on their channel conditions (i.e. SINR). The cost threshold C_{th} that each user should not exceed is equal to 1 W.

5.7.2 Results

We consider the case of DL communications aided by D2D relaying. The relay selection algorithm implemented is the one described in 5.4.3. The performance of this scenario is given as function of the number of relays K and the number of users N existing in each macro cell. We show in figure 5.8a that enabling D2D relaying improves the SINR of DL communications. Moreover, we study the effect of the number of relays K on the performance of DL communications. For this aim, we consider $N = 5$ users per cell and we plot both throughput metrics (i.e. Th_1 and Th_2 from equations 5.35 and 5.36) as function of the number of relays K (see figure 5.8b). This result illustrates that more the number of relays K increases more the gain achieved by D2D relaying is higher. However, one can see that without considering large number of relays, the D2D relaying can attain important gain in terms of spectral efficiency of DL communications.

In addition, we compare between both scenarios with and without D2D relaying as function of the number N of users dropped in each cell. We consider a scenario with $K = 4$ relays per cell. We show in 5.9 that for different values of $N \in \{2, 4, 8, 10, 12, 14\}$, the spectral efficiency of DL communications increases due to the deployment of our proposed algorithm of relay selection policy. However, in this figure it is not visible for the reader how much the spectral efficiency is increased. Therefore, for different values of N, we present in 5.10 the relative gain achieved by both Th_1 and Th_2 (i.e. given in equations 5.35 and 5.36) due to the implementation of our proposed algorithm of relay selection. A spectral efficiency enhancement of more than 50% is recorded.

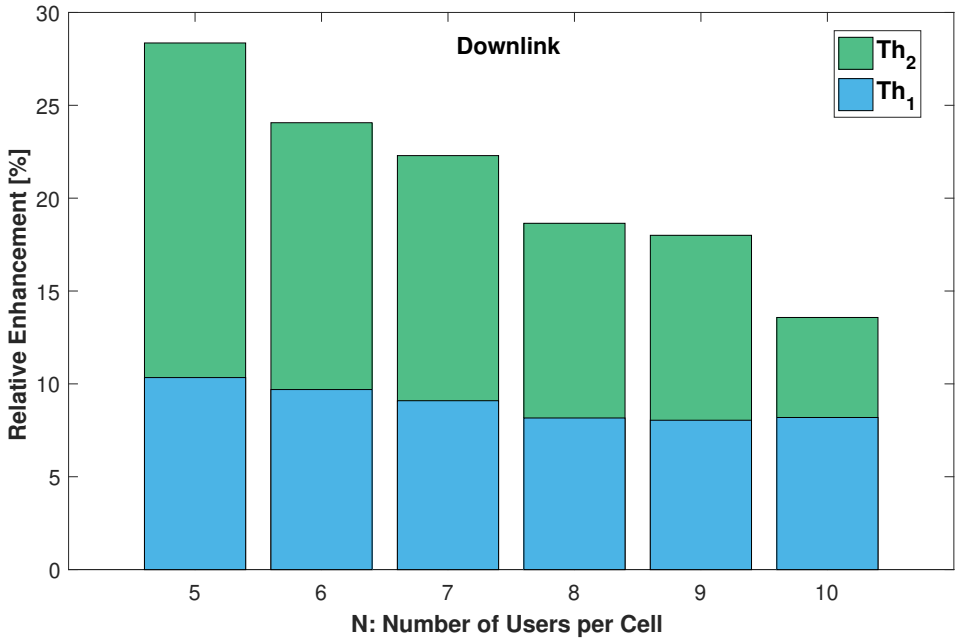


Figure 5.9: The spectral efficiency of DL communications as function of the number of users N for both scenarios

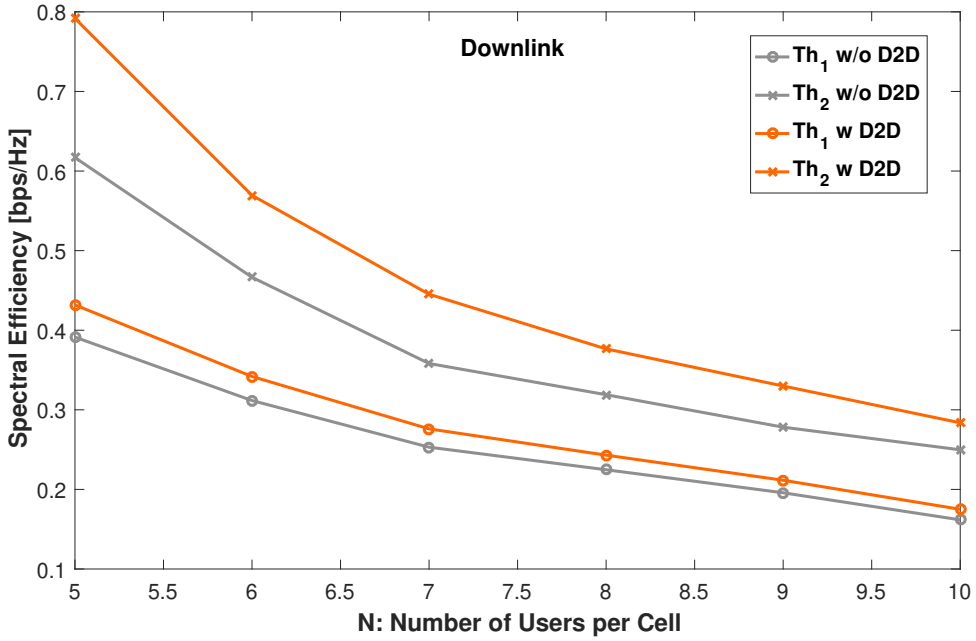


Figure 5.10: The relative enhancement of the spectral efficiency of DL communications as function of the number of users N

5.8 Conclusion

Finding the optimal relay selection policy is a challenging problem especially when the mobility of the relays is considered. In this chapter, we have developed a dynamic relays selection strategy that maximizes a certain performance metric of the cellular networks while guaranteeing some cost constraints. A CPOMDP problem has been formulated and its complexity is discussed. Thus, a greedy GCPBVI algorithm is addressed for achieving a low-complexity and close to optimal solution for the problem. Numerical results show the advantage of such approximation in reducing the problem complexity and prove the performance gain that such solution provides to mobile networks. Our claims are verified by implementing the suggested relay selection scheme in a system-level simulator of cellular networks.

6 | Conclusion and Perspective

6.1 Conclusion

Device-to-Device (D2D) communications that enable direct communication between nearby mobiles is a promising and innovative feature of the evolving Fifth-Generation (5G) cellular networks. However, several technical challenges need further investigations in order to make the implementation of D2D communications a reality. In this thesis, we study the three following problems: D2D vs. cellular mode selection, resource allocation of D2D communications and relay selection in D2D cooperative cellular network.

In chapter 3, we have proposed a queuing analysis of the two following scenarios: (i) cellular scenario where the communications between the users pass through the Base Station (BS) and (ii) D2D scenario where direct communications between the users are enabled via overlay D2D links. We show how this queuing theory approach captures the coupling between the queues and gives more realistic performance evaluation compared to a physical layer approach. We find the expressions of the exact stability regions for both scenarios as convex polytopes with finite number of vertices. In order to reduce the computational complexity of the exact expressions, we provide an approximated model and demonstrate the trade-off that exists between its preciseness and complexity. Based on these results, we derive a distance-based mode selection scheme which defines the cases where D2D communications are more advantageous than cellular ones. Numerical results show that applying this mode selection improves the throughput of cellular communications up to 25%.

In chapter 4, we considerD2D-enabled cellular network with limited feedback and we propose a channel adaptive resource allocation algorithms that minimize the D2D transmission power under throughput constraints. Based on Lyapunov optimization, we start by considering a centralized channel adaptive scheduling that improves the energy efficiency of D2D communications. In this case, D2D users send their channel state and transmission power information to the Base Station (BS). Since limited resources are available for Channel State Information (CSI) transmission, only a subset of D2D users are able to report their CSI at a given time-slot. Thus, the performance of centralized approaches is limited by the resources available for feedback exchange. Therefore, we develop a distributed approach where each D2D user benefits from its local channel state knowledge to share a short CSI indicator that is sufficient for deducing its corresponding energy efficiency metric. Some mechanisms are deployed to

avoid the collision that may occur between these different CSI indicators transmissions. When feedback collision is avoided, this new way of CSI reporting enables all the D2D users to share their CSI indicators at a given time-slot. Thus, in this case, the distributed approach achieves an energy efficiency performance comparable to the ideal solution where the BS has a global CSI knowledge of all the D2D users. Furthermore, we have discussed how both distributed and centralized algorithms can be simply implemented in existing Long Term Evolution (LTE) standards. Numerical results show that the distributed algorithm outperforms the centralized one. Compared to a round-robin scheduling, the proposed algorithms reduce the energy consumption between 70% to 98% and enhance the energy efficiency of D2D networks.

In chapter 5, we consider D2D cooperative cellular networks with mobile relays. We study a dynamic relay selection policy that aims to maximize the performance of cellular networks (e.g. in terms of throughput, reliability, coverage etc.) under cost constraints (e.g. in terms of energy or data consumption etc.). The mobility of the relays is captured by considering a Constrained Partially Observed Markov Decision Process (CPOMDP) formulation of the problem. The complexity of the optimal relay selection policy is discussed. We have proven the submodularity property of the reward and cost value functions in order to deduce a greedy relay selection policy that achieves close to optimal performance with a low computational complexity. Numerical results show that a user can improve its throughput up to 55% due to the aid of appropriately selected relays. The proposed relay selection policy is implemented in a system-level simulator to verify its performance in a realistic network. Simulator shows that D2D relay-aided scenario can enhance the performance of the network up to 55%.

6.2 Perspectives

The mode selection scheme proposed in chapter 3 considers overlay D2D where dedicated resources are allocated for D2D communications. A crucial but challenging extension of this work is to consider underlay D2D where cellular resources are reused by D2D communications. Underlay D2D scenarios improve the spectral efficiency of the network, however interference is generated between cellular and D2D communications which makes the queuing analysis proposed in chapter 3 a challenging problem. Furthermore, the throughput enhancement of this study can be visualized by implementing the proposed mode selection algorithm in the system-level simulator developed during this thesis.

The work in chapter 4 can serve as a baseline for future studies concerning energy efficiency of Multiple Input Multiple Output (MIMO) systems where CSI feedback plays an enormous challenge. In addition, this study proposes both distributed and centralized channel adaptive scheduling algorithms for optimizing the energy efficiency of D2D networks. However, this work can be generalized to any other performance metric that depends on the channel state of D2D links. Indeed, showing the performance of the distributed approach for different performance goals is an interesting extension of this work. Moreover, the performance of this study can be further investigated by implementing the proposed scheduling algorithms in our system-level simulator.

Based on the optimal relay selection strategy found in chapter 5, the following question may arise: what will happen if the mobility pattern of the relays is not communicated to the decision maker (neither the user nor the BS)? In this case, it is interesting to investigate some relay selection policies based on reinforcement learning. In addition, the states of the CPOMDP formulated in this work take only the relays' locations into account. This work can be extended to other parameters such as: queues' length, channel states, and relays' idle state etc. Moreover, online adaptive policies can be studied in order to instantaneously optimize the cellular performance under instantaneous cost constraints (i.e. not be limited to long term average criteria).

Beyond the future works related to this thesis, there still exist several open research questions regarding D2D communications. In particular, one can explore how D2D communications can act as an enabler of upcoming services in future cellular networks. Indeed, D2D communications contribute to the setup of low latency and ultra reliable services (e.g. industry automation) and especially for high speed scenarios (e.g. Vehicle-to-Vehicle (V2V) communications). This type of services involves a wide range of use cases and draws the attention of both industrial and academic researches. Moreover, another domain to examine is how D2D communications can be used for the deployment of unmanned aerial vehicle (UAV) as flying BS. There is a need to optimize this scenario in order to improve the performance of cellular networks in terms of capacity and coverage. Finally, D2D communications for Massive Machine Type Communications (mMTC) applications requires further investigations and especially for reducing the energy consumption of battery-limited devices.

Résumé

Cette thèse étudie les communications directes entre les mobiles, appelées communications D2D, en tant que technique prometteuse pour améliorer les futurs réseaux cellulaires. Cette technologie permet une communication directe entre deux terminaux mobiles sans passer par la station de base. Plusieurs études ont montré les avantages que les communications D2D apportent aux réseaux sans fil tels que: l'amélioration de l'efficacité énergétique et spectrale, l'augmentation du débit, la réduction des délais des communications et l'extension de la couverture du réseau. Bien que les communications D2D aient suscité l'intérêt de l'industrie, du milieu universitaire et des organismes de normalisation, divers problèmes nécessitent plus d'investigation afin d'assurer le déploiement le plus efficace des communications D2D dans les réseaux cellulaires au cours des prochaines années. Les principaux défis techniques rencontrés par les communications D2D sont les suivants : répartition du spectre, sélection de mode de communication, allocation de puissance et/ou des ressources, consommation d'énergie, gestion des interférences, communications D2D inter-opérateurs, confidentialité, authentification et sécurité.

La modélisation, l'évaluation et l'optimisation des différents aspects des communications D2D constituent les objectifs fondamentaux de cette thèse et sont réalisés principalement à l'aide des outils mathématiques suivants: la théorie des files d'attente, l'optimisation de Lyapunov et les Processus de Décision Markovien Partiellement Observable (POMDP). Les résultats de cette étude sont présentés en trois parties.

Dans la première partie, nous étudions un schéma de sélection entre mode cellulaire et mode D2D. Nous dérivons les régions de stabilité des scénarios suivants: réseaux cellulaires purs et réseaux cellulaires où les communications D2D sont activées. Nous fournissons une caractérisation complète de la région de stabilité de ces deux scénarios sous forme de polytopes convexes à nombre fini de sommets. Le couplage entre les files d'attente ainsi que la considération d'un débit adaptative à la qualité des liens D2D rendent la caractérisation de la région de stabilité un problème complexe à résoudre. Ainsi, une représentation approximative est étudiée afin de réduire la complexité de ce calcul. Une comparaison entre les deux scénarios cellulaire et D2D conduit à l'élaboration d'un algorithme de sélection entre le mode cellulaire et le mode D2D dont le but est d'améliorer la capacité du réseau.

Dans la deuxième partie, nous développons un algorithme d'allocation de ressources des communications D2D. En se basant sur l'optimisation de Lyapunov, nous présentons un algorithme centralisé et adaptatif à la qualité du canal dont l'objectif est de minimiser la puissance de transmission des utilisateurs D2D sous des contraintes de débit. Cette approche centralisée est limitée par le nombre de ressources disponibles pour les transmissions des informations liées à la qualité des liens D2D. Ainsi, nous proposons un algorithme distribué d'ordonnancement qui profite de la connaissance des utilisateurs de leur propre état de canal. L'idée principale est que les utilisateurs calculent localement leurs propres métriques d'efficacité énergétique et puis utilisent un nouveau système de signalisation pour partager cette information entre eux. En revanche, ce modèle de partage de connaissance déclenche une certaine collision entre les différents messages de signalisation. Certains mécanismes sont développés pour éviter ces collisions et amener à ce que l'efficacité énergétique de l'algorithme distribué dépasse celle de l'algorithme centralisé.

Dans la troisième partie, nous étudions une politique de sélection des relais D2D mobiles. La mobilité des relais représente un des défis que confronte toute stratégie de sélection de relais. L'objectif de la politique que l'on propose est d'optimiser les performances des réseaux cellulaires (ex : le débit, la fiabilité, la couverture) sous des contraintes de coût (ex : la puissance de transmission, le budget de puissance). Nous représentions la dynamique des relais par un processus de décision markovien où l'emplacement de chaque relais découverts ne peut être observé que si ce dernier avait été sélectionné. Par conséquent, le processus de décision de relais séquentiel sous contraintes de coût est plutôt modélisé par un processus contraint de décision markovien partiellement observable (CPOMDP). Comme les solutions exactes de ce type de problème sont difficiles à trouver, nous suggérons une solution approximative dont la complexité et la précision sont discutées.

Mots Clés: communications directes entre mobiles D2D, réseaux cellulaires, sélection de mode de communication, allocation de ressources, sélection des relais, théorie des files d'attente, optimisation Lyapunov, processus de décision markovien partiellement observable POMDP.

Appendices

7.3 D2D Mode Selection Proofs

7.3.1 Proof of theorem 3.3.1

We find the exact stability region for the three-UEs cellular scenario. The main challenge is to solve the complicated Markov Chain that models the queue Q_{BS} in order to find the probability that this queue is empty and deduce the service rate of the queues in the systems. The coupling between the queues leads to a multidimensional Markov Chain modeling of the system of queues. Since this model complicates the study of the stability region, we approach this problem in a different way. We use the priority policies definition (in section 3.2) to transform the multidimensional model to a 1D Markov Chain (one dimensional). Hence, for a given priority policy, each queue can be modeled by a 1D Markov Chain. In order to characterize the stability region, all the priority policies should be considered or at least the priority policies that achieve the corner points of the stability region.

We recall the three-UEs cellular scenario contains one *UE2UE communication* and one *UE2BS communication*. The *UE2UE communication* is modeled by the cascade of Q_s and Q_{BS} and the *UE2BS communication* is modeled by the queue Q_u. We consider the set of three rates $\{r_1, r_2, r_3\}$ with $r_1 = kr_2$ and $r_3 = 0$. So, if P is the number of packets transmitted at r_2 then kP is the number of packets transmitted at r_1. Assuming that the two sources queues Q_s and Q_u are saturated, we want to characterize the stability region.

The set of priority policies Ω_{Γ}^{ss} that consists of the corner points of the stability region is given by $\Omega_{\Gamma}^{ss} = \{\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4, \Gamma_5, \Gamma_6\}$ and is detailed in 3.3.3. For the cellular scenario, we denote by $q_n(S_i, S_j, S_k, \Gamma)$ the probability to transmit over the link_n (with $n \in \{s, d, u\}$) for a given policy Γ and a given state of the links: link_s at state S_s, link_d is at state S_d and link_u is at state S_u (with S_s, S_d and $S_u \in (S_1, S_2, S_3)$). Based on a simple analysis of the priority policies $\Gamma \in \Omega_{\Gamma}^{ss}$, these probabilities can be easily generated. Let us consider an example to clarify the procedure. Note that $q_s(S_i, S_j, S_k, \Gamma_3) = q_d(S_i, S_j, S_k, \Gamma_3) = q_u(S_i, S_j, S_k, \Gamma_3) = 0$ for all the the other combinations of states (S_i, S_j, S_k) not shown in the table. By analogy, the transmission probabilities in the cellular scenario for all $\Gamma \in \Omega_{\Gamma}^{ss}$ are computed.

(Step 1) Markov chain model of Q_{BS}: Cellular communications are modeled as coupled processor sharing queues where the service rates of Q_s (equivalent to the arrival rate of Q_{BS}) as well as the service rate of Q_u depend on the state (empty or not) of Q_{BS}. Let us study the Markov chain of the queue Q_{BS} in order to deduce the probability of being empty which means the probability of having at least one packet.

Q_{BS} corresponds to the queue at the BS side that can be modeled as a Markov chain with a transition probability from a state x_i to a state x_{i+1} is equal to the probability of receiving P packets per slot which is equivalent to the transmission probability of the queue Q_s at rate r_2 due to the fact that the packets arriving to Q_{BS} correspond to the packets transmitted by UE_s. Thus, the transition probability from a state x_i to a state x_{i+k} is equal to the probability of receiving kP packets per slot which is equivalent to the transmission probability of the queue Q_s at rate r_1. Moreover, the transition probability from a state x_i to a state x_{i-1} is equal to the transmission probability at the DL BS-UE_d at rate r_2 and the transition probability from a state x_i to a state x_{i-k} is equal to the transmission probability at the DL BS-UE_d at rate r_1. As we mentioned before, the scheduling decision depends on the state (empty or not) of Q_{BS} which means that the behavior of Q_s and Q_u is coupled to the state of Q_{BS} and the arrival probability when Q_{BS} is empty (state 0) is not equal to the arrival probability when Q_{BS} is not empty (state $\neq 0$). In figure 7.1, we present the discrete time Markov Chain with infinite states that describes the evolution of the queue Q_{BS}.

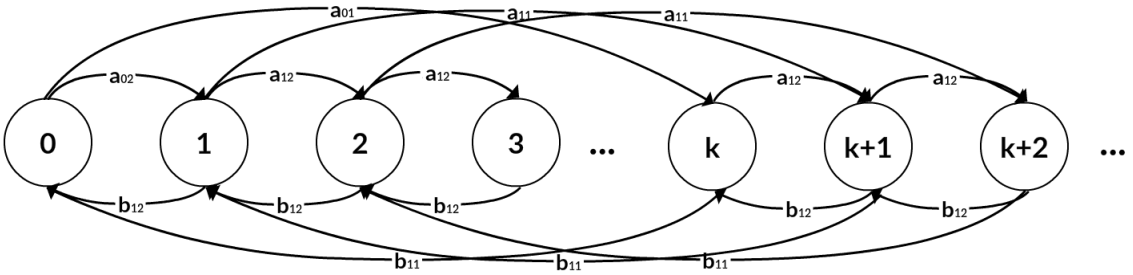


Figure 7.1: Markov Chain model of the queue Q_{BS}

We start by studying the queue Q_{BS} in terms of probability of transmitting and probability of receiving packets in order to deduce the probability that the queue Q_{BS} is empty.

(Step 1.a) Service probability of Q_{BS}: For Q_{BS}, two service probabilities exist the first one b_{11} corresponds to the probability of transmitting at rate r_1 and the second one b_{12} corresponds to the probability of transmitting at rate r_2.

$$b_{11}(\Gamma) = p_d^1 \sum_{i,j=1}^{3} p_s^i \, p_u^j \, q_d(S_i, S_j, S_1, \Gamma) \text{ and } b_{12}(\Gamma) = p_d^2 \sum_{i,j=1}^{3} p_s^i \, p_u^j \, q_d(S_i, S_j, S_2, \Gamma)$$

Parameters U, V, W', X', Y', Z', M and N from table 7.1 are used to make expressions simpler. These parameters depend on the considered priority policy Γ and their values are deduced from that of the transmission probabilities. The values of these parameters are specified in table 3.1 for all $\Gamma \in \Omega_\Gamma^{ss}$. Hence, the service probabilities

b_{11} and b_{12} of the queue Q_{BS} can be written as follows:

$$b_{11} = p_d^1 \left(M - p_s^1 W' - p_s^2 Y'\right) = p_d^1 \left(1 - \alpha_1 p_s^1 - \alpha_3 p_s^2\right) U \tag{7.1}$$

$$b_{12} = p_d^2 \left(N - p_s^1 X' - p_s^2 Z'\right) = p_d^2 \left(p_s^1 U + \bar{p}_s^1 V - \alpha_2 p_s^1 U - \alpha_4 p_s^2 V\right) \tag{7.2}$$

For each priority policy $\Gamma \in \Omega_\Gamma^{ss}$, we find the transmission probabilities of all the links (s, d and u) then we compute the parameters U, V, W', X', Y', Z', M and N using table 7.1.

(Step 1.b) Arrival probability of Q_{BS}: Due to the coupling between the queues Q_s and Q_u with the state (empty or not) of Q_{BS}, we should distinguish between (i) a_{01} and a_{02} that denote the arrival probabilities respectively at rate r_1 and r_2 when Q_{BS} is empty and (ii) a_{11} and a_{12} that denote the arrival probabilities respectively at rate r_1 and r_2 when Q_{BS} is not empty. For a given priority policy Γ, these probabilities are given by the following expressions with parameters U and V from table 7.1:

$$a_{01} = p_s^1 \sum_{j,k=1}^{3} p_u^j p_d^k q_s(S_1, S_j, S_k, \Gamma | Q_{BS}(t) = 0) = p_s^1 U \tag{7.3}$$

$$a_{02} = p_s^2 \sum_{j,k=1}^{3} p_u^j p_d^k q_s(S_2, S_j, S_k, \Gamma | Q_{BS}(t) = 0) = p_s^2 V \tag{7.4}$$

$$a_{11} = p_s^1 \sum_{j,k=1}^{3} p_u^j p_d^k q_s(S_1, S_j, S_k, \Gamma | Q_{BS}(t) > 0) = p_s^1 \left(\alpha_1 p_d^1 + \alpha_2 p_d^2 + p_d^3\right) U \tag{7.5}$$

$$a_{12} = p_s^2 \sum_{j,k=1}^{3} p_u^j p_d^k q_s(S_2, S_j, S_k, \Gamma | Q_{BS}(t) > 0) = p_s^2 \left(\alpha_3 p_d^1 U + \alpha_4 p_d^2 V + p_d^3 V\right) \tag{7.6}$$

Notation	Expression	Notation	Expression
U	$\sum_{i=1}^{3} p_u^i q_s(S_1, S_i, -, \Gamma \| Q_{BS}(t) = 0)$	V	$\sum_{i=1}^{3} p_u^i q_s(S_2, S_i, -, \Gamma \| Q_{BS}(t) = 0)$
W'	$\sum_{i=1}^{3} p_u^i q_s(S_1, S_i, S_1, \Gamma \| Q_{BS}(t) > 0) = \alpha_1 U$	X'	$\sum_{i=1}^{3} p_u^i q_s(S_1, S_i, S_2, \Gamma \| Q_{BS}(t) > 0) = \alpha_2 U$
Y'	$\sum_{i=1}^{3} p_u^i q_s(S_2, S_i, S_1, \Gamma \| Q_{BS}(t) > 0) = \alpha_3 U$	Z'	$\sum_{i=1}^{3} p_u^i q_s(S_2, S_i, S_2, \Gamma \| Q_{BS}(t) > 0) = \alpha_4 V$
W	$\sum_{i,j=1}^{3} \sum_{i=1}^{3} p_s^i p_d^j q_u(S_i, S_1, S_j, \Gamma \| Q_{BS}(t) = 0)$	X	$\sum_{i,j=1}^{3} \sum_{i=1}^{3} p_s^i p_d^j q_u(S_i, S_2, S_j, \Gamma \| Q_{BS}(t) = 0)$
Y	$\sum_{i,j=1}^{3} \sum_{i=1}^{3} p_s^i p_d^j q_u(S_i, S_1, S_j, \Gamma \| Q_{BS}(t) > 0)$	Z	$\sum_{i,j=1}^{3} \sum_{i=1}^{3} p_s^i p_d^j q_u(S_i, S_2, S_j \| Q_{BS}(t) > 0$
A	$\sum_{i=1}^{3} p_u^i q_s^d(S_1, S_i, \Gamma) = U$	B	$\sum_{i=1}^{3} p_u^i q_s^d(S_2, S_i, \Gamma) = V$
C	$\sum_{i=1}^{3} p_s^i q_u^d(S_i, S_1, \Gamma)$	D	$\sum_{i=1}^{3} p_s^i q_u^d(S_i, S_2, \Gamma)$
M	$p_s^1 U + p_s^2 U + p_s^3 U = U$	N	$p_s^1 U + p_s^2 V + p_s^3 V = p_s^1 U + \bar{p}_s^1 V$

Table 7.1: Notation to make expressions simpler

(Step 2) Probability that Q_{BS} empty Let us consider $a_0 = a_{01} + a_{02}$, $a_1 = a_{11} + a_{12}$ and $b_1 = b_{11} + b_{12}$. We solve the balance equations in order to find the stationary distribution Π of the Markov chain of Q_{BS} and more specifically the probability that Q_{BS} is empty Π_0. Recall that the stationary distribution of a Markov Chain exists if and only if the stability condition is verified. Hence, finding the stationary distribution based on the balance equations is a necessary and sufficient condition for guaranteeing the stability of Q_{BS}. The balance equations at each state i of the Markov chain (figure 7.1) are given by:

$$\text{For } i = 0: \quad \Pi_0 a_0 = \Pi_1 b_{12} + \Pi_k b_{11} \tag{7.7}$$

$$\text{For } i = 1: \quad \Pi_1 (a_1 + b_{12}) = \Pi_0 a_{02} + \Pi_2 b_{12} + \Pi_{k+1} b_{11} \tag{7.8}$$

$$\text{For } 2 \leq i \leq k-1: \quad \Pi_i (a_1 + b_1) = \Pi_{i-1} a_{12} + \Pi_{i+1} b_{12} + \Pi_{i+k} b_{11} \tag{7.9}$$

$$\text{For } i = k: \quad \Pi_k (a_1 + b_1) = \Pi_0 a_{01} + \Pi_{k-1} a_{12} + \Pi_{k+1} b_{12} + \Pi_{2k} b_{11} \tag{7.10}$$

$$\text{For } i > k: \quad \Pi_i (a_1 + b_1) = \Pi_{i-k} a_{11} + \Pi_{i-1} a_{12} + \Pi_{i+1} b_{12} + \Pi_{i+k} b_{11} \tag{7.11}$$

We look for a solution of the balance equations of the form $\Pi_n = x^n$. Then we construct a linear combination of these solutions (which also satisfies the boundary equations (7.7), (7.8), (7.9)) that satisfies the normalization equation $\sum_n \Pi_n = 1$. We start by substituting Π_n by x^n in (7.11) and then dividing by x^{n-k}. This yields the following polynomial equation:

$$P(x) = b_{11}x^{2k} + b_{12}x^{k+1} - (a_{12} + a_{11} + b_{12} + b_{11})\, x^k + a_{12}x^{k-1} + a_{11}$$

$$= (x-1)\left(b_{11} \sum_{i=k+1}^{2k-1} x^i + (b_{11} + b_{12})\, x^k - (a_{11} + a_{12})\, x^{k-1} - a_{11} \sum_{i=0}^{k-1} x^i \right) \tag{7.12}$$

The first root $x = 1$ but this is not a useful one, since we must be able to normalize the solution of the equilibrium equations. Stability conditions require that this polynomial has at least one root x with $|x| < 1$. Let us say we found R roots such that $|x| < 1$: $x_1, x_2, x_3 ... x_R$ with $1 \leq R \leq 2k-1$ (degree of the polynomial equation). Then the stationary distribution is given by the following linear combination:

$$\Pi_n = \sum_{i=1}^{R} c_i x_i^n \quad n = k+1, k+2 \tag{7.13}$$

Where Π_n for $0 \leq n \leq k$ as well as the coefficients c_i for $0 \leq i \leq R$ are computed by solving the following system of equations: (i) first $R + K$ balance equations and (ii) the normalization equation: $\sum_n \Pi_n = 1$.

For each choice of the coefficients c_k the linear combination satisfies (7.11). These coefficients add some freedom that can be used to also satisfy the initial balance equations for $0 \leq i \leq k$ as well as the normalization equation. Therefore, we have $R + k + 1$ unknowns to determine in order to find the stationary distribution of the queue Q_{BS}, these unknowns are: the R coefficients c_i and the $k + 1$ probabilities

Π_i for $0 \leq i \leq k$. These unknowns are found by solving the following linear system $AX = B$ with $X^T = [\Pi_0, \Pi_1, \Pi_2, ..., \Pi_{k-1}, \Pi_{k-1}, c_1, c_2, c_3..., c_{R-1}, c_R]$ (Note that X^T is the transpose of the vector X). By substituting (7.13) into the first $R + k - 1$ balance equation as well as the normalization equation, we deduce for $k > 2$ the matrix A and B of the linear system as follows. The special case of $k = 2$ is treated afterward.

$$A_0 = \begin{bmatrix}
State & \Pi_0 & \Pi_1 & \Pi_2 & \Pi_3 & \Pi_4 & ... & \Pi_{k-2} & \Pi_{k-1} & \Pi_k \\
i=0 & -a_0 & b_{12} & 0 & 0 & 0 & ... & 0 & 0 & b_{11} \\
i=1 & a_{02} & -(a_1+b_{12}) & b_{12} & 0 & 0 & ... & 0 & 0 & 0 \\
i=2 & 0 & a_{12} & -(a_1+b_{12}) & b_{12} & 0 & ... & 0 & 0 & 0 \\
... & 0 & 0 & a_{12} & -(a_1+b_{12}) & b_{12} & 0... & 0 & 0 & 0 \\
 & & & & & & ... & & & \\
i=k-1 & 0 & 0 & 0 & 0 & 0 & ... & a_{12} & -(a_1+b_{12}) & b_{12} \\
i=k & a_{0k} & 0 & 0 & 0 & 0 & ... & 0 & a_{12} & -(a_1+b_1) \\
i=k+1 & 0 & a_{1k} & 0 & 0 & 0 & ... & 0 & 0 & a_{12} \\
i=k+2 & 0 & 0 & a_{1k} & 0 & 0 & ... & 0 & 0 & 0 \\
i=k+3 & 0 & 0 & 0 & a_{1k} & 0 & ... & 0 & 0 & 0 \\
 & & & & & & ... & & & \\
i=2k & 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 & a_{1k} \\
i=2k+1 & 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 & 0 \\
... & 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 & 0 \\
i=R+k-1 & 0 & 0 & 0 & 0 & 0 & ... & 0 & 0 & 0 \\
Norm & 1 & 1 & 1 & 1 & 1 & ... & 1 & 1 & 1
\end{bmatrix}$$

$$A_i = \begin{bmatrix}
c_i \\
0 \\
b_{11}x_i^{k+1} \\
b_{11}x_i^{k+2} \\
b_{11}x_i^{k+3} \\
... \\
b_{11}x_i^{2k-1} \\
b_{12}x_i^{k+1} + b_{11}x_i^{2k} \\
-(a_1+b_1)\,x_i^{k+1} + b_{12}x_i^{k+2} + b_{11}x_i^{2k+1} \\
a_{12}x_i^{k+1} - (a_1+b_1)\,x_i^{k+2} + b_{12}x_i^{k+3} + b_{11}x_i^{2k+2} \\
a_{12}x_i^{k+2} - (a_1+b_1)\,x_i^{k+3} + b_{12}x_i^{k+4} + b_{11}x_i^{2k+3} \\
... \\
a_{12}x_i^{2k-1} - (a_1+b_1)\,x_i^{2k} + b_{12}x_i^{2k+1} + b_{11}x_i^{3k} \\
a_{1k}x_i^{k+1} + a_{12}x_i^{2k} - (a_1+b_1)\,x_i^{2k+1} + b_{12}x_i^{2k+2} + b_{11}x_i^{3k+1} \\
a_{1k}x_i^{k+2} + a_{12}x_i^{2k+1} - (a_1+b_1)\,x_i^{2k+2} + b_{12}x_i^{2k+3} + b_{11}x_i^{3k+2} \\
a_{1k}x_i^{R-1} + a_{12}x_i^{R+k-2} - (a_1+b_1)\,x_i^{R+k-1} + b_{12}x_i^{R+k} + b_{11}x_i^{R+2k} \\
\frac{1}{1-x_i} - \sum_{j=0}^{k} x_i^j
\end{bmatrix} ; \forall 1 \leq i \leq R$$

$$A = \begin{bmatrix} A_0 & A_1 & A_2 & ... & A_{R-1} & A_R \end{bmatrix}$$

$$B^T = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0...0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus, solving the linear system $AX = B$ with success induces two results: (i) stability of the queue Q_{BS} and (ii) the stationary distribution of Q_{BS} and especially what interests us is the probability that this queue is empty Π_0.

The result above holds for any $k \in \mathbb{N}^+$. However, we show the result for the case ($r_1 = 2r_2 \Rightarrow k = 2$) where $k - 1 = 1$ hence the case $2 \leq i \leq k - 1$ should not be taken into account for this special case. Thus, the set of balance equations is the following: For $i = 0$:

$$\Pi_0 a_0 = \Pi_1 b_{12} + \Pi_2 b_{11} \tag{7.14}$$

For $i = 1$:

$$\Pi_1 (a_1 + b_{12}) = \Pi_0 a_{02} + \Pi_2 b_{12} + \Pi_3 b_{11} \tag{7.15}$$

For $i = 2$:

$$\Pi_2 (a_1 + b_1) = \Pi_0 a_{01} + \Pi_1 a_{12} + \Pi_3 b_{12} + \Pi_4 b_{11} \tag{7.16}$$

For $i \geq 3$:

$$\Pi_i (a_1 + b_1) = \Pi_{i-2} a_{11} + \Pi_{i-1} a_{12} + \Pi_{i+1} b_{12} + \Pi_{i+2} b_{11} \tag{7.17}$$

By analogy to the general case of $r_1 = kr_2$, we look for a solution of the balance equations of the form $\Pi_{BS}(n) = x^n$ and then we construct a linear combination of these solutions also satisfying the boundary equations (7.14), (7.15), (7.16) as well as the normalization equation $\sum_n \Pi_{BS}(n) = 1$. Substituting $\Pi_{BS}(n)$ by x^n into (7.17) and then dividing by x^{n-2} yield the following polynomial equation:

$$\begin{aligned} P(x) &= b_{11}x^4 + b_{12}x^3 - (a_{12} + a_{11} + b_{12} + b_{11})\, x^2 + a_{12}x + a_{11} \\ &= (x - 1)\left(b_{11}x^3 + (b_{11} + b_{12})\, x^2 - (a_{11} + a_{12})\, x - a_{11}\right) \end{aligned} \tag{7.18}$$

Recall that the root $x = 1$ is not a useful one, since we must be able to normalize the solution of the equilibrium equations. Stability conditions require that this polynomial has at least one root x with $|x| < 1$ say x_1, x_2, x_3. We now consider the stationary distribution given by the following linear combination:

$$\Pi_n = \sum_{k=1}^{3} c_k x_k^n \quad n = 3, 4... \tag{7.19}$$

and Π_n for $= n = 0, 1, 2$ are given by the boundary equations. For each choice of the coefficients c_k the linear combination satisfies (7.17). These coefficients add some freedom that can be used to also satisfy the equations (7.14), (7.15), (7.16) and the normalization equation. Substituting of (7.19) into equations (7.14), (7.15), (7.16) we deduce Π_0, Π_1, Π_2 as function of c_1,c_1 and c_3. Then substituting of (7.19) for $n > 3$ as well as Π_0, Π_1, Π_2 as function of c_1,c_1 and c_3 into the normalization equation and the balance equations (7.17) for $n = 3, 4$ yield a set of 3 linear equations for 3 unknowns coefficients c_1,c_1 and c_3.

It can be shown that the unknowns Π_0, Π_1 and Π_2, c_1, c_2, c_3 are found by solving the following linear problem $AX = B$ such that:

$$A = \begin{bmatrix} \Pi_0 & \Pi_1 & \Pi_2 & c_1 & ... & c_3 \\ -a_0 & b_{12} & b_{11} & 0 & ... & 0 \\ a_{02} & -(a_1+b_1) & b_{12} & b_{11}x_1^3 & ... & b_{11}x_3^3 \\ a_{01} & a_{12} & -(a_1+b_1) & b_{12}x_1^3 + b_{11}x_1^4 & ... & b_{12}x_3^3 + b_{11}x_3^4 \\ 0 & a_{11} & a_{12} & -(a_1+b_1)x_1^3 + b_{12}x_1^4 + b_{11}x_1^5 & ... & -(a_1+b_1)x_3^3 + b_{12}x_3^4 + b_{11}x_3^5 \\ 0 & 0 & a_{11} & a_{12}x_1^3 - (a_1+b_1)x_1^4 + b_{12}x_1^5 + b_{11}x_1^6 & ... & a_{12}x_3^3 - (a_1+b_1)x_3^4 + b_{12}x_3^5 + b_{11}x_3^6 \\ 1 & 1 & 1 & \frac{1}{1-x_1} - 1 - x_1 - x_1^2 & ... & \frac{1}{1-x_3} - 1 - x_3 - x_3^2 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Solving the system of linear equations given by $AX = B$ gives the values of Π_i for $0 \leq i \leq k-1$ and of the coefficients c_j for $1 \leq j \leq R$. Therefore, we find the probability that the queue BS is empty which is Π_0.

(Step 3) Service rates of Q_s and Q_u: We follow the procedure below, based on queuing theory analysis of the network capacity, in order to derive the stability region of the cellular scenario. For the 3-UEs scenario, the stability region is characterized by computing the two service rates μ_s and μ_u of the *UE2UE* and *UE2BS communications*.

(Step 3.a) Service rate of Q_s: Let us start with the service rate of the *UE2UE communications* μ_s. If Q_{BS} is empty then the service rate of Q_s is denoted by μ_s^0, otherwise it is denoted by μ_s^1. Hence, the average service rate of Q_s for a given priority policy Γ is computed as follows:

$$\mu_s(\alpha, \Gamma) = \mathbb{P}[Q_{BS} = 0]\,\mu_s^0(\Gamma) + \mathbb{P}[Q_{BS} > 0]\,\mu_s^1(\alpha, \Gamma) = \Pi_0 \mu_s^0(\Gamma) + (1 - \Pi_0)\,\mu_s^1(\alpha, \Gamma) \tag{7.20}$$

with $\mu_s^0(\Gamma)$ and $\mu_s^1(\alpha, \Gamma)$ given by:

$$\mu_s^0(\Gamma) = \mathbb{E}[\mu_s(\Gamma, Q|Q_{BS} = 0)] = \sum_{i=1}^{3} r_i p_s^i \sum_{j=1}^{3} p_u^j q_s(S_i, S_j, -, \Gamma|Q_{BS} = 0)$$

$$= r_1 p_s^1 U + r_2 p_s^2 V \tag{7.21}$$

$$\mu_s^1(\alpha, \Gamma) = \mathbb{E}[\mu_s(\Gamma, Q|Q_{BS} > 0)] = \sum_{i=1}^{3} r_i p_s^i \sum_{j,k=1}^{3} p_u^j p_d^k q_s(S_i, S_j, S_k, \Gamma|Q_{BS} > 0)$$

$$= r_1 p_s^1 \left(\alpha_1 p_d^1 + \alpha_2 p_d^2 + p_d^3\right) U + r_2 p_s^2 \left(\alpha_3 p_d^1 U + \alpha_4 p_d^2 U + p_d^3 V\right) \tag{7.22}$$

Where the expressions of U and V depend on the considered priority policy Γ and their values for the six priority policies $\Gamma \in \Omega_\Gamma^{ss}$ are specified in table 3.1.

For each priority policy $\Gamma \in \Omega_\Gamma^{ss}$, we compute the probability that Q_{BS} is empty $\Pi_0(\alpha, \Gamma)$ from *step 2* then we substitute the above expressions of $\mu_s^0(\Gamma)$ and $\mu_s^1(\alpha, \Gamma)$ into the equation (7.20), we obtain (3.8) as the service rate $\mu_s(\alpha, \Gamma)$ of the queue Q_s.

(Step 3.b) Service rate of Q_u: After computing μ_s, let us compute the second element of the stability region which is the service rate μ_u of the *UE2UE communication*. If Q_{BS} is empty then the service rate of Q_u is denoted by μ_u^0 otherwise it is denoted by μ_u^1. By analogy to the Q_s, the service rate of Q_u is computed as follows:

$$\mu_u(\Gamma) = \Pi_0 \mu_u^0(\Gamma) + (1 - \Pi_0)\, \mu_u^1(\alpha, \Gamma) \tag{7.23}$$

with $\mu_u^0(\Gamma)$ and $\mu_u^1(\alpha, \Gamma)$ given by:

$$\mu_u^0(\Gamma) = \mathbb{E}\left[\mu_u(S_i, S_j, S_k, \Gamma\,, |Q_{BS} = 0)\right] = \sum_{j=1}^{3} r_j p_u^j \sum_{i=1}^{3} p_s^i q_u(S_i, S_j, -, \Gamma | Q_{BS} = 0)$$

$$= r_1 p_u^1 W + r_2 p_u^2 X \tag{7.24}$$

$$\mu_u^1(\alpha, \Gamma) = \mathbb{E}\left[\mu_u(S_i, S_j, S_k, \Gamma\,, Q | Q_{BS} > 0)\right] = \textstyle\sum_{j=1}^{3} r_j p_u^j \sum\limits_{i,k=1}^{3} p_s^i p_d^k q_u(S_i, S_j, S_k, \Gamma | Q_{BS} > 0)$$

$$= r_1 p_u^1 Y + r_2 p_u^2 Z \tag{7.25}$$

where the parameters W, X, Y and Z from table 7.1 are used to make expressions simpler and depend on the considered priority policy Γ and their values for the six priority policies $\Gamma \in \Omega_\Gamma^{ss}$ are specified in table 3.1.

For each priority policy $\Gamma \in \Omega_\Gamma^{ss}$, we compute the probability that Q_{BS} is empty $\Pi_0(\alpha, \Gamma)$ from *step 2* then we substitute the above expressions of $\mu_u^0(\Gamma)$ and $\mu_u^1(\alpha, \Gamma)$ into the equation (7.23), we obtain (3.9) as the service rate $\mu_u(\alpha, \Gamma)$ of the queue Q_u.

(Step 4) Characterization of the stability region: Combining the results of the previous steps, we deduce the exact stability region of the three-UEs scenario. Supposing that the arrival and service processes of Q_s and Q_u are strictly stationary and ergodic then their stability which is determined using **Loyne's** criterion is given by the condition that the average arrival rate is smaller than the average service rate.

The following procedure is pursued for capturing the stability region of the scenario. We start by considering a priority policies that correspond to the corner points of the stability region ($\Gamma \in \Omega_\Gamma^{ss}$ with $|\Omega_\Gamma^{ss}| = 6$). Then, for this priority policy, we vary $\alpha \in [0, 1]$. For each value of α, we find the probability that Q_{BS} is empty in order to deduce the service rates of the queues in the system. This procedure is applied for all the priority policies $\Gamma \in \Omega_\Gamma^{ss}$.

Therefore, the stability region for the 3-UEs cellular scenario is characterized by the set of mean arrival rates λ_s and λ_u in$\mathcal{R}_c^{ss}$ such that :

$$\mathcal{R}_c^{ss} = co\left(\bigcup_{\Gamma \in \Omega_\Gamma^{ss}} \bigcup_{\alpha \in [0,1]^4} \{\mu_s(\alpha, \Gamma), \mu_u(\alpha, \Gamma)\}\right)$$

where $\mu_s(\alpha, \Gamma)$ and $\mu_u(\alpha, \Gamma)$ are respectively given by (3.8) and (3.9).

(Step 5) Arrival rates within the stability region is equivalent to the stability of the system of queues: We prove that if the mean arrival rates λ_s

and λ_u are in the region $\mathcal{R}_c^{ss}$ is equivalent to the stability of the system of queues. To do so we prove that having $\boldsymbol{\lambda} \in \mathcal{R}_c^{ss}$ gives the stability of the queues and vice versa. We suppose the following notation:

- $a_i(t)$: arrival process at queue Q_i for $i = s, d, u$
- λ_s and λ_u the mean arrival rates at respectively the *UE2UE* and *UE2BS communications*. By the law of large numbers, we have with probability 1:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} a_i(\tau) = \mathbb{E}\left[a_i(t)\right] = \lambda_i$$

- The second moments of the arrival processes $\mathbb{E}\left[a_i(t)^2\right]$ are assumed to be finite.
- $b_i(t)$: departure process at queue Q_i for $i = s, d, u$
- $S(t) = (S_i(t), S_j(t), S_k(t))$: channel state vector where each SNR state is within $\{S_1, S_2, S_3\}$
- $\Gamma(t)$: policy of scheduling on slot t

Thus for each channel $i \in [s, d, u]$ the queuing dynamics are given by:

$$Q_i(t+1) = [Q_i(t) - b_i(t)]^+ + a_i(t)$$

where $b_i(t)$ represents the amount of service offered to channel i on slot t and is defined by a function $\hat{b}_i(S(t), Q(t), \Gamma(t))$:

$$b_i(t) = \hat{b}_i(S(t), Q(t), \Gamma(t)) = r_i q_i(S(t), Q(t), \Gamma(t))$$

Further, by the law of large numbers, we have with probability 1:

$$\mu_i(\Gamma) = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} b_i(\tau|\Gamma) = \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \hat{b}_i(S(\tau), Q(\tau)|\Gamma)$$

$$= \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \sum_{Q^*\in\mathbb{Q}} \hat{b}_i(S(\tau)|Q^*, \Gamma)\, \mathbb{1}_{[Q^*]}$$

$$= \lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} \left[\hat{b}_i(S(\tau)|Q_{BS}(\tau) = 0, \Gamma)\, \mathbb{1}_{[Q_{BS}(\tau)=0]} + \hat{b}_i(S(\tau)|Q_{BS}(\tau) > 0, \Gamma)\, \mathbb{1}_{[Q_{BS}(\tau)>0]}\right]$$

$$= \sum_{(S_i,S_j,S_k)\in\mathbb{S}} p_s^i p_u^j p_d^k r_i q_i(S_i, S_j, S_k|Q_{BS} = 0, \Gamma)\, \mathbb{P}[Q_{BS} = 0]$$

$$+ \sum_{(S_i,S_j,S_k)\in\mathbb{S}} p_s^i p_u^j p_d^k r_i q_i(S_i, S_j, S_k|Q_{BS} > 0, \Gamma)\, \mathbb{P}[Q_{BS} > 0]$$

$$= \mu_i^0(\Gamma)\, \mathbb{P}[Q_{BS} = 0] + \mu_i^1(\Gamma)\, \mathbb{P}[Q_{BS} > 0] = \mu_i^0(\Gamma)\, \Pi_0 + \mu_i^1(\Gamma)(1 - \Pi_0)$$

(Step 5.a) $\lambda \in \mathcal{R}_c^{ss} \Rightarrow$ ***Stability of the queues:*** $\lambda \in \mathcal{R}_c^{ss} \Rightarrow$ then for each channel $i = s, d, u$ there exists a $\mu_i^* = \sum_{\Gamma} \Pi_\Gamma \mu_i(\Gamma)$ as combination of $\mu_i(\Gamma)$ for different scheduling policies Γ such that $\lambda_i \leq \mu_i^*$. We have by the law of large numbers that:

$$\lim_{t\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} b_i(\tau) = \lim_{t\to\infty} \frac{1}{t} \sum_{\Gamma^*\in\Omega_\Gamma} \sum_{\tau=0}^{t-1} b_i(\tau|\Gamma=\Gamma^*) \mathbb{1}_{\Gamma=\Gamma^*} = \lim_{t\to\infty} \frac{1}{t} \sum_{\Gamma^*\in\Omega_\Gamma} \Pi_{\Gamma^*} \sum_{\tau=0}^{t-1} b_i(\tau|\Gamma=\Gamma^*)$$

$$= \sum_{\Gamma^*\in\Omega_\Gamma} \Pi_{\Gamma^*} \mu_i(\Gamma^*) = \mu_i^*$$

Hence $\lambda_i \leq \mu_i^* = \lim_{\to\infty} \frac{1}{t} \sum_{\tau=0}^{t-1} b_i(\tau)$ which gives that queue i is stable for $i = s, d, u$. We deduce that if $\lambda \in \mathcal{R}_c^{ss}$ then the system of queues is stable.

(Step 5.b) Stability of the queues $\Rightarrow \lambda \in \mathcal{R}_c^{ss}$***:*** If the queues are stable then each queue $i \in \{s, d, u\}$ has a stationary distribution Π_i. The mean service rate is given by:

$$\mu_i = \mathbb{E}\left[\hat{b}_i(S(t), Q(t), \Gamma(t))\right]$$

$$= \sum_{\Gamma^*\in\Omega_\Gamma} \Pi_{\Gamma^*} \left[\mathbb{E}\left[\hat{b}_i(S(t)|Q_{BS}=0, \Gamma=\Gamma^*) \Pi_{BS}(0)\right] + \mathbb{E}\left[\hat{b}_i(S(t)|Q_{BS}>0, \Gamma=\Gamma^*) \bar{\Pi}_{BS}(0)\right]\right]$$

$$= \sum_{\Gamma^*\in\Omega_\Gamma} \Pi_{\Gamma^*} \mu_i(\Gamma^*)$$

Since each queue $i \in \{s, d, u\}$ is stable, then $\lambda_i \leq \mu_i$ which gives $\lambda_i \leq \sum_{\Gamma^*\in\Omega_\Gamma} \Pi_{\Gamma^*} \mu_i(\Gamma^*)$ which is a combination of the limit of the stability region $\mathcal{R}_c^{ss}$ hence $\lambda \in \mathcal{R}_c^{ss}$.

7.3.2 Proof of lemma 3.3.2

We propose an approximated model for the three-UEs cellular scenario and we define the condition that the fraction vector α should verify in order to satisfy the stability of the approximated BS queue $\tilde{Q}_{BS}$.

(Step 1) Approximated system model: Cellular communications are modeled as coupled processor sharing queues where the service rates of Q_s (equivalent to the arrival rate of Q_{BS}) as well as the service rate of Q_u depend on the state (empty or not) of Q_{BS}. We have already studied the stability region of the cellular scenario by considering the real Markov chain representing the queue Q_{BS}. This real Markov chain considers the transition probabilities of both rates r_1 and r_2. We note that finding such stability region is done by varying the vector $\alpha \in [0, 1]^4$ and solving for each α the linear system of equations (see subsection 7.3.1) in order to find the probability that Q_{BS} is empty.

Let us present an approximation of the stability region for which we provide an explicit formula of the points (μ_s, μ_u) as well as the exact fraction vector $\tilde{\alpha}^*$ for which the limit of the stability region is achieved. Therefore, we avoid the need of considering all the $\alpha \in [0, 1]^4$ and solving the system of equations.

To do so, we consider the average number of packets received and transmitted per time slot at the BS side in order to find the probability of receiving and transmitting a packet per time slot at the BS level. Thus, the goal is to have an approximated queue $\tilde{Q}_{BS}$ that can be modeled by a simple birth and death Markov chain. We remark first that the average number of packets received by the BS is given by $Pa_{02} + kPa_{01}$ when the queue is empty and this average is equal to $Pa_{12} + kPa_{11}$ otherwise, we note second that the average number of transmitted packets by the BS is equal to $Pb_{12} + kPb_{11}$. In other words, with probability $a_{12} + ka_{11}$ the BS is receiving P packets per time slot and with probability $b_{12} + kb_{11}$ the BS is transmitting P packets per time slot.

The approximation is based on the simplification of the Markov Chain model that helps us to reduce the complexity of the problem. Therefore, the approximated Markov Chain still take into account the multiple rate model at the level of the probabilities of transition. Indeed, the multiple rate model is integrated in such a way that the probability of transmitting P packets at rate r_1 is given by the sum of the two following terms: (i) k times the probability of having a transmission at a rate r_1 and (ii) the probability of having a transmission at a rate r_2.

(Step 1.a) Approximated Markov Chain $\tilde{Q}_{BS}$***:*** As we mentioned before, the scheduling decision depends on the state (empty or not) of Q_{BS} which means that the behavior of Q_s and Q_u is coupled to the state of Q_{BS}. Thus, the arrival probability when Q_{BS} is empty (state 0 of Markov chain) is not equal to the arrival probability when Q_{BS} is not empty (state $\neq 0$ of Markov chain). The approximated queue $\tilde{Q}_{BS}$ corresponds to the queue at the BS side for which the transition probability from a state x_i to a state x_{i+1} is equivalent to the probability of transmitting P packets per time slot by Q_s which is equal to the probability of receiving P packets per time slot by $\tilde{Q}_{BS}$ which means a probability of $a_{02} + ka_{01}$ when $\tilde{Q}_{BS}$ is empty and $a_{12} + ka_{11}$ otherwise. Moreover, the transition probability from a state x_i to a state x_{i-1} is equivalent to the probability of transmitting P packets per time slot at the DL BS-UE_d which means a probability of $b_{12} + kb_{11}$. In figure 7.2, we present the discrete time Markov Chain with infinite states which describes the evolution of the approximated queue $\tilde{Q}_{BS}$.

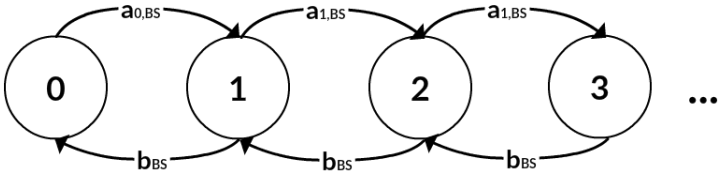


Figure 7.2: Approximated Markov Chain model of the BS queue: $\tilde{Q}_{BS}$

The transition probabilities of the Markov chain corresponding to $\tilde{Q}_{BS}$ are:

- Service probability: $b_{BS}(\Gamma) = b_{12}(\Gamma) + kb_{11}(\Gamma)$.
- Arrival probability when $\tilde{Q}_{BS} = 0$: $a^0_{BS}(\Gamma) = a_{02}(\Gamma) + ka_{01}(\Gamma)$.

- Arrival probability when $\tilde{Q}_{BS} > 0$: $a^1_{BS}(\alpha, \Gamma) = a_{12}(\alpha, \Gamma) + k a_{11}(\alpha, \Gamma)$.

We see that these probabilities depend on the priority policy Γ, but for clarity reasons of the expressions we may use the notation b_{BS}, a^0_{BS} and a^1_{BS} instead of $b_{BS}(\Gamma)$, $a^0_{BS}(\Gamma)$ and $a^1_{BS}(\alpha, \Gamma)$. For the approximated Markov chain we consider the fraction vector $\tilde{\alpha} = [\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4]$.

(Step 1.b) Probability that $\tilde{Q}_{BS}$ is empty: The probability that $\tilde{Q}_{BS}$ is empty is deduced from the stationary distribution $\tilde{\Pi}$ of the approximated Markov Chain of $\tilde{Q}_{BS}$ (figure 7.2) which is computed by applying the balance equations as follows:

$$\tilde{\Pi}_1 = \frac{a^0_{BS}}{b_{BS}} \tilde{\Pi}_0 \; ; \; \tilde{\Pi}_2 = \frac{a^0_{BS} a^1_{BS}}{b_{BS}} \tilde{\Pi}_0 \text{ and } \tilde{\Pi}_n = \frac{a^0_{BS} \left(a^1_{BS}\right)^{n-1}}{\left(b_{BS}\right)^n} \tilde{\Pi}_0$$

Knowing that the stationary distribution satisfies $\sum_{n=1}^{\infty} \tilde{\Pi}_n = 1$; hence the probability that $\tilde{Q}_{BS}$ is empty for a given priority policy Γ is given by:

$$\tilde{\Pi}_0 = \mathbb{P}\left[\tilde{Q}_{BS} = 0\right] = \frac{b_{BS} - a^1_{BS}}{b_{BS} - a^1_{BS} + a^0_{BS}} \text{ iff } \; a_{BS} < b_{BS}. \tag{7.26}$$

(Step 2) Service and arrival probabilities of $\tilde{Q}_{BS}$: In order to find the stability conditions, we need to find the expressions of transitions probabilities of the approximated Markov chain of $\tilde{Q}_{BS}$. Therefore, both service and arrival probability o the queue $\tilde{Q}_{BS}$ are computed. Using parameters U, V and N from table 7.1 and based on equations (7.1) and (7.2), the service probability b_{BS} of the queue $\tilde{Q}_{BS}$ can be written as follows:

$$b_{BS} = k p^1_d \left(1 - \tilde{\alpha}_1 p^1_s - \tilde{\alpha}_3 p^2_s\right) U + p^2_d \left(N - \tilde{\alpha}_2 p^1_s U - \tilde{\alpha}_4 p^2_s V\right) \tag{7.27}$$

In order to compute the arrival probability of $\tilde{Q}_{BS}$, we should distinguish between a^0_{BS}, i.e. arrival probability when $\tilde{Q}_{BS}$ is empty and a^1_{BS}, i.e. arrival probability when $\tilde{Q}_{BS}$ is not empty. Using parameters U, V and N from table 7.1 and based on equations (7.3), (7.4), (7.5) and (7.6), the probabilities of arrival a^0_{BS} and a^1_{BS} of the approximated queue $\tilde{Q}_{BS}$ respectively when it is empty or not empty can be written as follows:

$$a^0_{BS} = k p^1_s U + p^2_s V \tag{7.28}$$

$$a^1_{BS} = k p^1_s \left(\tilde{\alpha}_1 p^1_d + \tilde{\alpha}_2 p^2_d + p^3_d\right) U + p^2_s \left(\tilde{\alpha}_3 p^1_d U + \tilde{\alpha}_4 p^2_d V + p^3_d V\right) \tag{7.29}$$

The values of the parameter above U, V and N are given in table 3.1 for all the considered priority policies Γ that correspond to the corner point of the stability region ($\Gamma \in \Omega^{ss}_{\Gamma}$). Since the arrival probability a_{BS} at the queue $\tilde{Q}_{BS}$ is given by:

$$\mathbb{P}\left[\tilde{Q}_{BS} = 0\right] a^0_{BS} + \mathbb{P}\left[\tilde{Q}_{BS} > 0\right] a^1_{BS} = \tilde{\Pi}_0 a^0_{BS} + \left(1 - \tilde{\Pi}_0\right) a^1_{BS}$$

$$\Rightarrow a_{BS} = \frac{b_{BS} a^0_{BS}}{b_{BS} - a^1_{BS} + a^0_{BS}} \tag{7.30}$$

(Step 3) Stability condition of $\tilde{Q}_{BS}$**:** Supposing that the arrival and service processes of $\tilde{Q}_{BS}$ are strictly stationary and ergodic then the stability of $\tilde{Q}_{BS}$ can be determined using **Loynes'** criterion which states that the queue is stable if and only if the average arrival rate is strictly less than the average service rate. Then the stability condition that the fraction vector $\tilde{\alpha} = (\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4)$ should verify in order to satisfy the stability of the queue $\tilde{Q}_{BS}$ is given by:

$$a_{BS} < b_{BS} \Leftrightarrow \frac{b_{BS} a^0_{BS}}{b_{BS} - a^1_{BS} + a^0_{BS}} < b_{BS} \Leftrightarrow a^1_{BS} < b_{BS}$$

Since a^1_{BS} and b_{BS} are given by (7.29) and (7.27) then the queue $\tilde{Q}_{BS}$ is stable if the fraction vector $\tilde{\alpha}$ verifies (3.10).

7.3.3 Proof of theorem 3.3.3

We study the approximated model for the three-UEs cellular scenario and we derive its corresponding stability region. For this aim, we proceed similarly to the proof in appendix 7.3.1. However, the first two steps for the approximated model (the modeling of the BS queue by an approximated Markov Chain as well as the computation of the probability that this queue is empty) are already proved in appendix 7.3.2. Here, there are two main challenges: (i) finding the approximated Markov Chain model of the queue Q_{BS} that will help us to achieve an explicit and simple form of the stability region and (ii) the computation of the optimal fraction vector that will reduce the complexity of the problem.

(Step 1) Service rate of $\mathbf{Q}_s$**:** We compute the approximated stability region of the 3-UEs cellular scenario. For that, we compute the service rate of both *UE2UE* and *UE2BS communications* for a given priority policy Γ. We follow the procedure below, based on queuing theory analysis of the network capacity, in order to derive the performance study of cellular system. If $\tilde{Q}_{BS}$ is empty then the service rate of Q_s is denoted by μ^0_s and by μ^1_s otherwise. Since the service rate of Q_s is computed by

$$\tilde{\mu}_s(\tilde{\alpha}, \Gamma) = \mathbb{P}\left[\tilde{Q}_{BS} = 0\right] \mu^0_s(\Gamma) + \mathbb{P}\left[\tilde{Q}_{BS} > 0\right] \mu^1_s(\tilde{\alpha}, \Gamma)$$

$$\Rightarrow \tilde{\mu}_s = \frac{\mu^0_s (b_{BS} - a^1_{BS}) + \mu^1_s a^0_{BS}}{b_{BS} - a^1_{BS} + a^0_{BS}} \tag{7.31}$$

$\mu^0_s(\Gamma)$ and $\mu^1_s(\alpha, \Gamma)$ are respectively given by (7.21) and(7.22), we obtain (7.32) as the service rate $\tilde{\mu}_1(\tilde{\alpha}, \Gamma)$ of the queue Q_s.

$$\tilde{\mu}_s(\tilde{\alpha}, \Gamma) = \frac{\left(r_1 p^1_s U + r_2 p^2_s V\right)\left(-2k\tilde{\alpha}_1 p^1_s p^1_d U - (k+1)\tilde{\alpha}_2 p^1_s p^2_d U - (k+1)\tilde{\alpha}_3 p^2_s p^1_d U - \tilde{\alpha}_4 p^2_s p^2_d V + k p^1_d + p^2_d N - \left(k p^1_s U + p^2_s V\right) p^3_d\right)}{-2k\tilde{\alpha}_1 p^1_s p^1_d U - (k+1)\tilde{\alpha}_2 p^1_s p_d 2U - (k+1)\tilde{\alpha}_3 p^2_s p^1_d U - 2\tilde{\alpha}_4 p^2_s p^2_d V + \left(k p^1_d + p^2_d N\right) + (k p^1_s U + p^2_s V)\left(1 - p^3_d\right)}$$

$$+ \frac{\left(k p^1_s U + p^2_s V\right)\left(r_1 \tilde{\alpha}_1 p^1_s p^1_d W + r_1 \tilde{\alpha}_2 p^1_s p^2_d U + r_2 \tilde{\alpha}_3 p^2_s p^1_d U + r_2 \tilde{\alpha}_4 p^2_s p^2_d V + \left(r_1 p^1_s U + r_2 p^2_s V\right) p^3_d\right)}{-2k\tilde{\alpha}_1 p^1_s p^1_d U - (k+1)\tilde{\alpha}_2 p^1_s p^2_d U - (k+1)\tilde{\alpha}_3 p^2_s p^1_d U - 2\tilde{\alpha}_4 p^2_s p^2_d V + \left(k p^1_d + p^2_d N\right) + (k p^1_s U + p^2_s V)\bar{p}^3_d}$$

$$= (r_1 p^1_s U + r_2 p^2_s V) + (k p^1_s U + p^2_s V) \times$$

$$\frac{\left(-r_1 \tilde{\alpha}_1 p^1_s p^1_d U - r_1 \tilde{\alpha}_2 p^1_s p^2_d U - r_2 \tilde{\alpha}_3 p^2_s p^1_d U - r_2 \tilde{\alpha}_4 p^2_s p^2_d V + \left(r_1 p^1_s U + r_2 p^2_s V\right) \bar{p}^3_d\right)}{2k\tilde{\alpha}_1 p^1_s p^1_d U + (k+1)\tilde{\alpha}_2 p^1_s p^2_d U + (k+1)\tilde{\alpha}_3 p^2_s p^1_d U + 2\tilde{\alpha}_4 p^2_s p^2_d V - \left(k p^1_d + p^2_d N\right) - (k p^1_s U + p^2_s V)\bar{p}^3_d}$$

$$\tilde{\mu}_s(\tilde{\alpha},\Gamma) = \tfrac{1}{2}\left(r_1 p_s^1 U + r_2 p_s^2 V\right) \times$$

$$\left(1 + \frac{(1-k)\tilde{\alpha}_2 p_s^1 p_d^2 U + (k-1)\tilde{\alpha}_3 p_s^2 p_d^1 U - \left(k p_d^1 + p_d^2 N\right) + \left(k p_s^1 U + p_s^2 V\right)\bar{p}_d^3}{2k\tilde{\alpha}_1 p_s^1 p_d^1 U + (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U + (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U + 2\tilde{\alpha}_4 p_s^2 p_d^2 V - \left(k p_d^1 + p_d^2 N\right) - (k p_s^1 U + p_s^2 V)\bar{p}_d^3}\right) \quad (7.32)$$

(Step 2) Service rate of Q_u: If $\tilde{Q}_{BS}$ is empty then the service rate of Q_u is denoted by μ_u^0 otherwise it is denoted by μ_u^1. By analogy to the Q_s, the service rate of Q_u is computed as follows:

$$\tilde{\mu}_u(\tilde{\alpha},\Gamma) = \mathbb{P}\left[\tilde{Q}_{BS} = 0\right]\mu_u^0(\Gamma) + \mathbb{P}\left[\tilde{Q}_{BS} > 0\right]\mu_u^1(\tilde{\alpha},\Gamma)$$

$$\Rightarrow \tilde{\mu}_u = \frac{\mu_u^0\left(b_{BS-} a_{BS}^1\right) + \mu_u^1 a_{BS}^0}{b_{BS} - a_{BS}^1 + a_{BS}^0} \quad (7.33)$$

With $\mu_u^0(\Gamma)$ and $\mu_u^1(\alpha,\Gamma)$ respectively given by (7.24) and(7.25), we obtain (7.34) as the service rate $\tilde{\mu}_2(\tilde{\alpha},\Gamma)$ of the queue Q_u.

$$\tilde{\mu}_u(\tilde{\alpha},\Gamma) = \frac{\left(r_1 p_u^1 W + r_2 p_u^2 X\right)\left(-2k\tilde{\alpha}_1 p_s^1 p_d^1 U - (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U - (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U - 2\tilde{\alpha}_4 p_s^2 p_d^2 V + k p_d^1 + p_d^2 N - \left(k p_s^1 U + p_s^2 V\right)p_d^3\right)}{-2k\tilde{\alpha}_1 p_s^1 p_d^1 U - (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U - (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U - 2\tilde{\alpha}_4 p_s^2 p_d^2 V + \left(k p_d^1 + p_d^2 N\right) + (k p_s^1 U + p_s^2 V)\bar{p}_d^3}$$

$$+ \frac{\left(r_1 p_u^1 Y + r_2 p_u^2 Z\right)\left(k p_s^1 U + p_s^2 V\right)}{-2k\tilde{\alpha}_1 p_s^1 p_d^1 U - (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U - (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U - 2\tilde{\alpha}_4 p_s^2 p_d^2 V + \left(k p_d^1 + p_d^2 N\right) + (k p_s^1 U + p_s^2 V)\bar{p}_d^3}$$

$$= \left(r_1 p_u^1 W + r_2 p_u^2 X\right) +$$

$$\frac{\left(r_1 p_u^1 W + r_2 p_u^2 X\right)\left(-k p_s^1 U + p_s^2 V\right) + \left(r_1 p_u^1 Y + r_2 p_u^2 Z\right)\left(k p_s^1 U + p_s^2 V\right)}{-2k\tilde{\alpha}_1 p_s^1 p_d^1 U - (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U - (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U - 2\tilde{\alpha}_4 p_s^2 p_d^2 V + \left(k p_d^1 + p_d^2 N\right) + (k p_s^1 U + p_s^2 V)\bar{p}_d^3}$$

$$\tilde{\mu}_u(\tilde{\alpha},\Gamma) = \left(r_1 p_u^1 W + r_2 p_u^2 X\right) +$$

$$\frac{\left(r_1 p_u^1 (W-Y) + r_2 p_u^2 (X-Z)\right)\left(k p_s^1 U + p_s^2 V\right)}{2k\tilde{\alpha}_1 p_s^1 p_d^1 U + (k+1)\tilde{\alpha}_2 p_s^1 p_d^2 U + (k+1)\tilde{\alpha}_3 p_s^2 p_d^1 U + 2\tilde{\alpha}_4 p_s^2 p_d^2 V - \left(k p_d^1 + p_d^2 N\right) - (k p_s^1 U + p_s^2 V)\bar{p}_d^3} \quad (7.34)$$

(Step 3) The optimal fraction vector α^*: We know that finding the corner points of approximated stability region are sufficient to characterize this region. Hence, we intend to reduce more the complexity by avoiding the consideration of all the $\tilde{\alpha} \in [0,1]^4$ and finding the set of optimum α^* that achieve the corner points of the approximated stability region. So, for each priority policy $\Gamma \in \Omega_\Gamma^{ss}$, we compute the optimum $\tilde{\alpha}^* = (\tilde{\alpha}_1^*, \tilde{\alpha}_2^*, \tilde{\alpha}_3^*, \alpha_4^*)$ that maximizes μ_s and μ_u such a border point of the stability region is achieved. In other words, finding $\tilde{\alpha}^*$ for each priority policy $\Gamma \in \Omega_\Gamma^{ss}$ allows us to find explicitly the corresponding corner point and to avoid the need to vary $\tilde{\alpha} \in [0,1]^4$ for obtaining this point of the stability region.

We start by the service rate $\tilde{\mu}_u$. We can see that for all the priority policies, $\tilde{\mu}_u(\tilde{\alpha},\Gamma)$ is inversely proportional to $\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3$ and $\tilde{\alpha}_4$ then $\tilde{\alpha}^*=(0,0,0,0)$ maximizes $\tilde{\mu}_u(\tilde{\alpha},\Gamma)$ for all $\Gamma \in \Omega_\Gamma^{ss}$. It remains to maximize $\tilde{\mu}_s$ by solving the following optimization problem for each priority policy:

$$\max_{\tilde{\alpha}} \ \tilde{\mu}_s\left(\tilde{\alpha}, \Gamma\right)$$

$$s.t.\, C_0 = \begin{cases} C_{01}: \ (3.10)\, is\, verified \\ C_{02}: \ 0 \leq \tilde{\alpha}_1 \leq 1, \ 0 \leq \tilde{\alpha}_2 \leq 1 \\ 0 \leq \tilde{\alpha}_3 \leq 1\, and\, 0 \leq \tilde{\alpha}_4 \leq 1 \end{cases} \tag{7.35}$$

The optimization problem above is a Linear-Fractional Programming (LFP) where the objective function is the ratio of two linear functions with a set of linear constraints. This LFP has a bounded and non-empty feasible region and can be transformed into an equivalent Linear Problem (LP) which the solution (solved by any LP solution method such that simplex algorithm) yields that of the original LFP problem.

The solution of the optimization problem above belongs to the scenarios where at least one of the two constraints (C_{01},C_{02}) is reached which corresponds to different cases: (i) constraint C_{02} is achieved ($(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4) \in \{0,1\}^4$) while the stability condition C_{01} is verified without equality and (ii) the stability condition C_{01} is achieved with $(\tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4) \in [0,1]^4$. These cases will be studied in order to come up with the set, called $\mathbb{S}_\alpha$, of potential solutions of the optimization problem (7.35) . Several cases will be detailed where from each one we can deduce an element of the set $\mathbb{S}_\alpha$.

Case 1: We start by supposing that at the optimal solution the constraint C_{01} is verified but not achieved while the second constraint C_{02} is reached. The optimum in this case is the combination of $\tilde{\alpha} = \tilde{\alpha}_1, \tilde{\alpha}_2, \tilde{\alpha}_3, \tilde{\alpha}_4 \in \{0,1\}^4$ that maximizes $\tilde{\mu}_s\left(\tilde{\alpha}, \Gamma\right)$ while verifying the stability condition C_{01}.

Case 2: Now, we suppose that the stability condition C_{01} is achieved at the optimum. Hence, the optimization problem is converted to the problem (7.36) with $F_1 = M$ and $F_2 = F_1 - 2kp_s^1 p_d^1 U - 2p_s^2 p_d^2 V$. Recall that $M = (kp_d^1 + p_d^2 N) - (kp_s^1 U + p_s^2 V)\, p_d^3$.

By analogy to (7.35), we solve the problem (7.36) by considering that at least one of the linear constraints is reached which corresponds to the two following cases: (i) the first one assumes that C_{12} constraint is achieved and (ii) the other one considers that C_{11} constraint is achieved. We study both cases then the optimum corresponds to the case that maximizes $\tilde{\mu}_s\left(\tilde{\alpha}, \Gamma\right)$. We start by supposing that at the optimal solution the constraint C_{11} is not achieved then the optimum in this case is the combination of $\tilde{\alpha}_2, \tilde{\alpha}_3 \in \{0,1\}^2$ that maximizes $\tilde{\mu}_s\left(\tilde{\alpha}, \Gamma\right)$ and for which there exists $\tilde{\alpha}_1$ and $\tilde{\alpha}_4 \in [0,1]$ that satisfy (7.36).

$$\max_{\tilde{\alpha}_2, \tilde{\alpha}_3} \ \tfrac{1}{2}\left(r_1 p_s^1 U + r_2 p_s^2 V\right)\left(1 + \frac{(k-1)\tilde{\alpha}_2 p_s^1 p_d^2 U + (1-k)\tilde{\alpha}_3 p_s^2 p_d^1 U + \left(kp_d^1 + p_d^2 N\right) - \left(kp_s^1 U + p_s^2 V\right)\bar{p}_d^3}{kp_s^1 U + p_s^2 V}\right)$$

$$s.t.\, C_1 = \begin{cases} C_{11}: F_2 \leq (k+1)\,\tilde{\alpha}_2 p_s^1 p_d^2 U + (k+1)\,\tilde{\alpha}_3 p_s^2 p_d^1 U \leq F_1 \\ C_{12}: 0 \leq \tilde{\alpha}_2 \leq 1\, and\, 0 \leq \tilde{\alpha}_3 \leq 1 \end{cases} \tag{7.36}$$

Case 3: For this case, we suppose that the constraint C_{11} is active at the optimum by achieving either the lower bound F_1 (for $\tilde{\alpha}_1 = \tilde{\alpha}_4 = 0$) or the upper bound F_2 (for $\tilde{\alpha}_1 = \tilde{\alpha}_4 = 1$). For each bound F_1 and F_2 respectively, $\tilde{\alpha}_2$ and $\tilde{\alpha}_3$ are chosen in a way to verify respectively the following equations: $(k+1)\,\tilde{\alpha}_2 p_s^1 p_d^2 U + (k+1)\,\tilde{\alpha}_3 p_s^2 p_d^1 U = F_i$ for $i = 1, 2$. For this case, we suppose that the constraint C_{11} is active at the optimum then the optimization problem can be written as in (7.37).

The LFP problem (7.37) depends only from $\tilde{\alpha}_2$ where the objective function is linearly depending on $\tilde{\alpha}_2$ with a positive coefficient. Hence, the objective function is maximized by increasing as much as possible $\tilde{\alpha}_2$ such that the constraints (7.37) are verified.

Combining the results of the three cases above, for each priority policy Γ, we find the set $\mathbb{S}_\alpha$ to which corresponds the optimum α^* that maximizes $\tilde{\mu}_s(\tilde{\alpha}, \Gamma)$ and achieves the corresponding corner point of the stability region. Hence, we hugely reduce the complexity by finding the explicit value of α^* that gives the corner points that are sufficient to characterize the approximated stability region. Finally, for each priority policy within $\Gamma \in \Omega_\Gamma^{ss}$, the optimum $\tilde{\alpha}^* = (\tilde{\alpha}_1^*, \tilde{\alpha}_2^*, \tilde{\alpha}_3^*, \tilde{\alpha}_4^*)$, that reaches the corner point of the approximated stability region corresponding to Γ, to the finite set (3.13) denoted by $\mathbb{S}_\alpha$.

$$\max_{\tilde{\alpha}_2} \ \frac{1}{2}\left(r_1 p_s^1 U + r_2 p_s^2 V\right)\left(1 + \frac{2(k-1)\tilde{\alpha}_2 p_s^1 p_d^2 U + \frac{(1-k)}{(1+k)} F_i + \left(k p_d^1 + p_d^2 N\right) - \left(k p_s^1 U + p_s^2 V\right)\bar{p}_d^3}{k p_s^1 U + p_s^2 V}\right)$$

$$s.t.\ C_1 = \begin{cases} C_{21} : (k+1)\,\tilde{\alpha}_2 p_s^1 p_d^2 U + (k+1)\,\tilde{\alpha}_3 p_s^2 p_d^1 U = F_i \\ C_{22} : 0 \le \tilde{\alpha}_2 \le 1 \\ C_{23} : \ for\ F_1 :\ \tilde{\alpha}_1 = \tilde{\alpha}_4 = 0\ ;\ for\ F_2 :\ \tilde{\alpha}_1 = \tilde{\alpha}_4 = 1 \end{cases} \tag{7.37}$$

$$\mathbb{S}_\alpha = \left(\begin{array}{c} \tilde{\alpha}_1 = 0, \tilde{\alpha}_2 = 1, \tilde{\alpha}_3 = \frac{F_1 - (k+1) p_s^1 p_d^2}{(k+1) p_d^1 p_s^2 U}, \tilde{\alpha}_4 = 0 \\ \tilde{\alpha}_1 = 1, \tilde{\alpha}_2 = 1, \tilde{\alpha}_3 = \frac{F_2 - (k+1) p_s^1 p_d^2 U}{(k+1) p_d^1 p_s^2 U}, \tilde{\alpha}_4 = 1 \\ \tilde{\alpha}_1 = 0, \tilde{\alpha}_2 = \frac{F_1}{(k+1) p_s^1 p_d^2 U}, \tilde{\alpha}_3 = 0, \tilde{\alpha}_4 = 0 \\ \tilde{\alpha}_1 = 0, \tilde{\alpha}_2 = \frac{F_1 - (k+1) p_d^1 p_s^2 U}{(k+1) p_s^1 p_d^2 U}, \tilde{\alpha}_3 = 1, \tilde{\alpha}_4 = 0 \\ \tilde{\alpha}_1 = 1, \tilde{\alpha}_2 = \frac{F_2}{(k+1) p_s^1 p_d^2 U}, \tilde{\alpha}_3 = 0, \alpha_4 = 1 \\ \tilde{\alpha}_1 = 1, \tilde{\alpha}_2 = \frac{F_2 - (k+1) p_d^1 p_s^2 U}{(k+1) p_s^1 p_d^2 U}, \tilde{\alpha}_3 = 1, \tilde{\alpha}_4 = 1 \end{array} \right)$$

(Step 4) Characterization of the approximated stability region: Combining the results of the previous steps, we deduce the approximation of the stability region for the three-UEs cellular scenario. Supposing that the arrival and service processes of Q_s and Q_u are strictly stationary and ergodic then the stability of $\tilde{Q}_{BS}$ can be determined using **Loynes'** criterion which states that the queue is stable if and only if the average arrival rate is strictly less than the average service rate.

We proceed as follows for characterizing the approximated stability region. We start by considering a priority policy that corresponds to a corner point of the region ($\Gamma \in \Omega_{\Gamma}^{ss}$ with $|\Omega_{\Gamma}^{ss}| = 6$). Then, for each considered Γ, we find the corresponding optimum fraction vector α^* within the finite set (3.13). We deduce the service rate of the system of queues. Finally, these steps are applied over all the priority policies $\Gamma \in \Omega_{\Gamma}^{ss}$. Thus, the approximated stability region $\tilde{\mathcal{R}}_c^{ss}$ for the 3-UEs scenario is characterized by the set of the mean arrival rates (λ_s, λ_u) in $\tilde{\mathcal{R}}_c^{ss}$ such that:

$$\tilde{\mathcal{R}}_c^{ss} = co\left(\bigcup_{\Gamma \in \Omega_{\Gamma}^{ss}} \left\{\tilde{\mu}_s\left(\tilde{\alpha}^*, \Gamma\right), \tilde{\mu}_u\left(\tilde{\alpha}^*, \Gamma\right)\right\}\right)$$

where $\tilde{\mu}_s\left(\tilde{\alpha}^*, \Gamma\right)$ and $\tilde{\mu}_u\left(\tilde{\alpha}^*, \Gamma\right)$ are respectively given by (3.11) and (3.12).

(Step 5) Arrival rates within the stability region is equivalent to the stability of the system of queues: Proving that having a set of mean arrival rates (λ_s, λ_u) in $\tilde{\mathcal{R}}_c^{ss}$ is equivalent to the stability of the system queue is identical to the demonstration done in step 5 of the appendix 7.3.1. For brevity, we avoid repetition and remove this proof.

Policy	$q_1^d(S_i, S_j)$ and $q_2^d(S_i, S_j)$
1	$q_1^d(S_1, S_{1,2,3}) = q_1^d(S_2, S_{1,2,3}) = 1$ $q_2^d(S_3, S_{1,2}) = 1$
2	$q_1^d(S_{1,2}, S_3) = 1$ $q_2^d(S_{1,2,3}, S_1) = q_2^d(S_{1,2,3}, S_2) = 1$
3	$q_1^d(S_1, S_{1,2,3}) = q_1^d(S_2, S_3) = 1$ $q_2^d(S_{2,3}, S_1) = q_2^d(S_{2,3}, S_2) = 1$
4	$q_1^d(S_1, S_{2,3}) = q_1^d(S_2, S_{2,3}) = 1$ $q_2^d(S_{1,2,3}, S_1) = q_2^d(S_3, S_2) = 1$
5	$q_1^d(S_1, S_{1,2,3}) = q_1^d(S_2, S_{2,3}) = 1$ $q_2^d(S_{2,3}, S_1) = q_2^d(S_3, S_2) = 1$
6	$q_1^d(S_1, S_{2,3}) = q_1^d(S_2, S_3) = 1$ $q_2^d(S_{1,2,3}, S_1) = q_2^d(S_{2,3}, S_2) = 1$

Table 7.2: D2D Scenario: Probability of transmission over link 1 and 2 for different policies and different states of channel

7.3.4 Proof of proposition 1

For the 3-UE D2D scenario: the UE2UE communication is established directly between UE_s and UE_d and modeled by the queue Q_s^d while UE2BS communication is modeled

by Q_u^d. For this scenario, we denote by $q_n^d\left(S_i, S_j, \Gamma\right)$ the probability to transmit over the link_n (with $n \in \{s, u\}$) for a given policy Γ and a given state of the links: link_s at state S_s and link_u is at state S_u (with S_s and $S_u \in (S_1, S_2, S_3)$). The values of the corresponding transmission probabilities for all $\Gamma \in \Omega_\Gamma^{ss}$ are computed and for all the SNR states (S_i, S_j) are given in table 7.2. Note that $q_s^d\left(S_i, S_j, \Gamma_4\right) = q_s^d\left(S_i, S_j, \Gamma_4\right) = 0$ for all the the other combinations of not shown in the table. For D2D scenario, we apply the classical methodology for computing the stability region. Based on queuing theory analysis of the network capacity, we derive the following analytic expressions of the queues service rates for a given policy $\Gamma \in \Omega_\Gamma^{ss}$:

$$\mu_s^d = \mathbb{E}\left[\mu_1^d\left(S_i, S_j, \Gamma\right)\right] = \sum_{i=1}^{3} r_i p_{sd}^i \sum_{j=1}^{3} p_u^j q_s^d(S_i, S_j, \Gamma) = r_1 p_{sd}^1 \sum_{j=1}^{3} p_u^j q_s^d(S_1, S_j, \Gamma) + r_2 p_{sd}^2 \sum_{j=1}^{3} p_u^j q_s^d(S_2, S_j, \Gamma)$$

$$\mu_u^d = \mathbb{E}\left[\mu_u^d\left(S_i, S_j, \Gamma\right)\right] = \sum_{j=1}^{3} r_j p_u^j \sum_{i} p_{sd}^i q_u^d(S_i, S_j, \Gamma) = r_1 p_u^1 \sum_{i=1}^{3} p_{sd}^i q_u^d(S_i, S_1, \Gamma) + r_2 p_u^2 \sum_{i=1}^{3} p_{sd}^i q_u^d(S_i, S_2, \Gamma)$$

Parameters A, B, C and D from table 7.1 are used to make expressions simpler. These parameters depend on the considered priority policy Γ and their values, deduced from that of $q_1^d\left(S_i, S_j, \Gamma\right)$ and $q_2^d\left(S_i, S_j, \Gamma\right)$ in 7.2, are specified in table 3.1.
Hence, the service rates of Q_s^d and Q_u can respectively be written as (3.14) and (3.15) for all the priority policy $\Gamma \in \Omega_\Gamma^{ss}$. Thus, the stability region for the 3-UEs D2D scenario is given by the proposition (1).

7.3.5 Proof of theorem 3.3.5

For a given priority policy Γ, the D2D distance threshold d_{th} (as well as the path-loss threshold PL_{th}) corresponds to the values achieved when the following equality is verified:

$$\mu_s^d\left(\Gamma\right) = \mu_s^d\left(\Gamma\right) \Rightarrow r_1 p_{sd}^1 U + r_2 p_{sd}^2 V = \mu_s^d\left(\Gamma\right) \Rightarrow$$

$$r_1 U\left(\exp\left(-\frac{\gamma_1^{UL} N_0 PL_{th}}{P_{UL}}\right)\right) + r_2 V\left(\exp\left(-\frac{\gamma_2^{UL} N_0 PL_{th}}{P_{UL}}\right) - \exp\left(-\frac{\gamma_1^{UL} N_0 PL_{th}}{P_{UL}}\right)\right) = \mu_1^d\left(\Gamma\right)$$

If we consider $a = \frac{\gamma_1^{UL} N_0}{P_{UL}}$, $b = \frac{\gamma_2^{UL} N_0}{P_{UL}}$, and $y = PL_{th} z^{\frac{1}{a}} \left(\frac{V r_2}{\mu_1}\right)^{\frac{1}{b}}$ This trinomial equation has similar form as $y^a + z y^b = 1$ with $z = \left(r_1 U + r_2 V\right)\left(\frac{V r_2}{\mu_1}\right)^{-\frac{a}{b}}$. We can easily prove that $a > b$ since $\gamma_1^{UL} \geq \gamma_2^{UL}$. Hence, the solution of the trinomial equation can be found by means of Lagrange inversion see [108] series or also by Mellin transform (see [109] as a summary of his work) as:

$$y = \frac{1}{a} \sum_{n=0}^{\infty} \frac{\Gamma\left(\frac{1+nb}{a}\right)}{\Gamma\left(\frac{1+nb}{a} + 1 - n\right) n!} (-1)^n z^{-\frac{b}{a} n}$$

Thus, PL_{th} is given by:

$$PL_{th} = \left(\frac{V r_2}{\mu_1}\right)^{\frac{-1}{b}} z^{\frac{-1}{a}} y$$

7.3.6 Proof of theorem 3.3.4

We compare between the real and approximated stability region for the three-UEs scenario. For this aim, we proceed similarly to the proof in appendix 7.3.1.

(Step 1) Comparison between the Markov chains of Q_{BS} and $\tilde{Q}_{BS}$: We start by presenting the used notation for facilitating the readability of the expressions:

- The comparison between the two Markov chain models (exact and approximation) is done for each priority policy $\Gamma \in \Omega_{\Gamma}^{ss}$, hence to improve the clarity of the expressions we use the following: $a_{01} = a_{01}(\alpha, \Gamma)$, $a_{02} = a_{02}(\alpha, \Gamma)$, $a_{11} = a_{11}(\alpha, \Gamma)$, $a_{12} = a_{12}(\alpha, \Gamma)$, $b_{11} = b_{11}(\alpha, \Gamma)$, $b_{12} = b_{12}(\alpha, \Gamma)$
- $a_0 = a_{01} + a_{02}$, $a_1 = a_{11} + a_{12}$, $b_1 = b_{11} + b_{12}$
- $\Pi_n = \Pi_n(\alpha, \Gamma)$ probability that Q_{BS} is at state n of the Markov chain which means that the queue contains n P-packets
- $\tilde{\Pi}_n = \tilde{\Pi}_n(\alpha, \Gamma)$ probability that $\tilde{Q}_{BS}$ is at state n of the approximated Markov chain which means that the approximated queue $\tilde{Q}_{BS}$ contains n P-packets.

The main challenges of the exact stability region computation lies on the following to facts: (i) this region has no closed form expression since the final solution requires solving a linear system and (ii) the complexity of the exact stability region is high. Therefore, the advantage of the proposed approximation is to have a close upper bound that has an simple and explicit solution.

The Markov Chain modeling is restrictively used for computing the probability that the corresponding queues are empty. Thus, the comparison between the exact and the approximated Markov Chains is done by finding the absolute error between the probability that Q_{BS} is empty (Π_0) and the probability that $\tilde{Q}_{BS}$ is empty ($\tilde{\Pi}_0$). For a given priority policy Γ and a given fraction vector α this error is defined as follows:

$$\delta_{\Pi}(\alpha, \Gamma) := \Pi_0(\alpha, \Gamma) - \tilde{\Pi}_0(\alpha, \Gamma)$$

Let us consider the Markov chain of the queue Q_{BS} given in figure (7.1) and its corresponding balance equations given by (7.7), (7.8), (7.9) and (7.11). The generating function corresponding to this Markov chain is given by:

$$G(z) = \sum_{n=0}^{\infty} \Pi_n z^n$$

We deduce $G(z)$ as function of Π_0 and Π_1 by multiplying the n^{th} balance equation by z^n and by adding all these equations ($for\, n = 0, ..., \infty$) as follows:

$$\Pi_0 a_0 + \sum_{n=1}^{\infty} \Pi_n (a_1 + b_{12}) z^n + b_{11} \sum_{n=k}^{\infty} \Pi_n z^n$$

$$= a_{11} \sum_{n=k+1}^{\infty} \Pi_{n-k} z^n + a_{12} \sum_{n=2}^{\infty} \Pi_{n-1} z^n + b_{12} \sum_{n=0}^{\infty} \Pi_{n+1} z^n + b_{11} \sum_{n=0}^{\infty} \Pi_{n+k} z^n + a_{02} \Pi_0 z + a_{01} \Pi_0 z^k$$

$$\Rightarrow \sum_{n=0}^{\infty} \Pi_n \left(a_1 + b_1\right) z^n + \Pi_0 \left(a_0 - a_1 - b_1\right) - b_{11} \sum_{n=1}^{k-1} \Pi_n z^n$$

$$= a_{11} \sum_{n=1}^{\infty} \Pi_n z^{n+k} + a_{12} \sum_{n=1}^{\infty} \Pi_n z^{n+1} + b_{12} \sum_{n=1}^{\infty} \Pi_n z^{n-1} + b_{11} \sum_{n=k}^{\infty} \Pi_n z^{n-k} + a_{02} \Pi_0 z + a_{01} \Pi_0 z^k$$

$$\Rightarrow G\left(z\right) \left(a_{11} \sum_{n=2k-1}^{k+1} z^n + a_1 z^k - b_1 z^{k-1} - b_{11} \sum_{n=k-2}^{0} z^n \right)$$

$$= \Pi_0 \left(\left(a_{11} - a_{01}\right) \sum_{n=2k-1}^{k+1} z^n + \left(a_1 - a_0\right) z^k - b_1 z^{k-1} - b_{11} \sum_{n=k-2}^{0} z^n \right) - b_{11} \left(\sum_{i=k-1}^{0} z^i \right) \sum_{n=1}^{k-1} \Pi_n z^n$$

Since normalization constraint gives $G\left(1\right) = 1$, then Π_0 is given by:

$$\Pi_0 = \frac{b_{12} + k b_{11} - a_{12} - k a_{11}}{b_{12} + k b_{11} - a_{12} - k a_{11} + a_{02} + k a_{01}} - \frac{k b_{11} \sum_{n=1}^{k-1} \Pi_n}{b_{12} + k b_{11} - a_{12} - k a_{11} + a_{02} + k a_{01}}$$

Moreover, given equation (7.26), the expression of $\tilde{\Pi}_0$ can be written as follows:

$$\tilde{\Pi}_0 = \frac{b_1 + \left(k-1\right) b_{11} - a_1 - \left(k-1\right) a_{11}}{b_1 + \left(k-1\right) b_{11} - a_1 - \left(k-1\right) a_{11} + a_0 + \left(k-1\right) a_{01}}$$

$$= \frac{b_{12} + k b_{11} - a_{12} - k a_{11}}{b_{12} + k b_{11} - a_{12} - k a_{11} + a_{02} + k a_{01}}$$

Thus, we deduce that $\delta_{\Pi}\left(\alpha, \Gamma\right) := \Pi_0\left(\alpha, \Gamma\right) - \tilde{\Pi}_0\left(\alpha, \Gamma\right)$ is given by:

$$\delta_{\Pi}\left(\alpha, \Gamma\right) = -\frac{k b_{11} \sum_{n=1}^{k-1} \Pi_n\left(\alpha, \Gamma\right)}{b_{12} + k b_{11} - a_{12} - k a_{11} + a_{02} + k a_{01}} \tag{7.38}$$

(Step 2) $\tilde{\mathcal{R}}_c^{ss}$ an upper bound of $\mathcal{R}_c^{ss}$: We prove that this approximation is a close upper bound for the exact stability region of the 3-UEs cellular scenario. We use the following absolute difference between $\mu_s\left(\alpha, \Gamma\right)$ and $\tilde{\mu}_s\left(\alpha, \Gamma\right)$:

$$\delta_{\mu}\left(\alpha, \Gamma\right) := \tilde{\mu}_s\left(\alpha, \Gamma\right) - \mu_s\left(\alpha, \Gamma\right)$$

We know that $\tilde{\mu}_s\left(\alpha, \Gamma\right)$ is deduced from the service rate of the queue Q_s when $\tilde{Q}_{BS}$ is empty ($\tilde{\mu}_s^0\left(\Gamma\right)$) and when $\tilde{Q}_{BS}$ is not empty ($\tilde{\mu}_s^1\left(\alpha, \Gamma\right)$) as follows:

$$\tilde{\mu}_s\left(\tilde{\alpha}, \Gamma\right) = \tilde{\Pi}_0\left(\alpha, \Gamma\right) \tilde{\mu}_s^0\left(\Gamma\right) + \left(1 - \tilde{\Pi}_0\left(\alpha, \Gamma\right)\right) \tilde{\mu}_s^1\left(\alpha, \Gamma\right)$$

$$= \frac{\left(b_1 + \left(k-1\right) b_{11} - a_1 - \left(k-1\right) a_{11}\right) \tilde{\mu}_s^0\left(\Gamma\right) + \left(a_0 + \left(k-1\right) a_{01}\right) \tilde{\mu}_s^1\left(\tilde{\alpha}, \Gamma\right)}{b_1 + \left(k-1\right) b_{11} - a_1 - \left(k-1\right) a_{11} + a_0 + \left(k-1\right) a_{01}} \tag{7.39}$$

Similarly, $\mu_s\left(\alpha, \Gamma\right)$ can be written as follows:

$$\mu_s\left(\alpha, \Gamma\right) = \Pi_0\left(\alpha, \Gamma\right) \mu_s^0\left(\Gamma\right) + \left(1 - \Pi_0\left(\alpha, \Gamma\right)\right) \mu_s^1\left(\alpha, \Gamma\right)$$

Given that Π_0 is related to $\tilde{\Pi}_0$ by the equation (7.38) then:

$$\delta_\mu(\alpha,\Gamma) = \frac{kb_{11}\sum_{n=1}^{k-1}\Pi_n(\alpha,\Gamma)}{b_{12}+kb_{11}-a_{12}-ka_{11}+a_{02}+ka_{01}}\left(\mu_s^0(\Gamma)-\mu_s^1(\alpha,\Gamma)\right)\geq 0 \quad \forall\alpha,\forall\Gamma$$

$$\Rightarrow \mu_s(\alpha,\Gamma)\leq\tilde{\mu}_s(\alpha,\Gamma) \quad \forall\alpha\in[0,1],\forall\ \Gamma\in\Omega_\Gamma^{ss}$$

The highest transmission rate of Q_s, under a given priority policy Γ, corresponds to the case where Q_{BS}, therefore we can write $\mu_s^1(\alpha,\Gamma)\leq\mu_s^0(\Gamma)\ \forall\Gamma,\forall\alpha$. Hence, we prove that $\delta_\mu(\alpha,\Gamma)\geq 0$ for all $\alpha\in[0,1]^4$ and for all $\Gamma\in\Omega_\Gamma$. Thus, for all the fraction vectors α and all the priority policies Γ, the approximated service rate is always higher than the exact service rate. We deduce that $\tilde{\mathcal{R}}_c^{ss}$ is an UB of $\mathcal{R}_c^{ss}$.

(Step 3) Relative distance between $\mathcal{R}_c^{ss}$ and $\tilde{\mathcal{R}}_c^{ss}$: We define the relative distance between $\tilde{\mathcal{R}}_c^{ss}$ and $\mathcal{R}_c^{ss}$ as the maximum relative error between the service rate μ_s and $\tilde{\mu}_s$ over all the priority policies $\Gamma\in\Omega_\Gamma^{ss}$ (with $|\Gamma\in\Omega_\Gamma^{ss}|=6$). We denote by $\epsilon_\mu(\alpha,\Gamma)$ the following expression:

$$\epsilon_\mu(\alpha,\Gamma) := \frac{\tilde{\mu_s}(\alpha,\Gamma)-\mu_s(\alpha,\Gamma)}{\tilde{\mu_s}(\alpha,\Gamma)} = \frac{\frac{kb_{11}\sum_{n=1}^{k-1}\Pi_n(\alpha,\Gamma)}{b_{12}+kb_{11}-a_{12}-ka_{11}+a_{02}+ka_{01}}\left(\mu_s^0(\Gamma)-\mu_s^1(\alpha,\Gamma)\right)\Pi_1(\alpha,\Gamma)}{\frac{(b_{12}+kb_{11}-a_{12}-ka_{11})\mu_s^0(\Gamma)+(a_{02}+ka_{01})\mu_s^1(\alpha,\Gamma)}{b_{12}+kb_{11}-a_{12}-ka_{11}+a_{02}+ka_{01}}}$$

Using the notation introduced here for simplifying equation, we can write:

$$\mu_s^0 = r_1a_{01}+r_2a_{02} = r_2(ka_{01}+a_{02}) \text{ and } \mu_s^1 = r_1a_{11}+r_2a_{12} = r_2(ka_{11}+a_{12})$$

$$\Rightarrow \epsilon_\mu(\alpha,\Gamma) = \frac{kb_{11}\sum_{n=1}^{k-1}\Pi_n(\alpha)(a_{02}+ka_{01}-a_{12}-ka_{11})}{(b_{12}+kb_{11})(a_{02}+ka_{01})} \tag{7.40}$$

We recall that all the parameters in the equations above depend on the fraction vector α and on the priority policy Γ.

The transmission rate of Q_s, under a given priority policy Γ, achieved its highest value when Q_{BS} is empty, therefore we can write $\mu_s^1(\alpha,\Gamma)\leq\mu_s^0(\Gamma)\ \forall\Gamma,\forall\alpha$. Hence, $r_2(ka_{11}(\alpha)+a_{12}(\alpha))\leq r_2(ka_{01}+a_{02})$ thus $(ka_{11}(\alpha)+a_{12}(\alpha))\leq(ka_{01}+a_{02})$. Hence, we prove that $\epsilon_\mu(\alpha,\Gamma)\geq 0$ for all $\alpha\in[0,1]^4$ and for all $\Gamma\in\Omega_\Gamma^{ss}$ which validate the fact that the approximated stability region is an upper bound of the exact stability region of the cellular scenario.

Consider that a corner point of the stability region is achieved at α^* and that the limit of the approximated stability region is achieved at $\tilde{\alpha}^*$ then:

$$\mu_s(\alpha,\Gamma)\leq\mu_s(\alpha^*,\Gamma) \text{ and } \tilde{\mu_s}(\alpha,\Gamma)\leq\tilde{\mu}_s(\tilde{\alpha}^*,\Gamma)\ \forall\alpha\in[0,1]^4$$

We deduce that the service rate relative error $\epsilon_\mu(\alpha,\Gamma)$ for a given priority policy Γ is maximized as follows:

$$\epsilon_\mu(\alpha,\Gamma)\leq\frac{|\mu_s(\alpha^*,\Gamma)-\tilde{\mu}_s(\tilde{\alpha}^*,\Gamma)|}{\tilde{\mu}_s(\tilde{\alpha}^*,\Gamma)} = \frac{|\tilde{\mu}_s(\tilde{\alpha}^*,\Gamma)-\mu_s(\alpha^*,\Gamma)|}{\tilde{\mu}_s(\tilde{\alpha}^*,\Gamma)}$$

$$= \left| \frac{kb_{11} \sum_{n=1}^{k-1} \Pi_n(\Gamma)(a_{12} + ka_{11} - a_{02} - ka_{01})}{(b_{12} + kb_{11})(a_{02} + ka_{01})} \right|_{\tilde{\alpha}^*} := \epsilon^*(\Gamma)$$

We define the relative distance ϵ_R^* between $\mathcal{R}_c^{ss}$ and $\tilde{\mathcal{R}}_c^{ss}$ as the maximum of the relative service rate error $\epsilon^*(\Gamma)$ over all the priority policies $\Gamma \in \Omega_\Gamma^{ss}$. Considering all the corner points of the stability region which corresponds to all $\Gamma \in \Omega_\Gamma^{ss}$, we define the ϵ_R^* as follows :

$$\epsilon_R^* := \max_{\Gamma \in \Omega^{ss}} \epsilon^*(\Gamma) = \max_{\Gamma \in \Omega^{ss}} \left| \frac{kb_{11} \sum_{n=1}^{k-1} \Pi_n(\Gamma)(a_{12} + ka_{11} - a_{02} - ka_{01})}{(b_{12} + kb_{11})(a_{02} + ka_{01})} \right|_{\tilde{\alpha}^*}$$

We deduce the following bound of the exact service rate of μ_s:

$$\epsilon_\mu(\alpha, \Gamma) = \frac{\tilde{\mu_s}(\alpha, \Gamma) - \mu_s(\alpha, \Gamma)}{\tilde{\mu_s}(\alpha, \Gamma)} \leq \epsilon_R^* \Rightarrow \mu_s(\alpha, \Gamma) \geq (1 - \epsilon_R^*)\, \tilde{\mu_s}(\alpha, \Gamma)$$

$$\forall \alpha \in [0, 1] \text{ and } \forall\ \Gamma \in \Omega_\Gamma^{ss}$$

(Step 4) Upper and lower bound of $\mathcal{R}_c^{ss}$: We proved in step 2 that the approximated service rate is an upper bound of the exact service rate hence we can write: $\mu_s(\alpha, \Gamma) \leq \tilde{\mu_s}(\alpha, \Gamma)$; $\forall \alpha \in [0, 1]$ and $\forall\ \Gamma \in \Omega_\Gamma^{ss}$. We proved in step 3 that the relative error between the exact and approximated service rate of Q_s has a maximum value equal to ϵ_R^* hence we can write: $\mu_s(\alpha, \Gamma) \geq (1 - \epsilon_R^*)\, \tilde{\mu_s}(\alpha, \Gamma)$; $\forall \alpha \in [0, 1]$ and $\forall\ \Gamma \in \Omega_\Gamma^{ss}$. We verify that the exact stability region is bounded as it follows:

$$(1 - \epsilon_R^*)\, \tilde{\mathcal{R}}_c^{ss} \subseteq \mathcal{R}_c^{ss} \subseteq \tilde{\mathcal{R}}_c^{ss}$$

These bounds are verified by proving the following two statements:

- When a set of mean arrival rate $(\lambda_s, \lambda_u) \in \mathcal{R}_c^{ss}$ (which means the stability of the exact model of the queues) then $(\lambda_s, \lambda_u) \in \tilde{\mathcal{R}}_c^{ss}$ (which means the stability of the approximated model of the queues).
- When a set of mean arrival rate $(\lambda_s, \lambda_u) \in (1 - \epsilon_R^*)\, \tilde{\mathcal{R}}_c^{ss}$ gives that $(\lambda_s, \lambda_u) \in \mathcal{R}_c^{ss}$.

To do so, we use the same notation as in step 5 of appendix 7.3.1 and we recall that $\mu_i(\Gamma) = \mu_i^0(\Gamma)\, \Pi_0 + \mu_i^1(\Gamma)(1 - \Pi_0)$.

(Step 4.a) $\boldsymbol{\lambda} \in \mathcal{R}_c^{ss} \Rightarrow \boldsymbol{\lambda} \in \tilde{\mathcal{R}}_c^{ss}$: $\boldsymbol{\lambda} \in \mathcal{R}_c^{ss} \Rightarrow$ for each channel $i = s, d, u$ there exists a $\mu_i^* = \sum_\Gamma \Pi_\Gamma \mu_i(\Gamma)$ as combination of $\mu_i(\Gamma)$ for different scheduling policies Γ such that $\lambda_i \leq \mu_i^*$. Given that in ***step 2*** we verified that $\mu(\alpha, \Gamma) \leq \tilde{\mu}(\alpha, \Gamma)$ for all $\alpha \in [0, 1]$ and all $\Gamma \in \Omega_\Gamma^{ss}$, then there exists $\tilde{\mu}_i^* \in \tilde{\mathcal{R}}_c^{ss}$ such that $\mu_i^* \leq \tilde{\mu}_i^*$. Thus, we deduce that $\lambda_i \leq \mu_i^* \leq \tilde{\mu}_i^*$ which means that $\boldsymbol{\lambda} \in \tilde{\mathcal{R}}_c^{ss}$ and that the queues are stable in the approximated model.

(Step 4.b) $\boldsymbol{\lambda} \in (1-\epsilon_R^*)\,\tilde{\mathcal{R}}_c^{ss} \Rightarrow \boldsymbol{\lambda} \in \mathcal{R}_c^{ss}$**:** $\boldsymbol{\lambda} \in (1-\epsilon_R^*)\,\tilde{\mathcal{R}}_c^{ss}$ then for each channel $i = s,d,u$ there exists a $\tilde{\mu}_i^* = \sum_{\Gamma} \Pi_\Gamma \tilde{\mu}_i(\Gamma)$ as combination of $\tilde{\mu}_i(\Gamma)$ for different scheduling policies Γ such that $\lambda_i \leq (1-\epsilon_R^*)\,\mu_i^*$. Given that in ***step 3*** we verified that $\mu(\alpha,\Gamma) \geq (1-\epsilon_R^*)\,\tilde{\mu}(\alpha,\Gamma)$ for all $\alpha \in [0,1]$ and all $\Gamma \in \Omega_\Gamma^{ss}$, then there exists $\mu_i^* \in \mathcal{R}_c^{ss}$ such that $\mu_i^* \geq (1-\epsilon_R^*)\,\tilde{\mu}_i^*$. Thus, we deduce that $\lambda_i \leq (1-\epsilon_R^*)\,\tilde{\mu}_i^* \leq \mu_i^*$ which means that the queues are stable in the exact model and that $\boldsymbol{\lambda} \in \mathcal{R}_c^{ss}$.
From ***Step 4.a*** and ***Step 4.b***, we deduce the following bounds of the exact stability region:

$$(1-\epsilon_R^*)\,\tilde{\mathcal{R}}_c^{ss} \subseteq \mathcal{R}_c^{ss} \subseteq \tilde{\mathcal{R}}_c^{ss}$$

7.3.7 Proof of theorem 3.4.1

We find the exact stability region for the multi-UEs cellular scenario. We pursue the same procedure as the three-UEs cellular scenario in appendix 7.3.1. The main challenge is to find the set of mean arrival rates $\boldsymbol{\lambda}$ that verify the stability of the users queues while guaranteeing the stability conditions of the queues $Q_{i,BS}$ at the BS level (for $1 \leq i \leq K$). The computation of the exact stability region proves the complexity of such procedure and motivates us to search for less complex solutions.

We recall that we have two types of communications: *UE2UE communication* and *UE2BS communications*. The *UE2UE communications* pass thorough the BS such that each *UE2UE communication* i is modeled as the cascade of the queues $\mathrm{Q}_{i,s}$ at the $\mathrm{UE}_{i,s}$ level and $\mathrm{Q}_{i,BS}$ at the BS level (see Fig.3.2). We consider two rates $\{r_1, r_2\}$ with $r_2 = 0$. Assuming that the sources queues are saturated, we want to characterize the stability region of the multi-UEs scenario. Γ is the priority policy according to which the users are prioritized for transmission (i.e. $\Gamma \in \Omega_\Gamma$).

(Step 1) Markov chain model of $\boldsymbol{Q_{i,BS}}$**:** Cellular communications are modeled as coupled processor sharing queues where the service rates of $\mathrm{Q}_{i,s}$ (equivalent to the arrival rate of $\mathrm{Q}_{i,BS}$) as well as the service rate of $\mathrm{Q}_{j,u}$ depend on the state (empty or not) of all the queues $\mathrm{Q}_{k,BS}$ for $1 \leq k \leq K$. Each queue $\mathrm{Q}_{i,BS}$ can be modeled by a simple birth and death Markov chain (see Fig.7.3) which describes the evolution of the queue $\mathrm{Q}_{i,BS}$. The transition probabilities of the Markov chain corresponding to $\mathrm{Q}_{i,BS}$ depend on: (i) the fraction vector $\alpha = (\alpha_1, ..., \alpha_K)$ of all the queues $\mathrm{Q}_{i,BS}$, (ii) the priority policy Γ under which the queues are prioritized, and (iii) the state empty of not of the queues $\mathrm{Q}_{i,BS}$ (for $1 \leq i \leq K$). These probabilities are denoted by:

- Service probability: $b_{i,BS}(\alpha,\Gamma)$.
- Arrival probability when $\mathrm{Q}_{i,BS} = 0$: $a_{i,BS}^0(\alpha,\Gamma)$.
- Arrival probability when $\mathrm{Q}_{i,BS} > 0$: $a_{i,BS}^1(\alpha,\Gamma)$.

(Step 1.a) Arrival and service probabilities of $\boldsymbol{Q_{i,BS}}$**:** We should differ between $a_{i,BS}^0$: arrival probability when $\mathrm{Q}_{i,BS}$ is empty and $a_{i,BS}^1$: arrival probability when

$Q_{i,BS}$ is not empty and these probabilities of arrival are respectively given by:

$$a^0_{i,BS}(\alpha,\Gamma) = p_{i,1} X_i(\alpha,\Gamma) \tag{7.41}$$

$$a^1_{i,BS}(\alpha,\Gamma) = p^i_s\left(\alpha_i p^3_d + \bar{p}^i_d\right) X_i(\alpha,\Gamma) \;\; \forall 1 \le i \le K \tag{7.42}$$

with $X_i(\alpha,\Gamma)$ the probability that the resources are allocated to the i^{th} *UE2UE communication* for a given priority policy Γ and a given fraction vector α:

$$X_i(\alpha,\Gamma) = \prod_{m\in \mathbf{U}_\Gamma\{i\}} \bar{p}^m_u \prod_{n\in \mathbf{K}_\Gamma\{i\}} \bar{p}^n_s \left[\Pi^n_0 + \bar{p}^n_d \bar{\Pi}^n_0\right]$$

with Π^n_0 the probability that the queue $Q_{n,BS}$ is empty and $\bar{\Pi}^n_0$ the probability that this queue is not empty. It will be computed in **Step 1.c**.

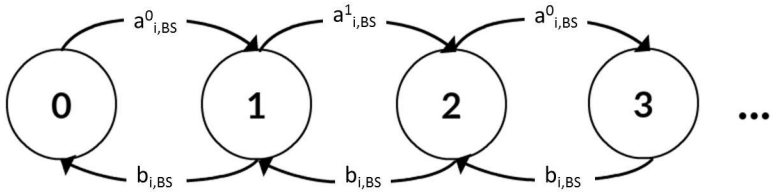


Figure 7.3: Markov chain model of each $Q_{i,BS}$ in the multi-UE scenario

The service probability $b_{i,BS}$ of $Q_{i,BS}$ can be written as follows: (for $1 \le i \le K$)

$$b_{i,BS} = p^i_d\left(1 - \alpha_i p^i_s\right) X_i(\alpha,\Gamma) \tag{7.43}$$

(Step 1.b) Stability condition of $Q_{i,BS}$***:*** Supposing that the arrival and service processes of $Q_{i,BS}$ are strictly stationary and ergodic then the stability of $Q_{i,BS}$ can be determined using **Loynes'** criterion which states that the queue is stable if and only if the average arrival rate is strictly less than the average service rate. Then the stability condition that the fraction parameter α_i should verify in order to satisfy the stability of the queue $Q_{i,BS}$ is:

$$a_{i,BS} < b_{i,BS}$$

Since the arrival probability $a_{i,BS}$ at the queue $Q_{i,BS}$ is given by (7.30) then this condition can be written:

$$b_{i,BS} a^0_{i,BS} b_{i,BS} - a^1_{i,BS} + a^0_{i,BS} < b_{i,BS} \Rightarrow a^1_{i,BS} < b_{i,BS}$$

Since $a^1_{i,BS}$ and $b_{i,BS}$ are given by (7.42) and (7.43) then the queue $Q_{i,BS}$ is stable if the fraction parameter α_i verifies:

$$p^i_s X_i\left(\alpha_i p^i_d + 1 - p^i_d\right) \le p^i_d\left(1 - \alpha_i p^i_s\right) X_i \Rightarrow 2\alpha_i p^i_s p^i_d \le p^i_d - p^i_s + p^i_s p^i_d$$

$$\alpha_i \le \frac{p^i_d - p^i_s + p^i_s p^i_d}{2 p^i_s p^i_d} := \alpha^*_i \tag{7.44}$$

We define the maximum fraction vector $\alpha^* = (\alpha_1^*, \alpha_2^*, ..., \alpha_K^*)$ as the vector containing the maximum fraction parameter α_i^* of all the *UE2UE communications* for which the system is stable. Thus, the stability region is defined by varying the fraction vector α within $[0, \alpha^*]$: $\alpha \prec \alpha^*$ ($\prec$ is a component wise inequality).

(Step 1.c) Probability that $Q_{i,BS}$ is empty: The probability that $Q_{i,BS}$ is empty is deduced from the stationary distribution Π^i of the Markov Chain of $Q_{i,BS}$ (Fig. 7.3) which is deduced from the balance equations. As demonstrated in (7.26), iff $a_{i,BS} < b_{i,BS}$ then the probability that $Q_{i,BS}$ is empty is given by:

$$\Pi_0^i(\alpha, \Gamma) = \frac{b_{i,BS}(\alpha, \Gamma) - a_{i,BS}^1(\alpha, \Gamma)}{b_{i,BS}(\alpha, \Gamma) - a_{i,BS}^1(\alpha, \Gamma) + a_{i,BS}^0(\alpha, \Gamma)} \; \forall 1 \leq i \leq K \tag{7.45}$$

$$\Rightarrow \Pi_0^i = \frac{-2\alpha_i p_s^i p_d^i + p_d^i - p_s^i \bar{p}_d^i}{-2\alpha_i p_s^i p_d^i + p_d^i + p_s^i p_d^i} \tag{7.46}$$

$$X_i(\alpha, \Gamma) = \prod_{m \in \mathbf{U}_\Gamma\{i\}} \bar{p}_u^m \prod_{n \in \mathbf{K}_\Gamma\{i\}} \bar{p}_s^n \left[1 + \frac{p_s^n p_d^n}{2\alpha_n^* p_s^n p_d^n - (1 + p_s^n) p_d^n}\right] \tag{7.47}$$

(Step 2) Service rates of $Q_{i,s}$ and $Q_{j,u}$: We compute the exact stability region of the multi-UEs cellular scenario. For that, we compute the service rate of each *UE2UE* and *UE2BS communications* for a given priority policy Γ. For brevity, we present the results for the i^{th} *UE2UE communication* and j^{th} *UE2BS communication*. These results are applied for all $1 \leq i \leq K; 1 \leq j \leq U$. We follow the procedure below, based on queuing theory analysis of the network capacity, in order to derive the performance of the system of queues.

We start by computing the service rate of $\mathrm{Q}_{i,s}$. If $\mathrm{Q}_{i,BS}$ is empty then the service rate of $\mathrm{Q}_{i,s}$ is denoted by $\mu_{i,s}^0$ and by $\mu_{i,s}^1$ otherwise then the service rate of $\mathrm{Q}_{s,1}$ is computed as follows:

$$\begin{aligned}\mu_{i,s}(\alpha, \Gamma) &= \mathbb{P}[Q_{i,BS} = 0]\, \mu_{i,s}^0(\alpha, \Gamma) + \mathbb{P}[Q_{i,BS} > 0]\, \mu_{i,s}^1(\alpha, \Gamma) \\ &= \Pi_0^i \mu_{i,s}^0(\alpha, \Gamma) + \left[1 - \Pi_0^i\right] \mu_{i,s}^1(\alpha, \Gamma)\end{aligned} \tag{7.48}$$

with $\mu_{i,s}^0$ and $\mu_{i,s}^1$ given by:

$$\mu_{i,s}^0(\alpha, \Gamma) = \mathbb{E}\left[\mu_{i,s}(S(t), \Gamma, Q(t) | Q_{i,BS}(t) = 0)\right]$$

$$\mu_{i,s}^1(\alpha, \Gamma) = \mathbb{E}\left[\mu_{i,s}(S(t), \Gamma, Q(t) | Q_{i,BS}(t) > 0)\right]$$

Based on (7.47), $\mu_{i,s}^0(\alpha, \Gamma)$ and $\mu_{i,s}^1(\alpha, \Gamma)$ can be written as follows for all $1 \leq i \leq K$:

$$\mu_{i,s}^0(\alpha, \Gamma) = r_1 p_s^i \prod_{m \in \mathbf{U}_\Gamma\{i\}} \bar{p}_u^m \prod_{n \in \mathbf{K}_\Gamma\{i\}} \bar{p}_s^n \left[1 + \frac{p_s^n p_d^n}{2\alpha_n p_s^n p_d^n - (1 + p_s^n) p_d^n}\right] \tag{7.49}$$

$$\mu_{i,s}^1(\alpha, \Gamma) = r_1 p_s^i \left(\alpha_i p_d^1 + 1 - p_d^1\right) \prod_{n \in \mathbf{K}_\Gamma\{i\}} \bar{p}_s^n \left[1 + \frac{p_s^n p_d^n}{2\alpha_n p_s^n p_d^n - (1 + p_s^n) p_d^n}\right] \tag{7.50}$$

Substituting all the Π_0^i by their values from equation (7.46) and based on the above expressions of $\mu_{i,s}^0$ and $\mu_{i,s}^1$, equation (7.48) gives that the formula (3.17) that describes the service rate $\mu_{i,s}(\alpha, \Gamma)$ of the i^{th} *UE2UE communications* in the multi-UE cellular scenario for a given fraction vector α and a given priority policy Γ.

Now, we compute the service rate of $Q_{j,u}$. For a given policy Γ, the service rate of each $Q_{j,u}$ with $1 \leq j \leq U$ depends also on the state (empty or not) of all the queues $Q_{m,BS}$ for $m \in \mathbf{K}_\Gamma\{j+K\}$ and is computed as follows for all $1 \leq j \leq U$:

$$\mu_{j,u} = \mathbb{E}\left[\mu_{j,u}\left(S(t), \Gamma, Q(t)|Q_{BS}(t)\right)\right] = r_1 p_{j,u} X_{j+K}(\alpha, \Gamma)$$

$$= r_1 p_{j,u} \prod_{m \in \mathbf{U}_\Gamma\{j+K\}} \bar{p}_u^m \prod_{n \in \mathbf{K}_\Gamma\{j+K\}} \bar{p}_s^n \left[\Pi_0^n + \bar{p}_d^n \bar{\Pi}_0^n\right]$$

Substituting all Π_0^i by their values from equation (7.46) we can verify that the service rate $\mu_{i,s}(\alpha, \Gamma)$ of the j^{th} *UE2BS communications* in the Multi-UEs cellular scenario is given by (3.17) for a given fraction vector α and a given priority policy Γ.

(Step 3) Characterization of the multi-UEs cellular stability region: Combining the results of the previous steps, we deduce the exact stability region of the multi-UEs scenario. Supposing that the arrival and service processes of the queues are strictly stationary and ergodic then their stability which is determined using **Loyne's** criterion is given by the condition that the average arrival rate is smaller than the average service rate.

The following procedure is pursue for capturing the stability region of the scenario. We consider the set of all the possible priority policies Ω_Γ with $|\Omega_\Gamma| = (K+U)!$). Then, for each priority policy, we vary $\boldsymbol{\alpha} \in [0, \boldsymbol{\alpha}^*]$ in order to guarantee the stability of the $Q_{i,BS}$ queues (with $1 \leq i \leq K$). For each value of α, we find the probabilities that $Q_{i,BS}$ (with $1 \leq i \leq K$) are empty in order to deduce the service rates of the queues in the system. This procedure is applied for all the priority policies $\Gamma \in \Omega_\Gamma$.

Therefore, the stability region for the multi-UEs cellular scenario is characterized by the set of mean arrival rates $\boldsymbol{\lambda} \in \mathcal{R}_c$ such that:

$$\mathcal{R}_c = co\left(\bigcup_{\Gamma \in \Omega_\Gamma} \bigcup_{\alpha \in [0, \alpha^*]} \{\boldsymbol{\mu}(\boldsymbol{\alpha}, \Gamma)\}\right)$$

where the elements of $\boldsymbol{\mu}(\Gamma)$ which are $\mu_{i,s}(\Gamma)$ and $\mu_{j,u}(\Gamma)$ (for $1 \leq i \leq K$, $1 \leq j \leq U$) are respectively given by (3.17) and (3.18).

(Step 4) Arrival rates within the stability region is equivalent to the stability of the system of queues: Proving that having a set of mean arrival rates $\boldsymbol{\lambda}$ in $\tilde{\mathcal{R}}_c$ is equivalent to the stability of the system queue is identical to the demonstration done in step 5 of the appendix 7.3.1. For brevity, we remove this proof to avoid repetition.

7.3.8 Proof of theorem 3.4.2

In order to reduce the computation complexity of the stability region for multi-UEs cellular scenario, we start by reducing the set of α that is studied for characterizing this stability region. We prove that considering only the border values of each fraction parameter α_i (which corresponds to 0 and α_i^* given by equation (7.44)) is sufficient for characterizing the stability region $\mathcal{R}_c$. Hence, the procedure for computing the stability region $\mathcal{R}_c$, as shown in appendix 7.3.7, remains unaffected thus for brevity only the modifications will be presented in this proof.

Here, we prove that there is no need to vary each $\alpha_i \in [0, \alpha_i^*]$ for finding the stability region $\mathcal{R}_c$; but it is ample to consider only the border values of $\alpha_i \in \{0, \alpha_i^*\}$. For each α_i (for all $1 \leq i \leq K$), only these two values $\{0, \alpha_i^*\}$ have to be taken into account, thus the computation complexity of the stability region $\mathcal{R}_c$ decreases largely due to the fact that there is no need anymore to vary α_i within $\in [0, \alpha_i^*]$. Let us verify that for each fraction parameter α_i (that corresponds to the i^{th} *UE2UE communication*), it is sufficient to consider these two values of α_i: $\{0, \alpha_i^*\}$ for characterizing the exact stability region $\mathcal{R}_c$. We prove that, for all $1 \leq i \leq K$, the service rate vector $\boldsymbol{\mu}(\alpha_1, .., \alpha_i, ..., \alpha_K, \Gamma)$ can be written as a convex combination of $\boldsymbol{\mu}(\alpha_1, .., 0, ..., \alpha_K, \Gamma)$ and $\boldsymbol{\mu}(\alpha_1, .., \alpha_i^*, ..., \alpha_K, \Gamma)$. This approach is applied separately over all the fraction parameters α_i (with $1 \leq i \leq K$). The mathematical form of the convex combination consists of verifying that for each $\alpha_i \in [0, \alpha_i^*]$ there exists a $\gamma_i \in [0, 1]$ such that:

$$\boldsymbol{\mu}(\alpha_1, .., \alpha_i, ..., \alpha_K, \Gamma) = \gamma_i \boldsymbol{\mu}(\alpha_1, .., 0, ..., \alpha_K, \Gamma) + (1 - \gamma_i)\boldsymbol{\mu}(\alpha_1, .., \alpha_i^*, ..., \alpha_K, \Gamma) \tag{7.51}$$

Expressing $\boldsymbol{\mu}(\alpha_1, .., \alpha_i, ..., \alpha_K, \Gamma)$ as a convex combination of $\boldsymbol{\mu}(\alpha_1, .., 0, ..., \alpha_K, \Gamma)$ and $\boldsymbol{\mu}(\alpha_1, .., \alpha_i^*, ..., \alpha_K, \Gamma)$ means that, for all $\alpha_i \in [0, \alpha_i^*]$, the service vector $\boldsymbol{\mu}$ is an interior point of the convex hull $co\{\boldsymbol{\mu}(\alpha_1, .., 0, ..., \alpha_K, \Gamma), \boldsymbol{\mu}(\alpha_1, .., \alpha_i^*, ..., \alpha_K, \Gamma)\}$. Hence, these two corner points $\boldsymbol{\mu}(\alpha_1, .., 0, ..., \alpha_K, \Gamma)$ and $\boldsymbol{\mu}(\alpha_1, .., \alpha_i^*, ..., \alpha_K, \Gamma)$ are sufficient for describing the convex hull over all the points $\boldsymbol{\mu}(\alpha_1, .., \alpha_i, ..., \alpha_K, \Gamma)$ with α_i varying within $[0, \alpha_i^*]$.

In order to find the value of γ_i, we start by analyzing the expressions of $\boldsymbol{\mu}$ and especially how the parameter α_i affects this formula. Indeed, the service rate of each communication in the system is affected by the parameter α_i in one of the two following forms:

$$\textbf{First form}: 1 + \frac{(1 - p_{i,s})\, p_{i,d}}{-2\alpha_i p_{i,s} p_{i,d} + (1 + p_{i,s})\, p_{i,d}} := F_1(\alpha_i)$$

$$\textbf{Second form}: 1 - \frac{p_{i,s} p_{i,d}}{-2\alpha_i p_{i,s} p_{i,d} + (1 + p_{i,s})\, p_{i,d}} := F_2(\alpha_i)$$

The first form $F_1(\alpha_i)$ appears only at the level of the service rate of the i^{th} *UE2UE communication* ($\mu_{1,i}(\boldsymbol{\alpha}, \Gamma)$) while the second form $F_2(\alpha_i)$ appears at the level of the service rate of all the other communications. $\boldsymbol{\mu}(\boldsymbol{\alpha}, \Gamma)$ can be written as in equation (7.51) if and only if both forms ($F_1(\alpha_i)$ and $F_2(\alpha_i)$) respectively verified the equation (7.51). In other words, we need to verify that $F_1(\alpha_i)$ (respectively $F_2(\alpha_i)$) can be

written as a convex combination of $F_1(0)$ and $F_1(\alpha_i^*)$ (respectively $F_2(0)$ and $F_2(\alpha_i^*)$) with γ_i as coefficient:

$$F_1(\alpha_i) = \gamma_i F_1(0) + (1-\gamma_i) F_1(\alpha_i^*) \text{ and } F_2(\alpha_i) = \gamma_i F_2(0) + (1-\gamma_i) F_2(\alpha_i^*)$$

Knowing that $\alpha \in [0, \alpha_i^*]$, we can write $\alpha_i = \lambda \alpha_i^*$ with $\lambda \in [0,1]$. Let us consider the following notation:

$$\gamma_i = \frac{(1+p_{i,s})\, p_{i,d}\,(1-\lambda)}{(1+p_{i,s})\, p_{i,d} - \lambda\,(p_{i,s}p_{i,d} + p_{i,d} - p_{i,s})} \tag{7.52}$$

We show that $0 \leq \gamma_i \leq 1$. First, both numerator and denominator are positives then $\gamma_i \geq 0$. Second,

$$-\lambda\,(1+p_{i,s})\, p_{i,d} \leq -\lambda\,(1+p_{i,s})\, p_{i,d} + \lambda p_{i,s}$$

$$\Rightarrow (1+p_{i,s})\, p_{i,d}\,(1-\lambda) \leq (1+p_{i,s})\, p_{i,d} - \lambda\,(p_{i,d} + p_{i,s}p_{i,d} - p_{i,s}) \Rightarrow \gamma_i \leq 1$$

Knowing that γ_i (given by (7.52)) is within $[0,1]$, we can easily verify the following:

$$F_1(\alpha_i) = \gamma_i F_1(0) + (1-\gamma_i) F_1(\alpha_i^*) \text{ and } F_2(\alpha_i) = \gamma_i F_2(0) + (1-\gamma_i) F_2(\alpha_i^*)$$

where $F_1(0) = 1 + \frac{\bar{p}_{i,s}}{1+p_{i,s}}$, $F_1(\alpha_i^*) = 1 + \frac{\bar{p}_{i,s}p_{i,d}}{p_{i,s}}$, $F_2(0) = 1 - \frac{p_{i,s}}{1+p_{i,s}}$ and $F_1(\alpha_i^*) = \bar{p}_{i,d}$.

Hence, for all $1 \leq i \leq K$ the service rate vector $\boldsymbol{\mu}$ can be written as in equation (7.51) with γ_i given by (7.52). Therefore, we show that taking the convex hull of the vector $\boldsymbol{\mu}$ over all $\alpha_i \in [0, \alpha_i^*]$ is equivalent to considering the convex hull over $\alpha_i \in \{0, \alpha_i^*\}$. This remains true for all the $1 \leq i \leq K$.

The same procedure, as in appendix 7.3.7, is pursued for capturing the less complex stability region of the scenario: (i) we consider the set of all the possible priority policies Ω_Γ with $|\Omega_\Gamma| = (K+U)!$), (ii) for each priority policy, we vary $\alpha_i \in \{0, \alpha_i^*\}$ (with $1 \leq i \leq K$), (iii) for each value of α, we find the probabilities that $\mathrm{Q}_{i,BS}$ (with $1 \leq i \leq K$) are empty in order to deduce the service rates of the queues in the system, (iv) this procedure is applied for all the priority policies $\Gamma \in \Omega_\Gamma$. The only difference with that in appendix 7.3.7 is at the level of the step (ii) where we consider a limited set of values of the fraction vector $\boldsymbol{\alpha}$ compared to that in appendix 7.3.7. We see that the complexity in this case is reduced to: $2^K\,(K+U)!$.

For the multi-UEs cellular scenario, the stability region $\mathcal{R}_c$ is simplified as follows:

$$\mathcal{R}_c = co\left(\bigcup_{\Gamma\in\Omega_\Gamma}\bigcup_{\alpha\in\mathbb{S}_\alpha}\{\boldsymbol{\mu}(\boldsymbol{\alpha},\Gamma)\}\right)$$

$$\text{with } \mathbb{S}_\alpha = \{\boldsymbol{\alpha} \,|\, \alpha_i \in \{0, \alpha_i^* = \frac{p_{i,d} - p_{i,s} + p_{i,s}p_{i,d}}{2p_{i,s}p_{i,d}}\}\ \forall 1 \leq i \leq K\}$$

where the elements of $\boldsymbol{\mu}(\Gamma)$ which are $\mu_{i,s}(\Gamma)$ and $\mu_{j,u}(\Gamma)$ (for $1 \leq i \leq K$, $1 \leq j \leq U$) are respectively given by (3.17) and (3.18).

7.3.9 Proof of theorem 3.4.5

Similarly to the proof in appendix 7.3.8, the demonstration here follows the same steps as in the appendix 7.3.7. The only difference is at the level of the set of the considered priority policies that will be studied for characterizing the stability region. Actually, in appendix 7.3.7 we consider all the possible priority policies $\Gamma \in \Omega_\Gamma$ with $|\Omega_\Gamma| = (K+U)!$ while in this proof we reduce the subset of the considered priority policies to $\Omega_\Gamma^{K_0} \subset \Omega_\Gamma$ where $|\Omega_\Gamma^{K_0}| = \frac{(K+U)!}{(K+U-K_0)!}$ with $K_0 <<<< K+U$.We denote by $p_s = p_{i,s} = p_{j,u}$ and $p_d = p_{i,d}$ overall $1 \leq i \leq K$ and $1 \leq j \leq U$.

From appendix 7.3.8 we can see that finding the stability region of the cellular symmetric case involves the consideration of $|\Omega_\Gamma| = (K+U)!$ policies and 2^K values of the fraction vector. On the aim of reducing this $2^K\,(K+U)!$ complexity, we remark that there exists a $K_0 \in \mathbb{N}$ such than when the user has a priority less than K_0 then the improvement in terms of performance (i.e. service rate of its queue) is small. Hence, we avoid the search among all the priority policies Ω_Γ and we limit the search to the set $\Omega_\Gamma^{K_0}$ which consists of the priority policies of all the subsets of K_0 elements within the $K+U$ communications. The computational complexity is reduced to $2^{K_0}\frac{(K+U)!}{(K+U-K_0)!}$.

Considering ϵ as the preciseness parameter, we pursue the following procedure for finding the $\epsilon-$ approximation of the stability region of the multi-UE symmetric case: **step 1** gives the best service rate that a user can have when it is at the k^{th} priority level, **step 2** deduces the priority level $K_0\,(\epsilon)$ (as function of the percentage of preciseness ϵ) such that beyond this priority level the error of the users' service rate is bounded by ϵ; this enables us to deduce the set of the priority policies $\Omega_\Gamma^{K_0}$ that is sufficient for describing the ϵ-approximation of the stability region, ***step 3*** combines the results of the previous steps to provide the ϵ close upper bound of the stability region of the symmetric case of the multi-UEs scenario.

(Step 1) The best service rate of a user at a priority level k***:*** We consider the symmetric case of the multi-UE scenario with K *UE2UE communications* and U *UE2BS communications* ($2K+U$ users). A user with a priority level k has, at the best case, a service rate equals to:

$$\mu^k = r_1 p_s \bar{p}_s^{k-1} \tag{7.53}$$

This service rate corresponds to the most optimistic case where the k^{th} user as well as all the more prioritized UEs ($1 \leq i \leq k$) have *UE2BS communications*. Actually, considering that the user at priority level k has K_1 more prioritized *UE2UE communications* and K_2 more prioritized *UE2BS communications* then from (3.17) and (3.18) we deduce the following best service rates:

- If the k^{th} priority level corresponds to a *UE2UE communication* then its most optimist service rate corresponds to $\alpha_k = \alpha_k^*$ and $\alpha_i = 0$ for all $1 \leq i \leq K_1$:

$$\mu_s^k = \frac{1}{2} r_1 p_s \left(1 + \frac{\bar{p_s} p_d}{p_s}\right) \bar{p}_s^{K_1+K_2} \left(\frac{1}{1+p_s}\right)^{K_1}$$

- If the k^{th} priority level corresponds to a *UE2BS communication* then its most optimist service rate corresponds to $\alpha_i = 0$ for all $1 \leq i \leq K_1$:

$$\mu_u^k = r_1 p_s \bar{p}_s^{K_1+K_2} \left(\frac{1}{1+p_s}\right)^{K_1}$$

A more optimist case corresponds to the one where all the more prioritized users having *UE2BS communications*, which means $K_1 = k-1$ and $K_2 = 0$.

- If the k^{th} is participating to a *UE2UE communication* then its most optimist service rate is:

$$\mu_s^k = \frac{1}{2} r_1 p_s \left(1 + \frac{\bar{p}_s p_d}{p_s}\right) \bar{p}_s^{k-1}$$

- If the k^{th} is participating to a *UE2BS communication* then its most optimist service rate is:

$$\mu_u^k = r_1 p_s \bar{p}_s^{k-1}$$

It is obvious that $\mu_u^k \geq \mu_s^k$ since $\frac{1}{2}\left(1 + \frac{\bar{p}_s p_d}{p_s}\right) \leq 1$. Hence, the service rate of a user at a priority level k (denoted by μ^k) is maximized as follows:

$$\mu^k \leq r_1 p_s \bar{p}_s^{k-1} := \mu^{k*}$$

We deduce that the service rate of any user, at priority level k, is lower or equal to $\mu^{k*} := r_1 p_s \bar{p}_s^{k-1}$. Our approximation is based on finding the priority level K_0 such that the maximum service rate $\mu^{K_0*} = r_1 p_s \bar{p}_s^{K_0-1}$ at this priority level is lower than ϵ. Then, we assume that the service rates of the communications that have a higher priority level (greater than K_0) are equal to zero.

(Step 2) The priority policies subset $\Omega_\Gamma^{K_0}$ as function of ϵ: The level K_0 corresponds to the priority policy level at which the service rate is negligible which means lower than ϵ. Knowing that the service rate of a user at a priority level k is bounded by equation (7.53)), hence for the priority level K_0 we can write:

$$\mu^{K_0*} = r_1 p_s \bar{p}_s^{K_0-1} \leq \epsilon \Rightarrow K_0 = \left\lceil 1 + \frac{\log\left(\frac{\epsilon}{r_1 p_s}\right)}{\log\left(\bar{p}_s\right)} \right\rceil$$

We should be aware that K_0 is limited by the total number of communications $K+U$.

$$K_0 = min\left\{K+U, \left\lceil 1 + \frac{\log\left(\frac{\epsilon}{r_1 p_s}\right)}{\log\left(\bar{p}_s\right)} \right\rceil\right\} \tag{7.54}$$

We define $\Omega_\Gamma^{K_0}$ as follows:

- We consider all the possible sets of K_0 communications within the all existing $K+U$ communications. In total, there is $\frac{(K+U)!}{(K+U-K_0)!K_0!}$ subsets of K_0 elements.

- Then for each of the above subsets, we find all the possible priority sorting of these K_0 communications. We know that $K_0!$ is the number of the possible permutation of the K_0 elements of each subset.

We deduce that $\Omega_{\Gamma}^{K_0}$, the set of all the possible priority policies of all the subset of K_0 elements within the total of $K+U$ communications, consists of $\frac{(K+U)!}{(K+U-K_0)!}$ elements.

(Step 3) Characterizing the approximated stability region: The same procedure, as in appendix 7.3.8, is pursued for capturing the ϵ-approximated stability region of multi-UEs symmetric case. However, here the set of the studied priority policies is restricted in order to reduce the complexity and proposing an ϵ close upper bound of the exact stability region. We proceed as follows : (i) for a given ϵ (for a given precision), we find K_0 from equation (7.54) and we deduce the subset $\Omega_{\Gamma}^{K_0}$ of the required priority policies for characterizing the ϵ upper bound of the exact stability region (see ***step 2*** where $|\Omega_{\Gamma}^{K_0}| = \frac{(K+U)!}{(K+U-K_0)!}$), (ii) for each priority policy $\Gamma \in \Omega_{\Gamma}^{K_0}$, we vary $\alpha_i \in \{0, \alpha_i^*\}$ (with $1 \leq i \leq K$), (iii) for each value of α, we find the probabilities that the queues $Q_{i,BS}$ (with $1 \leq i \leq K$) are empty in order to deduce the service rates of the queues in the system. The main difference between this procedure and the one in appendix 7.3.8 is at the level of the step (i) where we consider a limited set of priority policies $\Gamma \in \Omega_{\Gamma}^{K_0}$ and not all the possible priority policies $\Omega_{\Gamma}^{K_0}$. For these priority policies we consider that the service rates of the communications, that have a priority level higher than K_0, are equal to zero.

We define μ_i^k as the service rate of the i^{th} communication (with $1 \leq i \leq K+U$) when it has a priority level k. Hence, the approximated model (with $\tilde{\boldsymbol{\mu}}$ denotes the service rate vector) consists of considering the following:

- $\tilde{\mu}_i^k = \mu_i^k$ for $1 \leq i \leq K+U$ and for all $k \leq K_0$.
- $\tilde{\mu}_i^k = 0$ for $1 \leq i \leq K+U$ and for all $k \geq K_0$.

Thus, we can deduce that the service rate of all the communications ($1 \leq i \leq K+U$) for any priority policy $\Gamma \in \Omega_{\Gamma}^{K_0}$ is bounded as follows:

$$\tilde{\mu}_i \leq \mu_i \leq \tilde{\mu}_i + \epsilon \tag{7.55}$$

The above expression can be proved by verifying that it remains true for both cases: (i) priority levels $k \leq K_0$ and priority levels $k \geq K_0$.

A similar analysis as the one done in ***step 4*** of appendix 7.3.6, let us deduce the bounds of the stability region based on the bounds of the service rates (equation (7.55)). Thus, for the symmetric case of the multi-UEs cellular scenario, the exact stability region can be bounded as follows:

$$\tilde{\mathcal{R}}_c \subseteq \mathcal{R}_c \subseteq \tilde{\mathcal{R}}_c + \epsilon$$

Where the ϵ close upper bound of the exact stability region $\tilde{\mathcal{R}}_c$ is given by the set of $\boldsymbol{\lambda} \in \tilde{\mathcal{R}}_c$ such that:

$$\tilde{\mathcal{R}}_c = co\left(\bigcup_{\Gamma \in \Omega_{\Gamma}^{K_0}} \bigcup_{\alpha \in \mathbb{S}_{\alpha}^{K_0}} \{\boldsymbol{\mu}(\boldsymbol{\alpha}, \Gamma)\}\right)$$

with

$$K_0 = min\left\{K+U, \left\lceil 1+\frac{\log\left(\frac{\epsilon}{r_1 p_s}\right)}{\log\left(\bar{p}_s\right)}\right\rceil\right\}$$

where $\Omega_{\Gamma}^{K_0}$ is the set of the feasible priority policies of K_0 communications among all the $K+U$ communications, $\mathbb{S}_{\alpha}^{K_0}$ is the set of the fraction vector $\boldsymbol{\alpha}$ where each element $\alpha_i \in \{0, \alpha_i^*\}$ and the elements of $\boldsymbol{\mu}(\Gamma)$ which are $\mu_{i,s}(\Gamma)$ and $\mu_{j,u}(\Gamma)$ (for $1 \leq i \leq K$, $1 \leq j \leq U$) are respectively given by (3.17) and (3.18).

7.3.10 Proof proposition 3.4.6

For the multi-UEs D2D scenario, direct UE2UE communications are considered. We apply the classical methodology for computing the stability region. Based on queuing theory analysis of the network capacity, we derive for a given policy Γ, the following analytic expressions of respectively the service rates of the i^{th} UE2UE queue (with $1 \leq i \leq K$) as well as the j^{th} UE2BS queue (with $1 \leq j \leq U$):

$$\mu_{i,s}^{d}(\Gamma) = \mathbb{E}\left[\mu_{i,s}^{d}\left(S(t), \Gamma\right)\right] = r_1 p_{sd}^{i} X_i^{d}(\Gamma)$$

$$\mu_{j,u}^{d}(\Gamma) = \mathbb{E}\left[\mu_{j,u}^{d}\left(S(t), \Gamma\right)\right] = r_1 p_{u}^{j} X_i^{d}(\Gamma)$$

with $X_i^{d}(\Gamma)$ is the probability that the chain i transmits in the D2D scenario which means it is the probability that none of the more prioritized chain are transmitting which gives:

$$X_i^{d}(\Gamma) = \prod_{n \in \mathbf{K}_{\Gamma}\{i\}} \bar{p}_n^{sd} \prod_{m \in \mathbf{U}_{\Gamma}\{i\}} \bar{p}_m^{u} \tag{7.56}$$

Based on (7.56), we can verify that for a given priority policy Γ, the service rates of the UE2UE communication as well as that of the UE2BS communications for the multi-UEs D2D scenario are respectively given by (3.19) and (3.20). Hence, for the multi-UEs D2D scenario,the stability region for $\mathcal{R}_c$ can be simplified as follows:

$$\mathcal{R}_d = co\left(\bigcup_{\Gamma \in \Omega_{\Gamma}} \{\boldsymbol{\mu}^{d}(\Gamma)\}\right)$$

where the elements of $\boldsymbol{\mu}^{d}(\Gamma)$ which are $\mu_{i,s}^{d}(\Gamma)$ and $\mu_{j,u}^{d}(\Gamma)$ (for $1 \leq i \leq K$, $1 \leq j \leq U$) are respectively given by (3.19) and (3.20).

7.4 D2D Resource Allocation Proofs

7.4.1 Proof of proposition 2 and 3

We denote by Γ^{cent} the proposed centralized scheduling and by Γ_c^* the *best* centralized scheduling that optimizes OP (4.2) in a limited-feedback D2D network. In this case, BS has the global statistical knowledge of the D2D CSI but has the instantaneous CSI of a subset of user (i.e. number of users in this subset is limited to the number of resources available for CSI reporting).

Lyapunov technique (see [110]) is applied for verifying proposition 3 that compares the time average of users' power consumption of the proposed centralized algorithm Γ^{cent} to that of the *best* centralized limited-feedback scheduling Γ_c^*. In order to solve the optimization problem (4.2), we proposed to introduce the following virtual queues (given by equation (4.3)).

$$Q_n(t+1) = (Q_n(t) - R_n(t))^+ + R_{th}$$

From queuing theory [31], it is known that strong stability of the queues means that the average arrival rate is less than the average departure rate. Using the above definition of the virtual queue given, we can establish the equivalence between the throughput constraint of the optimization problem and the strong stability of the system of virtual queues. It means that the goal is to minimize the energy consumption of the D2D network while guaranteeing the stability of the queues. The proof is based on Lyapunov optimization theory. We first define the Lyapunov function $L_{\boldsymbol{Q}}(t)$ given by the following:

$$L_{\boldsymbol{Q}}(t) := \frac{1}{2}\sum_{i=1}^{\mathrm{N}} Q_i^2(t)$$

We denote by $\boldsymbol{Q}(t) = (Q_1(t), ..., Q_N(t))$ the vector of the current virtual queue backlogs and by $\boldsymbol{H}(t) = [h_1, h_2, ..., h_N]$ the vector of channel states. Applying the methodology in [110] gives that the drift-plus-penalty expression is upper bounded by the following:

$$\Delta(\boldsymbol{Q}(t)) := \Delta(\boldsymbol{L}(t)) + V\sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[P_i(t)\,|\boldsymbol{Q}(t)\right]$$

$$\leq C + \sum_{i=1}^{\mathrm{N}} Q_i(t) R_{th} - \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} Q_i(t) R_i\left(\Gamma^{cent}(t), \boldsymbol{H}(t)\right)|\boldsymbol{Q}(t)\right]$$

$$+ V\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} P_i\left(\Gamma^{cent}(t), \boldsymbol{H}(t)\right)|\boldsymbol{Q}(t)\right] \tag{7.57}$$

where $\Gamma^{cent}(t)$ is the proposed centralized scheduling policy and C is a finite constant such that

$$\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} \frac{R_{th}^2 + R_i^2(t)}{2}|\boldsymbol{Q}(t)\right] \leq C$$

At each time-slot t, the proposed algorithm computes, based on global statistical CSI, the subset Λ^* of users that will transmit its CSI feedback (given by equation (4.5)). Then, the scheduled user is the one that has the lowest value $v_i(t)$. Based on equation (4.5) of the chosen subset Λ^*, the proposed centralized algorithm $\Gamma(t)$ verifies the inequality (7.58) for any scheduling policy $\Gamma(t)$ including the optimal policy of (4.2) $\Gamma_c^*(t)$ when a centralized limited-feedback scenario is considered.

$$\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}}\left[VP_i\left(\Gamma^{cent}(t),\boldsymbol{H}(t)\right)-Q_i(t)R_i\left(\Gamma^{cent}(t),\boldsymbol{H}(t)\right)\right]|\boldsymbol{Q}(t)\right]$$

$$\leq \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}}\left[VP_i\left(\Gamma(t),\boldsymbol{H}(t)\right)-Q_i(t)R_i\left(\Gamma(t),\boldsymbol{H}(t)\right)\right]|\boldsymbol{Q}(t)\right] \tag{7.58}$$

The optimal policy $\Gamma_c^*(t)$ verifies the OP (4.2) which implies that the corresponding throughput constraint is satisfied, thus this policy guarantees the stability of the virtual queues. Supposing that the arrival rate vector of these virtual queues is interior to the stability region of the system of queue and $\epsilon > 0$, we have:

$$\mathbb{E}\left[R_i\left(\Gamma_c^*(t)\right)|\boldsymbol{Q}(t)\right]=\mathbb{E}\left[R_i\left(\Gamma_c^*(t)\right)\right]\geq R_{th}+\epsilon$$

Same methodology as in [110] gives the following upper bound of the total average backlogs of the virtual queues:

$$\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{\mathrm{N}}\mathbb{E}\left[Q_i(t)\right]\leq\frac{C+B}{\epsilon} \tag{7.59}$$

with B a finite constant such that $\limsup_{T\to\infty}\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{\mathrm{N}}\mathbb{E}\left[VP_i\left(\Gamma_c^*(t)\right)\right]\leq B$. Thus, proposition 2 is verified and all the virtual queues in the system are strongly stable. Hence, when the arrival rate at the virtual queues is less than its average departure rate then the proposed centralized scheduling satisfies the throughput constraint in (4.2). Pursuing with the Lyapunov optimization of queuing networks leads to the proposition 3 as follows:

$$\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{\mathrm{N}}\mathbb{E}\left[P_i\left(\Gamma^{cent}(t)\right)\right]\leq\frac{C}{V}+\frac{\mathbb{E}\left[L_{\boldsymbol{Q}}(0)\right]}{VT}+\frac{1}{T}\sum_{t=0}^{T-1}\sum_{i=1}^{\mathrm{N}}\mathbb{E}\left[P_i\left(\Gamma_c^*(t)\right)\right]$$

Therefore, the proposed centralized algorithm achieves the performance of the *best* solution of the limited-feedback problem (4.2) with a distance of $O\left(\frac{1}{V}\right)$ and a time average queue backlog of $O(V)$.

7.4.2 Proof of proposition 4

We denote by Γ^{dist} the proposed distributed scheduling and by Γ_{id}^* the *ideal* scheduling policy that optimizes problem (4.2) while assuming the global knowledge of the instantaneous CSI of D2D links.

Lyapunov technique (see [110]) is applied for verifying proposition 4 that compares the time average of users' power consumption of the distributed algorithm Γ^{dist} to that of the *ideal* scheduling. In order to solve the optimization problem (4.2), we proposed to introduce the virtual queues given by (4.3).

From queuing theory [31], it is known that strong stability of the queues means that the average arrival rate is less than the average departure rate. Using the virtual queue defined by equation (4.3), we can establish the equivalence between the throughput constraint of the optimization problem and the strong stability of the system of virtual queues. It means that the goal is to minimize the energy consumption of the D2D network while guaranteeing the stability of the queues. The proof is based on Lyapunov optimization theory. We first define the following Lyapunov function $L_{\boldsymbol{Q}}(t)$ is given by:

$$L_{\boldsymbol{Q}}(t) := \frac{1}{2}\sum_{i=1}^{\mathrm{N}} Q_i^2(t)$$

We denote by $\boldsymbol{Q}(t) = (Q_1(t), ..., Q_N(t))$ the vector of the current virtual queue backlogs and by $\boldsymbol{H}(t) = [h_1, h_2, ..., h_N]$ the vector of channel states. Applying the methodology in [110] gives that the drift-plus-penalty expression is upper bounded by the following:

$$\Delta(\boldsymbol{Q}(t)) := \Delta(\boldsymbol{L}(t)) + V\sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[P_i(t)\,|\boldsymbol{Q}(t)\right]$$

$$\leq C + \sum_{i=1}^{\mathrm{N}} Q_i(t) R_{th} - \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} Q_i(t) R_i\left(\Gamma^{dist}(t), \boldsymbol{H}(t)\right)|\boldsymbol{Q}(t)\right]$$

$$+ V\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} P_i\left(\Gamma^{dist}(t), \boldsymbol{H}(t)\right)|\boldsymbol{Q}(t)\right] \tag{7.60}$$

where $\Gamma^{dist}(t)$ is the proposed distributed scheduling policy and C is a finite constant such that $\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} \frac{R_{th}^2 + R_i^2(t)}{2}|\boldsymbol{Q}(t)\right] \leq C$.

In order to compare between the proposed distributed scheduling and the *ideal* one, we recall the respective procedure of these policies. For the *ideal* scheduling Γ_{id}^*, we suppose the global CSI knowledge of the D2D communications. Based on these D2D channel states, both the transmission power and the throughput of all the D2D links are recognized. Thus, the scheduled user corresponds to the one that optimizes problem (4.2) (i.e. the one that minimizes the consumption power under the throughput constraint).

For the proposed distributed scheduling, at each time-slot t, each D2D user i estimates its channel state and deduces its energy efficiency metric $v_i(t)$. Due to an existing mapping between the discrete values of the energy efficiency metric and the $K^{(2)}$ available resource elements for CSI feedback; each user i sends a simple CSI indicator on the RE that maps with its discrete value $\tilde{v}_i(t)$. Supposing a non collision scenario, the scheduled user corresponds to the one that transmits its CSI indicator on the RE with the lowest index (i.e. the user that has the lowest value of

the energy efficiency metric $v_i(t)$). Thus, compared to any other scheduling policy $\Gamma(t)$, the proposed distributed scheduling verifies the following at each time slot t:

$$\sum_{i=1}^{\mathrm{N}} \left[VP_i\left(\Gamma^{dist}(t), \boldsymbol{H}(t)\right) - Q_i(t) R_i\left(\Gamma^{dist}(t), \boldsymbol{H}(t)\right)\right] |\boldsymbol{Q}(t)$$

$$\leq \sum_{i=1}^{\mathrm{N}} \left[VP_i\left(\Gamma(t), \boldsymbol{H}(t)\right) - Q_i(t) R_i\left(\Gamma(t), \boldsymbol{H}(t)\right)\right] |\boldsymbol{Q}(t)$$

The equation above is verified for any scheduling policy $\Gamma(t)$ including the *ideal* policy $\Gamma^*_{id}(t)$ where the network has the global CSI knowledge of D2D communications. Thus:

$$\mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} \left[VP_i\left(\Gamma^{dist}(t), \boldsymbol{H}(t)\right) - Q_i(t) R_i\left(\Gamma^{dist}(t), \boldsymbol{H}(t)\right)\right] |\boldsymbol{Q}(t)\right]$$

$$\leq \mathbb{E}\left[\sum_{i=1}^{\mathrm{N}} \left[VP_i\left(\Gamma^*_{id}(t), \boldsymbol{H}(t)\right) - Q_i(t) R_i\left(\Gamma^*_{id}(t), \boldsymbol{H}(t)\right)\right] |\boldsymbol{Q}(t)\right] \tag{7.61}$$

The ideal policy $\Gamma^*_{id}(t)$ verifies the OP (4.2) which implies that the corresponding throughput constraint is satisfied, thus this policy guarantees the stability of the virtual queues. Supposing that the arrival rate vector of these virtual queues is interior to the stability region of the system of queue and $\epsilon > 0$, we have:

$$\mathbb{E}\left[R_i\left(\Gamma^*_{id}(t)\right) |\boldsymbol{Q}(t)\right] = \mathbb{E}\left[R_i\left(\Gamma^*_{id}(t)\right)\right] \geq R_{th} + \epsilon$$

Same methodology as in [110] gives the following upper bound of the total average backlogs of the virtual queues:

$$\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[Q_i(t)\right] \leq \frac{C+B}{\epsilon} \tag{7.62}$$

with B a finite constant such that $\limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[VP_i\left(\Gamma^*_{id}(t)\right)\right] \leq B$. Thus, all the virtual queues in the system are strongly stable. Hence, when the arrival rate at the virtual queues is less than its average departure rate then the distributed scheduling policy satisfies the throughput constraint in (4.2). Pursuing with the Lyapunov optimization of queuing networks leads to the proposition 4 as follows:

$$\frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[P_i\left(\Gamma^{dist}(t)\right)\right] \leq \frac{C}{V} + \frac{\mathbb{E}\left[L_{\boldsymbol{Q}}(0)\right]}{VT} + \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^{\mathrm{N}} \mathbb{E}\left[P_i\left(\Gamma^*_{id}(t)\right)\right]$$

Therefore, we propose an algorithm based on a distributed approach that benefits from the D2D local knowledge of their CSI values. When collision is bypassed, the proposed distributed algorithm achieves the performance of the *ideal* solution of (4.2) with a distance of $O\left(\frac{1}{V}\right)$ and a time average queue backlog of $O(V)$.

7.4.3 Proof of proposition 5

Recall that $K^{(2)}$ REs are available for the transmission of the CSI indicators. A collision occurs at the level of the k^{th} RE when at least two users have the same quantized energy efficiency metric (i.e. $\tilde{v}_i = \tilde{v}_j = a_k$ with $i \neq j$ hence a collision occurs at the k^{th} element). Therefore, the probability of overall collision P_c corresponds to the probability that none of the D2D users in the network is scheduled (i.e. the CSI indicator's transmission of each D2D user collides with at least another one). Let's compute this probability of collision P_c assuming a single bit rate model (M = 1) with R and S the corresponding bit-rate and SNR threshold.

For each element $a_j \in \mathbb{S}_v$ (with $1 \leq j \leq K^{(2)}$), we define the following two events whose probabilities of occurrence are computed in the sequel:

- $A_{i,j}$: i^{th} D2D link has $\tilde{v}_i = a_j$
- $B_{i,j}$: i^{th} D2D link has $\tilde{v}_i \geq a_{\min\{j+1,K^{(2)}\}}$

A Rayleigh fading channel h_i with zero mean and unit variance is considered, hence the squared magnitude $|h_i|^2$ has an exponential distribution of parameter one. Therefore, we can deduce the probability of the two above events. For the simplification of coming expressions, we use the following notation:

$$c_{i,j} = -\frac{V\mathsf{S}N_o}{(a_j + Q_i\mathsf{R})\, L_i}$$

1. **The probability of $A_{i,j}$ (for $1 \leq j \leq K^{(2)}$)**

$$\mathbb{P}(A_{i,j}) = \mathbb{P}(\tilde{v}_i = a_j) = \mathbb{P}(v_i \epsilon\,]a_j, a_{j+1}]) = \mathbb{P}\left(a_j < \frac{V\mathsf{S}N_o}{|h_i|^2 L_i} - Q_i\mathsf{R} \leq a_{j+1}\right)$$

$$= 2\left[\exp\left(-\frac{V\mathsf{S}N_o}{(a_j + Q_i\mathsf{R})\, L_i}\right) - \exp\left(-\frac{V\mathsf{S}N_o}{(a_{j+1} + Q_i\mathsf{R})\, L_i}\right)\right] = 2\left[\exp(c_{i,j}) - \exp(c_{i,j+1})\right]$$

where $a_{K^{(2)}+1}$ is equal to $+\infty$.

2. **The probability of $B_{i,j}$ (for $1 \leq j \leq K^{(2)} - 1$)**

$$\mathbb{P}(B_{i,j}) = \mathbb{P}(\tilde{v}_i \geq a_{j+1}) = 1 - 2\exp\left(-\frac{V\mathsf{S}N_o}{(a_{j+1} + Q_i\mathsf{R})\, L_i}\right) = 1 - 2\exp(c_{i,j+1})$$

$\bar{p}_c\{i,j\}$ denotes the probability that only the user i has its quantized value $\tilde{v}_i$ equals to $a_j \in \mathbb{S}_v$ for $1 \leq j \leq K^{(2)}$ and is given by:

$$\bar{p}_c\{i,j\} = \mathbb{P}(A_{i,j}) \prod_{k=1 \neq i}^{\mathrm{N}} \mathbb{P}\left(\bar{A}_{k,j}\right)$$

$$= 2\left[\exp\left(c_{i,j-1}\right) - \exp\left(c_{i,j}\right)\right] \times \prod_{k=1\neq i}^{N} \left[1 - 2\exp\left(c_{k,j-1}\right) + 2\exp\left(c_{k,j}\right)\right]$$

The probability of collision P_c is deduced as follows:

$$P_c = \mathbb{P}\{\text{no scheduled D2D user}\} = 1 - \sum_{i=1}^{\mathrm{N}} \mathbb{P}\{\text{D2D user } i \text{ scheduled}\}$$

$$= 1 - \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{K^{(2)}} \mathbb{P}\{\text{only the } i^{th} \text{ D2D user has } \tilde{v}_i = a_j\} \times \mathbb{P}\{j = \text{lowest RE index without collision}\}$$

$$P_c = 1 - \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{K^{(2)}} \bar{p}_c\{i,j\} \prod_{k=1}^{j-1} \left(1 - \sum_{l=1\neq i}^{\mathrm{N}} \bar{p}_c\{l,k\}\right)$$

7.4.4 Proof of theorem 4.5.1

The proof is split into 4 steps. In **step 1**, we start by expressing the lower bounds of $\mathbb{P}\left(A_{i,j}\right)$ and $\mathbb{P}\left(\bar{A}_{i,j}\right)$. These bounds are used, in **step 2**, for computing the upper and the lower bounds of the probability $\bar{p}_c\{i,j\}$ (i.e. probability that only the D2D user i transmits its CSI indicator on RE j). Therefore, in **step 3**, we find the upper bound of the collision probability P_c. In **step 4**, we conclude the value of the Lyapunov constant V that reduces the collision probability P_c to ϵ.

Step 1

Verify that for all $\{i,j\}$:

$$\mathbb{P}\left(A_{i,j}\right) \geq \mathbb{P}\left(B_{i,K^{(2)}-1}\right) \text{ and } \mathbb{P}\left(\bar{A}_{i,j}\right) \geq \mathbb{P}\left(B_{i,K^{(2)}-1}\right)$$

For $1 \leq j \leq K^{(2)} - 1$, we have the two following expressions:

$$\mathbb{P}\left(A_{i,j}\right) = 2\left[\exp\left(\frac{-VSN_0L_i^{-1}}{a_j + Q_iR}\right) - \exp\left(\frac{-VSN_0L_i^{-1}}{a_{j+1} + Q_iR}\right)\right]$$

$$\mathbb{P}\left(A_{i,K^{(2)}}\right) = \mathbb{P}\left(B_{i,K^{(2)}-1}\right) = \left[1 - 2\exp\left(\frac{-VSN_0L_i^{-1}}{a_{k^{(2)}} + Q_iR}\right)\right]$$

Then, $\mathbb{P}\left(A_{i,j}\right) \geq \mathbb{P}\left(B_{i,K^{(2)}-1}\right)$ and $\mathbb{P}\left(B_{i,j}\right) \geq \mathbb{P}\left(B_{i,K^{(2)}-1}\right)$ $\forall 0 \leq i \leq N$, $\forall 1 \leq j \leq K^{(2)}$.

Based on the definition of $\mathbb{P}\left(A_{i,j}\right)$ and $\mathbb{P}\left(B_{i,j}\right)$ one can see that:

$$\mathbb{P}\left(A_{i,j}\right) = \mathbb{P}\left(B_{i,j-1}\right) - \mathbb{P}\left(B_{i,j}\right) \Rightarrow \mathbb{P}\left(\bar{A}_{i,j}\right) = 1 - \mathbb{P}\left(B_{i,j-1}\right) + \mathbb{P}\left(B_{i,j}\right)$$

$$\Rightarrow \mathbb{P}\left(\bar{A}_{i,j}\right) \geq \mathbb{P}\left(B_{i,j}\right) \geq \mathbb{P}\left(B_{i,K^{(2)}-1}\right)$$

Step 2

Upper and lower bounds of $\bar{p}_c\{i,j\}$: Based on the lower bound of $\mathbb{P}(A_{i,j})$ and $\mathbb{P}\left(\bar{A}_{i,j}\right)$ we deduce that:

- Lower bound: $\bar{p}_c\{i,j\} \geq \prod_{i=i}^{\mathrm{N}} \mathbb{P}\left(B_{i,K^{(2)}}\right)$
- Upper bound: $\bar{p}_c\{i,j\} \leq \mathbb{P}\left(\bar{B}_{i+1 \bmod \mathrm{N},K^{(2)}}\right)$

Step 3

Upper bound of P_c: Based on the upper and lower bound of $\bar{p}_c\{i,j\}$, we deduce that:

$$P_c = 1 - \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{K^{(2)}} \bar{p}_c\{i,j\} \prod_{1}^{j-1} \left(1 - \sum_{l=1\neq i}^{\mathrm{N}} \bar{p}_c\{l,k\}\right)$$

$$P_c \leq 1 - \left[\prod_{m=1}^{\mathrm{N}} \mathbb{P}\left(B_{m,K^{(2)}-1}\right)\right] \times \sum_{i=1}^{\mathrm{N}} \sum_{j=1}^{K^{(2)}} \left[1 - \sum_{l=1\neq i}^{\mathrm{N}} \mathbb{P}\left(\bar{B}_{l+1 \bmod \mathrm{N},K^{(2)}-1}\right)\right]^{j}$$

$$P_c \leq 1 - \mathrm{N}K^{(2)} \prod_{m=1}^{\mathrm{N}} \left[1 - 2\exp\left(\frac{-V\mathsf{S}N_0 L_m^{-1}}{a_{K^{(2)}} + Q_m \mathsf{R}}\right)\right] \times \left[1 - 2\sum_{l=1}^{\mathrm{N}} \exp\left(\frac{-V\mathsf{S}N_0 L_l^{-1}}{a_{K^{(2)}} + Q_l \mathsf{R}}\right)\right]^{K^{(2)}}$$

In fact, for all $1 \leq i \leq N$, we can verify that:

$$\exp\left(\frac{-V\mathsf{S}N_0 L_l^{-1}}{a_{K^{(2)}} + Q_l \mathsf{R}}\right) \leq \exp\left(\frac{-V\mathsf{S}N_0 L_{min}^{-1}}{VP_{max} + R_{th}\mathsf{R}t_{simu}}\right)$$

where L_{min} is the path-loss over a D2D link with a peer distance equals to the minimum D2D peer distance (i.e. d_{min}). Hence,

$$P_c \leq 1 - \mathrm{N}K^{(2)} \left[1 - 2\mathrm{N}\exp\left(\frac{-V\mathsf{S}N_0 L_{min}^{-1}}{VP_{max} + R_{th}\mathsf{R}t_{simu}}\right)\right]^{\mathrm{N}+K^{(2)}}$$

Step 4

Deduce $V(\epsilon)$: In order to bound the collision probability P_c by ϵ, the Lyapunov constant $V(\epsilon)$ should verify theorem 4.5.1.

7.5 Mobile D2D Relay Selection Proofs

7.5.1 Proof of theorem 5.4.4

Proof. The following definition and lemma are used for proving theorem 5.4.4.

Definition 7.5.1. *The distance function of relay i with transition probability $\boldsymbol{P}_i$ and stationary distribution π_i is defined as follows:*

$$d_i(t) = \max_{s\in\mathcal{S}} ||\boldsymbol{P}_i^t(s,:) - \boldsymbol{\pi}_i|| \tag{7.63}$$

Lemma 7.5.2. *From [111], the distance function of relay i verifies the following properties:*

- $d_i(t) \leq \frac{\lambda_i^{*t}}{\pi_{i,min}}$ *with λ_i^* the highest eigenvalue of the matrix $\boldsymbol{P}_i$ and $\pi_{i,min}$ the lowest component of the stationary distribution corresponding to $\boldsymbol{P}_i$.*
- $d_i(t+a) \leq d_i(t) \, \forall a \in \mathbb{N}^+$

The density $\epsilon^{\mathcal{B}}$ of a belief set $\mathcal{B}(\boldsymbol{s},h)$ is given by:

$$\epsilon^{\mathcal{B}}(\mathcal{B}(\boldsymbol{s},h)) = \max_{\tilde{\boldsymbol{b}}\in\Delta} \min_{\boldsymbol{b}\in\mathcal{B}(\boldsymbol{s},h)} ||\boldsymbol{b} - \boldsymbol{b}'||_1$$

$$\leq \max_{\tilde{\boldsymbol{b}}\in\Delta} \min_{\boldsymbol{b}\in\mathcal{B}(\boldsymbol{s},h)} \sum_{i\in\mathcal{K}} \left[\sum_{s\in\mathcal{S}} |b^i(s) - \tilde{b}^i(s)| \right]$$

For any reachable belief point $\tilde{\boldsymbol{b}}^i$ of relay i, there exists an integer n and a state $S_i \in \mathcal{S}$ such that $\tilde{\boldsymbol{b}}^i = \boldsymbol{P}_i^n(S_i,:)$. The construction of the belief set $\mathcal{B}(\boldsymbol{s},h)$ induces that $\boldsymbol{b}^i$ in the equation above can be written as $\boldsymbol{b}^i = \boldsymbol{P}_i^m(s_i(0),:)$ with the integer $m \leq h$ and $s_i(0)$ the initial state of relay i.

We will consider both cases:

- if $n \leq h$, then the corresponding reachable point $\tilde{\boldsymbol{b}}^i$ has been taken into account in the belief set $\mathcal{B}(\boldsymbol{s},h)$, thus:
$$\sum_{s\in\mathcal{S}} |b^i(s) - \tilde{b}^i(s)| = 0 \tag{7.64}$$
- if $n > h$, the expression $\sum_{s\in\mathcal{S}} |b^i(s) - \tilde{b}^i(s)|$ can be replaced by:
$$\sum_{s\in\mathcal{S}} |b^i(s) - \tilde{b}^i(s)| = \sum_{s\in\mathcal{S}} |P_i^m(s_i(0),s) - P_i^n(S_i,s)|$$
$$\leq \sum_{s\in\mathcal{S}} |P_i^m(s_i(0),s) - \pi_i(s)| + \sum_{s\in\mathcal{S}} |P_i^n(s_i(0),s) - \pi_i(s)| \leq d_i(m) + d_i(n)$$
Based on lemma 7.5.2 and since $n > h$, then:
$$\sum_{s\in\mathcal{S}} |b^i(s) - \tilde{b}^i(s)| \leq 2d_i(h) \leq 2\frac{\lambda_i^{*h}}{\pi_{i,min}} \tag{7.65}$$

From equations (7.64) and (7.65) and summing over all the relays $i \in \mathcal{K}$ we prove the theorem 5.4.4. □

7.5.2 Proof of theorem 5.4.5

Proof. The studied CPOMDP of horizon T and an initial state $\boldsymbol{s}_0$ achieves an error bound of ϵ when η_T^r and η_T^c (given by equation (5.20)) are lower than ϵ. Limiting the belief set to $\mathcal{B}(\boldsymbol{s}_0, h)$ generates the following errors on the value functions:

$$\eta_T^r \leq \frac{(R_{max} - R_{min})}{(1-\gamma)^2} \sum_{i \in \mathcal{K}} 2 \frac{\lambda_i^{*h}}{\pi_{i,min}} \leq 2K \frac{(R_{max} - R_{min})\, \lambda^{*h}}{\pi_{min}\, (1-\gamma)^2} \tag{7.66}$$

and

$$\eta_T^c \leq 2K \frac{(C_{max} - C_{min})\, \lambda^{*h}}{\pi_{min}\, (1-\gamma)^2} \tag{7.67}$$

with $\lambda^* = \max\limits_{i \in \mathcal{K}} \lambda_i^*$; $\pi_{min}^* = \min\limits_{i \in \mathcal{K}} \pi_{i,min}^*$.

Therefore, limiting η_T^r and η_T^c to ϵ corresponds to choosing the parameter h of $\mathcal{B}(\boldsymbol{s}_0, h)$ in such a way that the expressions in equations (7.66) and (7.67) are bounded by ϵ. Hence, theorem 5.4.5 is deduced.

Note that for a discount factor $\gamma = 1$, the sum over the horizon T gives the following upper bounds of the errors η_T^r and η_T^c:

$$\eta_T^r \leq 2KT \frac{(R_{max} - R_{min})\, \lambda^{*h}}{\pi_{min}} \text{ and } \eta_T^c \leq 2KT \frac{(C_{max} - C_{min})\, \lambda^{*h}}{\pi_{min}}$$

Thus, for $\gamma = 1$, the expressions of $f_r(\epsilon)$ and $f_c(\epsilon)$ in equation 5.25 of theorem 5.4.5 are given by:

$$f_r(\epsilon) = \log\left(\frac{\epsilon \pi_{min}}{2KT\,(R_{max} - R_{min})}\right) ; f_c(\epsilon) = \log\left(\frac{\epsilon \pi_{min}}{2KT\,(C_{max} - C_{min})}\right)$$

□

7.5.3 Proof of theorem 5.4.6

Proof. Please note that the proof of theorem 5.4.6 is done considering the reward Q-function as an example. By analogy, we can deduce the same property for the cost Q-function. Based on the definition of discrete derivative given in section 2.4 (equation 2.29, we deduce the discrete derivative of the relay $e \in \mathcal{K}$ knowing that an action set $\hat{a}_M$ (corresponding to action $\boldsymbol{a}_M$) is taken as follows:

$$\Delta_{Q_t^\pi}(e|\hat{a}_M) := Q_t^\pi\left(\boldsymbol{b}^t, \hat{a}_M \cup e\right) - Q_t^\pi\left(\boldsymbol{b}^t, \hat{a}_M\right) \tag{7.68}$$

$$= \rho\left(\boldsymbol{b}^t, \hat{a}_M \cup e\right) - \rho\left(\boldsymbol{b}^t, \hat{a}_M\right) + \sum_{k=t+1}^{T} \left[G_k^\pi\left(\boldsymbol{b}^t, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}^t, \hat{a}_M\right)\right]$$

As shown in section 2.4, $Q_t^\pi(\boldsymbol{b}, \boldsymbol{a})$ is non-negative, monotone and submodular in $\boldsymbol{a}$ if the discrete derivative of the Q-function $\Delta_{Q_t^\pi}$ verifies the following:

$$\Delta_{Q_t^\pi}(e|\hat{a}_M) \geq 0 \tag{7.69}$$

and

$$\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right) \geq \Delta_{Q_t^\pi}\left(e|\hat{a}_M \cup \hat{a}_N\right) \tag{7.70}$$

In order to verify the two properties above, we compute the discrete derivative $\Delta_{Q_t^\pi}$ for a given relay e knowing that the action set is $\hat{a}_M$:

$$\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right) = \rho\left(\boldsymbol{b}^t, \hat{a}_M \cup e\right) - \rho\left(\boldsymbol{b}, \hat{a}_M\right) + \sum_{k=t+1}^{T}\left[G_k^\pi\left(\boldsymbol{b}^t, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}^t, \hat{a}_M\right)\right]$$

$$= \sum_{s\in\mathcal{S}} b_t^e\left(s\right) r_e\left(s\right) + \sum_{k=t+1}^{T}\left[G_k^\pi\left(\boldsymbol{b}^t, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}^t, \hat{a}_M\right)\right]$$

$$= \sum_{s\in\mathcal{S}} b_t^e\left(s\right) r_e\left(s\right) + \sum_{k=t+1}^{T}\gamma^k \sum_{\boldsymbol{z}^{t:k}}\left[P\left(\boldsymbol{z}^{t:k}|\boldsymbol{b}^t, \hat{a}_M \cup e, \pi\right)\rho\left(\boldsymbol{b}^{\boldsymbol{a}^\pi}_{\boldsymbol{z}^{t:k}}, \boldsymbol{a}^\pi\right) - P\left(\boldsymbol{z}^{t:k}|\boldsymbol{b}^t, \hat{a}_M, \pi\right)\rho\left(\boldsymbol{b}^{\boldsymbol{a}^\pi}_{\boldsymbol{z}^{t:k}}, \boldsymbol{a}^\pi\right)\right]$$

Since the total reward model, given by equation 5.1, is equal to the sum of the reward of each selected relay, then the difference $G_k^\pi\left(\boldsymbol{b}, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}, \hat{a}_M\right)$ will be limited to the reward of relay e as follows:

$$\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right) = \sum_{s\in\mathcal{S}} b_t^e\left(s\right) r_e\left(s\right) + \sum_{k=t+1}^{T}\gamma^k \sum_{z_e^{t:k}}\left[P\left(z_e^{t:k}|\boldsymbol{b}_t^e, \hat{a}_M \cup e, \pi\right)\rho\left(\boldsymbol{b}^e_{z_e^{t:k}}, \boldsymbol{a}^\pi\right) - P\left(z_e^{t:k}|\boldsymbol{b}_t^e, \hat{a}_M, \pi\right)\rho\left(\boldsymbol{b}^e_{z_e^{t:k}}, \boldsymbol{a}^\pi\right)\right]$$

where $z_e^{t:k}$ is the observation of the state of relay e between t and k epochs, $\boldsymbol{b}_t^e$ the belief vector of relay e at epoch t, $\boldsymbol{b}^e_{z_e^{t:k}}$ the belief vector of relay e after an observation $z_e^{t:k}$ between t and k epochs. In order to prove the properties 7.69 and 7.70 and deduce by that the submodularity of the Q-function, we compute $\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right)$ for the following two cases:

- Relay e has not been chosen by the policy π between the epochs $t+1$ and T: in this case, the belief vector $\boldsymbol{b}_n^e$ of relay e (for $n = t+1, ..., T$) is the same in both expressions $G_k^\pi\left(\boldsymbol{b}, \hat{a}_M\right)$ and $G_k^\pi\left(\boldsymbol{b}, \hat{a}_M \cup e\right)$. Hence,

$$G_k^\pi\left(\boldsymbol{b}, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}, \hat{a}_M\right) = 0 \text{ for all } k = t+1, ..., T$$

$$\Rightarrow \Delta_{Q_t^\pi}\left(e|\hat{a}_M\right) = \sum_{s\in\mathcal{S}} b_t^e\left(s\right) r_e\left(s\right)$$

  Thus, in this case the discrete derivative of the Q-function verifies the properties 7.69 and 7.70.

- Relay e has been chosen by the policy π at epoch t^*: in this case, the MU observes the state of relay e: $z_e^{t^*}$. For this case, $\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right)$ can be written as follows:

$$\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right) = \sum_{s\in\mathcal{S}} b_t^e\left(s\right) r_e\left(s\right) + \sum_{k=t+1}^{t^*}\gamma^k\left[G_k^\pi\left(\boldsymbol{b}, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}, \hat{a}_M\right)\right]$$

$$+\gamma^{t^*+1}\left[G^\pi_{t^*+1}\left(\boldsymbol{b}, \hat{a}_M \cup e\right) - G_{t^*+1}k^\pi\left(\boldsymbol{b}, \hat{a}_M\right)\right] + \sum_{k=t^*+2}^{T}\gamma^k\left[G_k^\pi\left(\boldsymbol{b}, \hat{a}_M \cup e\right) - G_k^\pi\left(\boldsymbol{b}, \hat{a}_M\right)\right]$$

Epoch	t	$t^* = t+1$	$t^* = t+2$	...	t^*	...
$\boldsymbol{b}^e$	$\boldsymbol{b}_t^e$	$\boldsymbol{P}_e.\boldsymbol{b}_t^e$	$\boldsymbol{P}_e^2.\boldsymbol{b}_t^e$	...	$\boldsymbol{P}_e^{t^*-t}.\boldsymbol{b}_t^e$	...
$Pr\left(z_e\|\boldsymbol{a},\boldsymbol{b}\right)$	$\left(\boldsymbol{P}_e.\boldsymbol{b}_t^e\right)_{z_e^t}$	$\left(\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e\right)_{z_e^{t+1}}$	$\left(\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e^2\right)_{z_e^{t+2}}$	...	$\left(\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e^{t^*-t}\right)_{z_e^{t^*}}$	...

Table 7.3: The values of the belief state $\boldsymbol{b}^e$ and the probability $Pr\left(z|\boldsymbol{a},\boldsymbol{b}\right)$ of the relay e in the case where the action at epoch t contains the relay e

The second term $\sum_{k=t+1}^{t^*}\gamma^k\left[G_k^\pi\left(\boldsymbol{b},\hat{a}_M\cup e\right)-G_k^\pi\left(\boldsymbol{b},\hat{a}_M\right)\right]=0$ since relay e has not been chosen between $t+1$ and t^* epochs. In addition, the fourth term $\sum_{k=t^*+2}^{T}\gamma^k\left[G_k^\pi\left(\boldsymbol{b},\hat{a}_M\cup e\right)-G_k^\pi\left(\boldsymbol{b},\hat{a}_M\right)\right]=0$ since relay e has been chosen at epoch t^*. Therefore, the policy π will have the same value for both $G_k^\pi\left(\boldsymbol{b},\hat{a}_M\cup e\right)$ and $G_k^\pi\left(\boldsymbol{b},\hat{a}_M\right)$. Thus,

$$\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right)=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\left[G_{t^*+1}^\pi\left(\boldsymbol{b},\hat{a}_M\cup e\right)-G_{t^*+1}k^\pi\left(\boldsymbol{b},\hat{a}_M\right)\right]$$

$$=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\sum_{z_e^{t:k}}\left[P\left(z_e^{t:k}|\boldsymbol{b}_t^e,\hat{a}_M\cup e,\pi\right)\rho\left(\boldsymbol{b}_{z_e^{t:k}}^e,\boldsymbol{a}^\pi\right)-P\left(z_e^{t:k}|\boldsymbol{b}_t^e,\hat{a}_M,\pi\right)\rho\left(\boldsymbol{b}_{z_e^{t:k}}^e,\boldsymbol{a}^\pi\right)\right]$$

$$=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\sum_{z_e^{t^*}}\left[\sum_{z_e^t}P\left(z_e^{t^*}|z_e^t\right)P\left(z_e^t|\boldsymbol{b}_t^e,\hat{a}_M\cup e,\pi\right)-P\left(z_e^{t^*}|\boldsymbol{b}_t^e,\hat{a}_M,\pi\right)\right]\sum_{s'\in\mathcal{S}}\boldsymbol{P}_e\left(z_e^{t^*},s'\right)r_e\left(s'\right)$$

$$=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\sum_{z_e^{t^*}}\left[\sum_{z_e^t}P\left(z_e^{t^*}|z_e^t\right)\left(\boldsymbol{P}_e.\boldsymbol{b}_t^e\right)\left(z_e^t\right)-\left(\boldsymbol{P}_e^{t^*-t+1}.\boldsymbol{b}_t^e\right)\left(z_e^{t^*}\right)\right]\sum_{s'\in\mathcal{S}}\boldsymbol{P}_e\left(z_e^{t^*},s'\right)r_e\left(s'\right)$$

$$=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\sum_{z_e^{t^*}}\left[\left(\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e^{t^*-t}\right)\left(z_e^{t^*}\right).\left(\boldsymbol{P}_e.\boldsymbol{b}_t^e\right)\left(z_e^t\right)-\left(\boldsymbol{P}_e^{t^*-t+1}.\boldsymbol{b}_t^e\right)\left(z_e^{t^*}\right)\right]\sum_{s'\in\mathcal{S}}\boldsymbol{P}_e\left(z_e^{t^*},s'\right)r_e\left(s'\right)$$

$$=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\sum_{z_e^{t^*}}\left[\left(\boldsymbol{P}_e^{t^*-t+2}.\boldsymbol{b}_t^e\right)\left(z_e^{t^*}\right)-\left(\boldsymbol{P}_e^{t^*-t+1}.\boldsymbol{b}_t^e\right)\left(z_e^{t^*}\right)\right]\sum_{s'\in\mathcal{S}}\boldsymbol{P}_e\left(z_e^{t^*},s'\right)r_e\left(s'\right)$$

$$=\sum_{s\in\mathcal{S}}b_t^e\left(s\right)r_e\left(s\right)+\gamma^{t^*+1}\sum_{s'\in\mathcal{S}}\left[\left(\boldsymbol{P}_e^{t^*-t+3}.\boldsymbol{b}_t^e\right)\left(s'\right)-\left(\boldsymbol{P}_e^{t^*-t+2}.\boldsymbol{b}_t^e\right)\left(s'\right)\right]r_e\left(s'\right)$$

$$=\sum_{s\in\mathcal{S}}r_e\left(s\right)\left[b_t^e\left(s\right)+\gamma^{t^*+1}\left(\boldsymbol{P}_e^{t^*-t+3}.\boldsymbol{b}_t^e\right)\left(s\right)-\gamma^{t^*+1}\left(\boldsymbol{P}_e^{t^*-t+2}.\boldsymbol{b}_t^e\right)\left(s\right)\right]$$

Lemma 7.5.3. *For a relay e, the two following statements are verified by induction:*

- *if* $\left(\boldsymbol{P}_e.\boldsymbol{b}_t^e\right)\left(s\right)\geq\boldsymbol{b}_t^e\left(s\right)$ *then* $\left(\boldsymbol{P}_e^n.\boldsymbol{b}_t^e\right)\left(s\right)\geq\left(\boldsymbol{P}_e^{n-1}.\boldsymbol{b}_t^e\right)\left(s\right)\forall n\in\mathbb{Z}\geq 2$
- *if* $\left(\boldsymbol{P}_e.\boldsymbol{b}_t^e\right)\left(s\right)\leq\boldsymbol{b}_t^e\left(s\right)$ *then* $\left(\boldsymbol{P}_e^n.\boldsymbol{b}_t^e\right)\left(s\right)\leq\left(\boldsymbol{P}_e^{n-1}.\boldsymbol{b}_t^e\right)\left(s\right)\forall n\in\mathbb{Z}\geq 2$

Based on the lemma 7.5.3, we deduce that, in this case, the discrete derivative $\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right)$ verifies the properties 7.69 and 7.70.

Since the discrete derivative $\Delta_{Q_t^\pi}\left(e|\hat{a}_M\right)$ verifies the properties 7.69 and 7.70 for the two possible cases studied above; theorem 5.4.6 is verified.

□

Epoch	t	$t^* = t+1$	$t^* = t+2$	...	t^*	...
$\boldsymbol{b}^e$	$\boldsymbol{b}_t^e$	$\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e$	$\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e$	...	$\boldsymbol{P}_e\left(z_e^t,:\right).\boldsymbol{P}_e^{t^*-t-1}$	...
$Pr\left(z_e\|\boldsymbol{a},\boldsymbol{b}\right)$	$\left(\boldsymbol{P}_e.\boldsymbol{b}_t^e\right)_{z_e^t}$	$\left(\boldsymbol{P}_e^2.\boldsymbol{b}_t^e\right)_{z_e^{t+1}}$	$\left(\boldsymbol{P}_e^3.\boldsymbol{b}_t^e\right)_{z_e^{t+2}}$	...	$\left(\boldsymbol{P}_e^{t^*-t+1}.\boldsymbol{b}_t^e\right)_{z_e^{t^*}}$	...

Table 7.4: The values of the belief state $\boldsymbol{b}^e$ and the probability $Pr\left(z|\boldsymbol{a},\boldsymbol{b}\right)$ of the relay e in the case where the action at epoch t does not contain the relay e

7.5.4 Proof of theorem 5.4.7

Proof. The proof is done by induction. We start by verifying equation (5.28) for $t = 1$:

$$V_1^{r,G}\left(\boldsymbol{b}\right) = greedy - \operatorname{argmax} \rho_r\left(\boldsymbol{b},\boldsymbol{a}\right) \text{ s.t. } \rho_c\left(\boldsymbol{b},\boldsymbol{a}\right) \leq C_{th}$$

and

$$V_1^{r,B}\left(\boldsymbol{b}\right) = \operatorname{argmax} \rho_r\left(\boldsymbol{b},\boldsymbol{a}\right) \text{ s.t. } \rho_c\left(\boldsymbol{b},\boldsymbol{a}\right) \leq C_{th}$$

Based on the results given in 2.4.2, the greedy maximization algorithm achieves a $1-\frac{1}{e}$ approximation. Thus,$V_1^{r,G} \geq \left(1-\frac{1}{e}\right) V_1^{r,B}$.

Then, we assume that equation (5.28) is verified for $t-1$ and prove that it remains verified for t.

$$V_{t-1}^{r,G}\left(\boldsymbol{b}\right) \geq \left(1-\frac{1}{e}\right)^{2t-2} V_{t-1}^{r,B}\left(\boldsymbol{b}\right)$$

$$\rho_r\left(\boldsymbol{b},\boldsymbol{a}\right)+\gamma\sum_{\boldsymbol{z}\in\mathcal{Z}} P\left(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}\right) V_{t-1}^{r,G}\left(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}\right) \geq \left(1-\frac{1}{e}\right)^{2t-2}\left[\rho_r\left(\boldsymbol{b},\boldsymbol{a}\right)+\gamma\sum_{\boldsymbol{z}\in\mathcal{Z}} P\left(\boldsymbol{z}|\boldsymbol{a},\boldsymbol{b}\right) V_{t-1}^{r,B}\left(\boldsymbol{b}_{\boldsymbol{z}}^{\boldsymbol{a}}\right)\right]$$

Based on the definition of Q_t^r function given by equation 5.15, we deduce:

$$Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}\right) \geq \left(1-\frac{1}{e}\right)^{2t-2} Q_t^{r,B}\left(\boldsymbol{b},\boldsymbol{a}\right) \forall \boldsymbol{a} \in \mathcal{A} \tag{7.71}$$

We use the following notations (all of these maximization are under cost constraint):

- $\boldsymbol{a}_{Q^G}^G = greedy - \operatorname{argmax} Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}\right)$
- $\boldsymbol{a}_{Q^G}^* = \operatorname{argmax} Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}\right)$
- $\boldsymbol{a}_{Q^B}^G = greedy - \operatorname{argmax} Q_t^{r,B}\left(\boldsymbol{b},\boldsymbol{a}\right)$
- $\boldsymbol{a}_{Q^B}^* = \operatorname{argmax} Q_t^{r,B}\left(\boldsymbol{b},\boldsymbol{a}\right)$

Thus,

$$V_t^{r,G}\left(\boldsymbol{b}\right) = greedy - \operatorname{argmax} Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}\right) = Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}_{Q^G}^G\right)$$

Given that $Q_t^{r,\pi}$-function is submodular and constrained to a modular $Q_t^{c,\pi}$, then results shown in 2.4.2 gives:

$$V_t^{r,G}\left(\boldsymbol{b}\right) = Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}_{Q^G}^G\right) \geq \left(1-\frac{1}{e}\right) Q_t^{r,G}\left(\boldsymbol{b},\boldsymbol{a}_{Q^G}^*\right)$$

From the definition of $\boldsymbol{a}^*_{Q^G}$ we know that $Q_t^{r,G}\left(\boldsymbol{b}, \boldsymbol{a}^*_{Q^G}\right) \geq Q_t^{r,G}\left(\boldsymbol{b}, \boldsymbol{a}\right)$ for all $\boldsymbol{a} \in \mathcal{A}$ including $\boldsymbol{a}^G_{Q^B}$. Therefore,

$$V_t^{r,G}\left(\boldsymbol{b}\right) \geq \left(1 - \frac{1}{e}\right) Q_t^{r,G}\left(\boldsymbol{b}, \boldsymbol{a}^G_{Q^B}\right)$$

For equation (7.71), $Q_t^{r,G}\left(\boldsymbol{b}, \boldsymbol{a}^G_{Q^B}\right) \geq \left(1 - \frac{1}{e}\right)^{2t-2} Q_t^{r,B}\left(\boldsymbol{b}, \boldsymbol{a}^G_{Q^B}\right)$. Thus:

$$V_t^{r,G}\left(\boldsymbol{b}\right) \geq \left(1 - \frac{1}{e}\right)^{2t-1} Q_t^{r,B}\left(\boldsymbol{b}, \boldsymbol{a}^G_{Q^B}\right)$$

Given that $Q_t^{r,\pi}$ function is submodular and constrained to a modular $Q_t^{c,\pi}$, then as shown in 2.4.2:

$$Q_t^{r,B}\left(\boldsymbol{b}, \boldsymbol{a}^G_{Q^B}\right) \geq \left(1 - \frac{1}{e}\right) Q_t^{r,B}\left(\boldsymbol{b}, \boldsymbol{a}^*_{Q^B}\right)$$

We deduce that:

$$V_t^{r,G}\left(\boldsymbol{b}\right) \geq \left(1 - \frac{1}{e}\right)^{2t} Q_t^{r,B}\left(\boldsymbol{b}, \boldsymbol{a}^*_{Q^B}\right) = V_t^{r,B}\left(\boldsymbol{b}\right)$$

Therefore, by induction, the theorem 5.4.7 is verified. □

7.6 Simulation Settings

As part of this thesis, we have developed a LTE system level simulator that considers DL, UL and D2D features. This simulator is mainly used for evaluating the performance of D2D communications. In particular, we expose the results of implementing the relay selection scheme proposed in chapter 5 into the developed system level simulator. In this section, we describe the different characteristics of this $C++$ system level simulator.

Layout

In the LTE system level simulator, we assume a **carrier frequency** of 2 GHz and a **bandwidth** of 10 MHz. We consider an **hexagonal** grid with 7 tri-sector macro sites. 21 hexagonal cells are formed as illustrated in figure 7.4. **Wrap-around** is deployed to avoid border effects and all the cells are supposed **synchronized**.

Two **urban macro** scenarios are considered with an Inter-Site Distance (ISD) of 500 m: (i) **option** 1: all the users are outdoor and (ii) **option** 2: with one RRH/Indoor Hotspot per cell (see section A.2.1.1.5 in [59] for more details about RRH/Indoor Hotzone). Unless specified otherwise, we implement the layout parameters specified for 3GPP case 1 in table A.2.1.1-1 of [59].

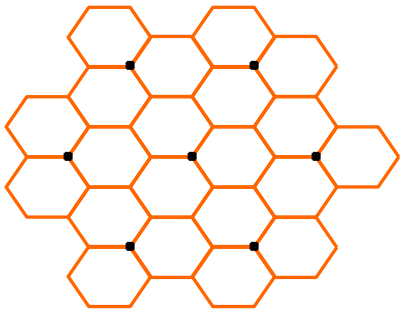
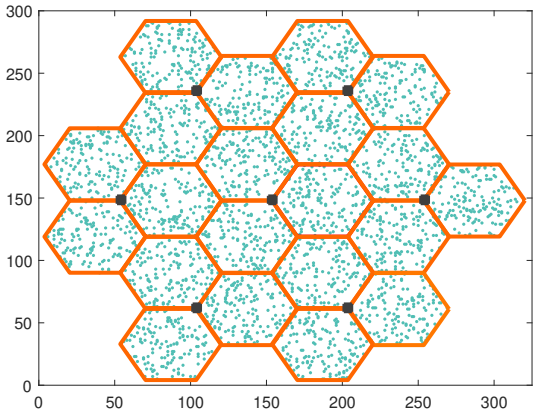


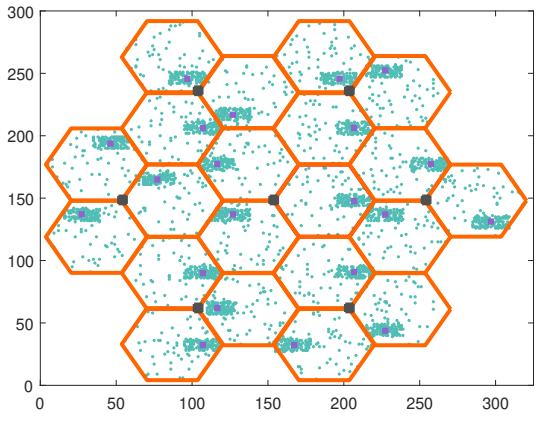
Figure 7.4: The hexagonal grid with 7 tri-sector macro sites

Users drop

The details concerning the users' drop considered in the simulator are given in table A.2.1.1-1 of [18]. We give a brief description in the sequel. For layout option 1, all the users are outdoor and they are randomly and uniformly dropped throughout the macro geographical area. For layout option 2, $^2/_3$ of the users are randomly and uniformly dropped within the indoor building of each cell and the remaining fraction $^1/_3$ of users are randomly and uniformly dropped throughout the macro geographical area. We consider 20% of the users outdoor and 80% indoor. This can be guaranteed by considering some of the users that are dropped outdoor as virtual indoor users. We assume the following constraints on the building and users drops: 3 m as the minimum distance between two users, 35m as the minimum distance between a user and the BS, 100 m as the minimum distance between the building center and the BS. In order to illustrate these two layouts, an example for 150 UEs is given in figure 7.5.



(a) Layout option 1: all users outdoor and randomly and uniformly distributed



(b) Layout option 2: an indoor rectangle building in each cell with $^2/_3$ of the users inside

Figure 7.5: The layouts of users drop that are implemented in the system-level simulator

Channel Model

For cellular communications, we consider the **antenna pattern** (horizontal and vertical) specified for 3GPP case 1 and detailed in table A.2.1.1-2 of [59]. An BS **antenna gain** of 14 dBi and an antenna tilt of 15° characterize the BS. The user is assumed to be equipped with omni-directional antennas. The BS **transmission power** is 46 dBm (almost 40W) and the users' maximum transmission power, i.e. on both UL and D2D, is 23 dBm (almost 0.2W).

The **distance dependent pathloss** for cellular communications (UL and DL) corresponds to the macro-to-UE model for 3GPP case 1 given in table A.1.1.2-3 of [59]. Hence, for a distance d Km between the BS and the UE, the cellular pathloss is given in dB as follows:

$$PL_{cell} = 128.1 + 37.6 \log 10\,(d)$$

In addition, the **LOS probability** of both UL and DL communications is the one specified in table A.1.1.2-3 of [59] for the 3GPP case 1. Thus, for a distance d Km between the BS and the UE, the LOS probability is given by:

$$\mathbb{P}_{LOS} = \min\{\frac{0.018}{d}, 1\} \times \left(1 - \exp\left(-\frac{d}{0.063}\right)\right) + \exp\left(-\frac{d}{0.063}\right)$$

When the user is indoor, then its UL and DL communications suffer from a penetration loss of mean 10dB.

For cellular links, both UL and DL, we consider large scale **shadowing** caused by the geographical irregularities of the environment due to the presence of obstacles between the BS and the UE (e.g. buildings). In this simulator, shadowing is modeled as a zero-mean Gaussian random variable, with standard deviation 8 dB. A shadowing map is constructed for each site in the network with a precision of 5 Pixels. Both spatial and time correlation exist between the shadowing maps of different sites with a correlation distance of 50 m and a time correlation of 1s. The shadowing maps are updated frequently with a periodicity of 1 s. The details of the shadowing model is described in the appendix A of [112].

For the case of D2D links, the distance dependent pathloss, the LOS probability, the penetration loss and the shadowing settings are given in table A.2.1.2 of [18] (i.e. for different scenarios outdoor-to-outdoor, outdoor-to-indoor and indoor-to-indoore). The antenna gain for D2D communications is considered 0 dBi (i.e. given in table A.2.1.1-1 of [18]).

For cellular communications, the **fast fading** has a Rayleigh distribution, i.e. square root of the sum of the squares of two normal distribution random variables $F_1(t)$ and $F_2(t)$. The fast fading is assumed to vary according to a Gauss-Markov model. If v is the speed traveled by a user then the correlation coefficient ρ_f is given by:

$$\rho_f = \exp\left(-\frac{2vf_c}{c}\right) \tag{7.72}$$

where f_c the carrier frequency and c the speed of light. We apply the correlation on the two component of the fast fading as follows:

$$F_1(t) = \rho_f F_1(t-1) + \sqrt{\frac{2}{\pi}\left(1-\rho_f^2\right)} Y_1(t)$$

$$F_2(t) = \rho_f F_2(t-1) + \sqrt{\frac{2}{\pi}\left(1-\rho_f^2\right)} Y_2(t)$$

where $Y_1(t)$ and $Y_2(t)$ are two random variables with normal distribution and the $\sqrt{\frac{2}{\pi}}$ as a correction that is added to guarantee a fast fading of average one.

In addition, $Y_1(t)$ and $Y_2(t)$ are generated based on the existing spatial correlation between the fast fading coefficients. We consier the correlation matrix R_{spat} defined in table B.2.2.1 of [113] with $\alpha = \beta = 0$ (i.e. low spatial correlation). In order to introduce spatial correlation into the channels coefficients according to a specific correlation matrix R_{spat}, we consider a vector of MIMO uncorrelated channel matrix (i.e. random variables with normal distribution) and multiply it by the lower triangular matrix corresponding to the Cholesky factorization of the correlation matrix R_{spat}.

The same model of fast fading is applied to D2D communications. The only difference is that the velocity v in equation 7.72 is modified as specified in section A.2.1.2.1 of [18].

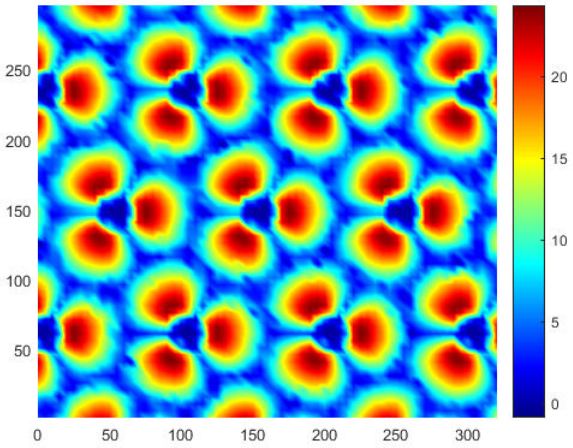
The **noise** parameters are the following: thermal noise = -174 dbm/Hz, the BS noise figure = 5 dB, UE noise figure = 9 dB and D2D noise figure = 9 dB.

On the UL, a **power control** is installed in order to reduce the power emitted by the users and limit by that the interference between neighbor cells. For a pathloss PL between the UE and the BS (i.e. taking into account the distance dependent pathloss, penetration loss, antenna pattern and shadowing loss), the user transmits at the following power:

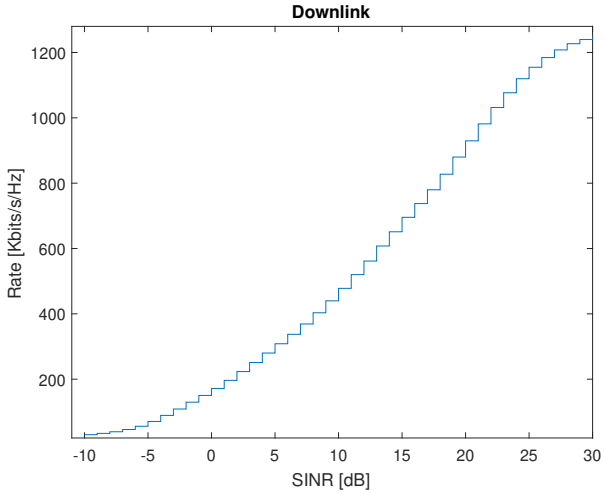
$$P = \min\{P_{max}, P_0 + \alpha PL\} \text{ in dBm}$$

where P_{max} is the maximum user's transmission power, $P_0 = -106$ dBm is the target received power at the BS and $\alpha = 1.0$ is the compensation factor. The same mechanism of Power control is applied for D2D communications.

We consider LTE codebook-based **precoding** for enabling transmit diversity mode. We assume that a single layer data stream is transmitted via two antennas. The following transmission schemes are simulated: 2×2 MIMO for DL and 1×2 MIMO for both UL and D2D communications. The precoding at the transmitter is done by choosing the precoding vector among 7.5 that maximizes the received power.



(a) SINR map for DL



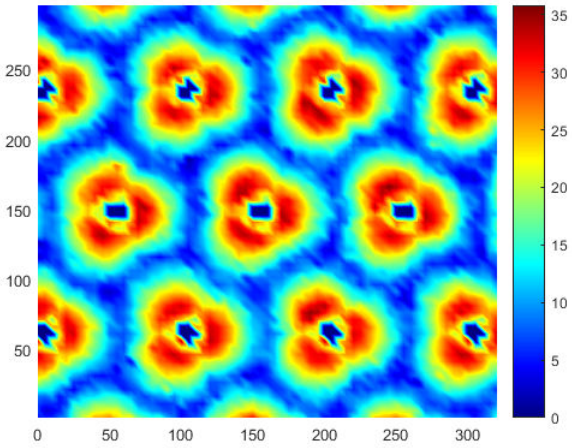
(b) SINR-rate mapping according to link-level simulations of DL communications

Codebook index	Rank 1	Rank 2
0	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$
1	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-1\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$
2	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\j\end{bmatrix}$	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\j & -j\end{bmatrix}$
3	$\frac{1}{\sqrt{2}}\begin{bmatrix}1\\-j\end{bmatrix}$	–

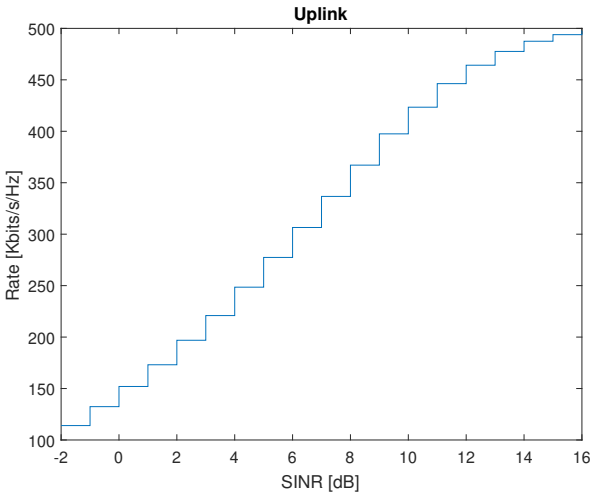
Table 7.5: Precoding vectors

Link curves

The SINR maps of the DL and UL channel models considered in the simulator are respectively given in figures 7.6a and 7.7a. Based on a link-level simulator, we deduce the mapping between the SINR and the rate (Kbits/s/Hz) for both DL (see figure 7.6b) and UL (see figure 7.7b) communications. We deduce the amount of data transmitted at each allocated RB. The UL link adaptation is applied to D2D communications; thus the UL SINR-rate mapping is considered for D2D links.



(a) SINR map for UL



(b) SINR-rate mapping according to link-level simulations of UL communications

Traffic Model

Considering a fix number of users, FTP2 model of traffic given in section A.2.1.3 of [59] is assumed (see figure 7.8). A DL file size of 0.5 MBytes and an UL file size of 0.25 Mbytes are considered. An exponential reading time D of parameter 0.2 separates the end of a file download and the arrival of the next file (i.e. the mean of D is 5 s).

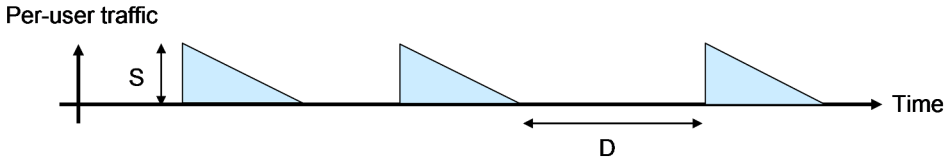


Figure 7.8: Traffic generation of FTP Model 2

Scheduling

We consider a round-robin scheduler that serves the users in a cycle way and independently of their radio conditions. At a given time-slot, only one UE is scheduled (i.e. all the available RBs are allocated to the selected UE). Since overlay D2D is assumed and the UL resources are often less utilized than the DL resources, we dedicate 20% of the UL resources for D2D communications.

D2D discovery

The discovery process allow a device to discover nearby users with whom D2D communication can be established. We implement the discovery process described in [18] and validate it by comparing our results to the one given in [18]. The following scenario is considered for evaluating the implementation of the discovery procedure. In each cell, among the 21 created macro cells, 150 UEs are dropped using one of the two dropping procedures (i.e. options 1 and 2) described in 7.6. The SINR of D2D links is illustrated in figure 7.9 and is compared to UL and DL SINR.

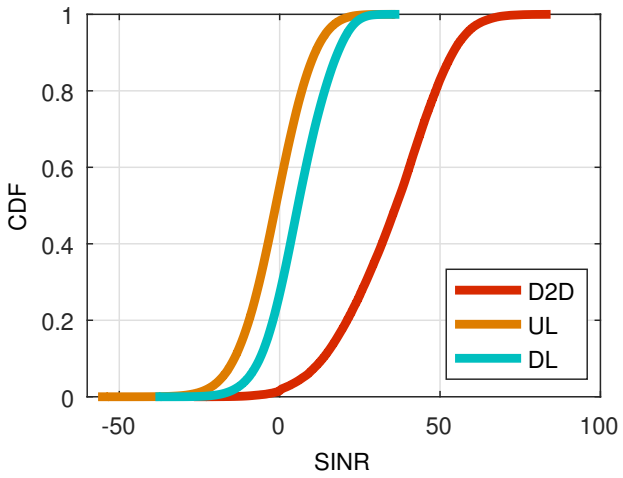
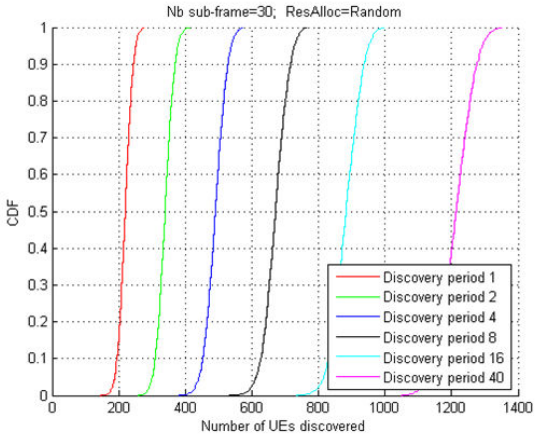


Figure 7.9: CDF of the D2D SINR

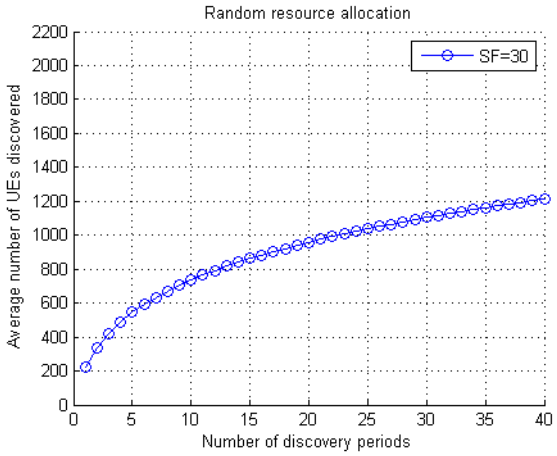
For each discovery period of 10 s, a limited number of subframes (i.e. 30 TTIs) are dedicated for the discovery process. Indeed, one RB pair is randomly allocated to each device among all the resources available in the 30 discovery subframes. Therefore, each user will transmit its discovery signal on its allocated physical RBs. When this discovery signal is decoded by a nearby user then the latter considers the device generating the discovery signal as a discovered user (i.e. minimum association received power is -107 dBm). Considering half duplex devices, then users transiting at the same TTI cannot discover each other. An in-band emissions described in table A.2.1.5-1 of [18] is evaluated in each non-allocated RB with $\{3, 6, 3, 3\}$ as W,X,Y,Z parameters .

For evaluating the discovery scheme described above, the two following metrics are considered and compared to those given in [18].

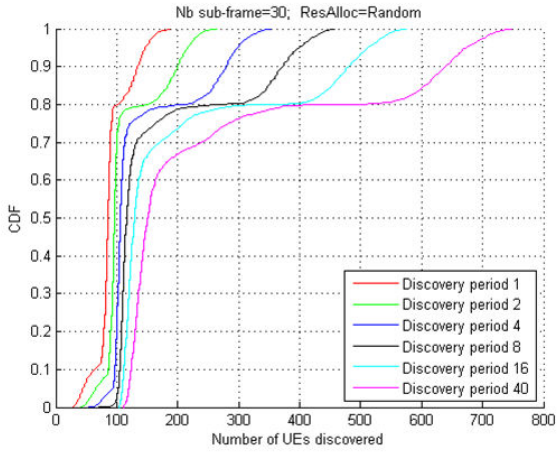
- CDF of the number of discovered UEs for different discovery periods (see figure 7.10a for layout option 1 and figure 7.11a for layout option 2).
- Number of discovered UEs as function of the number of discovery periods (see figure 7.10b for layout option 1 and figure 7.11b for layout option 2).



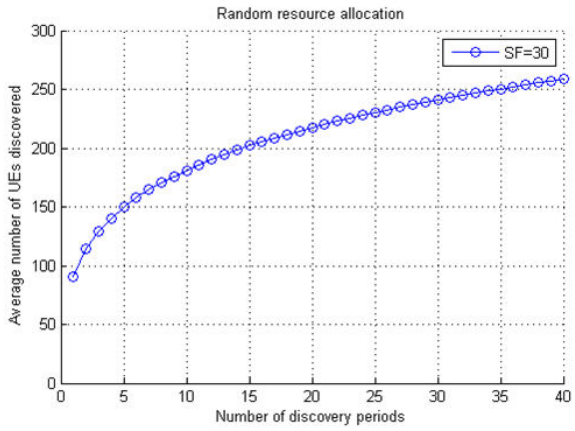
(a) CDF of the number of discovered UEs for different discovery periods for layout option 1



(b) Number of discovered UEs as function of the number of discovery periods for layout option 1



(a) CDF of the number of discovered UEs for different discovery periods for layout option 2



(b) Number of discovered UEs as function of the number of discovery periods for layout option 2

List of Publications

Journal Papers

1) R. Ibrahim, M. Assaad, B. Sayrac, A. Ephremides, "Stability Analysis of TDD Networks Revisited: A trade-off between Complexity and Precision", *journal submitted*, 2017

2) R. Ibrahim, M. Assaad, B. Sayrac, A. Gati, "Distributed vs. Centralized Scheduling in D2D-enabled Cellular Networks: Extended Version", *journal submitted*, 2018

3) R. Ibrahim, M. Assaad, B. Sayrac, "A Dynamic and Incentive Policy for Selecting D2D Mobile Relays", *journal submitted*, 2018

Conference Papers

4) [84] R. Ibrahim, M. Assaad, B. Sayrac, A. Ephremides, "Overlay D2D vs. Cellular communications: A stability region analysis", *2017 International Symposium on Wireless Communication Systems (ISWCS)*, Bologna, 2017, pp. 378-383..

5) [114] R. Ibrahim, M. Assaad, B. Sayrac, A. Gati, "When Distributed outperforms Centralized Scheduling in D2D-Enabled Cellular Networks", *2018 ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, Montreal, 2018.

6) R. Ibrahim, M. Assaad, B. Sayrac, "Greedy Algorithm for Selecting D2D Mobile Relays under Cost Constraints ", *paper submitted to the 2019 International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt)*, Avignon, 2019.

Patents

7) R. Ibrahim, B; Sayrac, M. Assaad, "Procédé centralisé d'allocation de ressources de transmission à des terminaux D2D dans un réseau d'accès cellulaire", Patent Number 1854184, filed on the 18^{th} of May 2018.

8) R. Ibrahim, B; Sayrac, M. Assaad, "Procédé distribué d'allocation de ressources de transmission à des terminaux D2D dans un réseau d'accès cellulaire", Patent Number 1854186, filed on the 18^{th} of May 2018.

Bibliography

[1] J. Liu, N. Kato, J. Ma, and N. Kadowaki, "Device-to-device communication in lte-advanced networks: A survey," *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 1923 – 1940, Fourthquarter 2015.

[2] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 4, pp. 1801 – 1819, Fourthquarter 2014.

[3] J. Andrews, S. Shakkottai, R. Heath, N. Jindal, M. Haenggi, R. Berry, D. Guo, M. Neely, S. Weber, S. Jafar, and A. Yener, "Rethinking information theory for mobile ad hoc networks," *IEEE Communications Magazine*, vol. 46, no. 12, pp. 94 – 101, December 2008.

[4] A. Orsino, G. Araniti, L. Militano, J. Alonso-Zarate, A. Molinaro, and A. Iera, "Energy efficient iot data collection in smart cities exploiting d2d communications," in *Sensors*, 2016.

[5] N. Vo, T. Q. Duong, H. D. Tuan, and A. Kortun, "Optimal video streaming in dense 5g networks with d2d communications," *IEEE Access*, vol. 6, pp. 209 – 223, 2018.

[6] X. Cheng, L. Yang, and X. Shen, "D2d for intelligent transportation systems: A feasibility study," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 4, pp. 1784 – 1793, Aug 2015.

[7] J. Lee and T. Q. S. Quek, "Device-to-device communication in wireless mobile social networks," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, May 2014, pp. 1 – 5.

[8] N. Golrezaei, A. F. Molisch, and A. G. Dimakis, "Base-station assisted device-to-device communications for high-throughput wireless video networks," *Proc. IEEE ICC*, 2012.

[9] J. Liu, Y. Kawamoto, H. Nishiyama, N. Kato, and N. Kadowaki, "Device-to-device communications achieve efficient load balancing in lte-advanced networks," 2014.

[10] J. Li, J. B. Song, and Z. Han, "Network connectivity optimiza- tion for device-to-device wireless system with femtocells," 2013.

[11] S. E. Elayoubi, A. M. Masucci, J. Roberts, and B. Sayraç, "Optimal d2d content delivery for cellular network offloading - special issue on device-to-device communication in 5g networks," *MONET*, vol. 22, pp. 1033 – 1044, 2017.

[12] T. V. Santana, R. Combes, and M. Kobayashi, "Device-to-device aided multicasting," *2018 IEEE International Symposium on Information Theory*, 2018.

[13] Y. Li, Z. Su, L. Huang, and W. Song, "A speed-aware joint handover approach for clusters of d2d devices," *Proceedings in IEEE Vehicular Technology Conference (VTC'18 Fall)*, 2018.

[14] G. Fodor, S. Parkvall, S. Sorrentino, P. Wallentin, Q. Lu, and N. Brahmi, "Device-to-device communications for national security and public safety," *IEEE Access*, vol. 2, pp. 1510 – 1520, 2014.

[15] X. Lin, J. G. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3gpp device-to-device proximity services," 2014.

[16] TR22803, "Feasibility study for proximity services (prose) (release 12)," 2012.

[17] TR23703, "Study on architecture enhancements to support proximity services (prose) (release 12)," 2013.

[18] TR36843, "Technical report 3GPP TR 36.843: Study on LTE Device-to-Device proximity services v12.0.1 2014," 2014.

[19] TS36201, "3gpp ts 36.201 evolved universal terrestrial radio access (e-utra); lte physical layer; general description," 2018.

[20] H. Xing and S. Hakola, "The investigation of power control schemes for a device-to-device communication integrated into ofdma cellular system," in *21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Sept 2010, pp. 1775 – 1780.

[21] X. Lin, J. G. Andrews, and A. Ghosh, "Spectrum sharing for device-to-device communication in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 12, pp. 6727 – 6740, Dec 2014.

[22] C. H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," 2009.

[23] B. Cho, K. Koufos, and R. Jäntti, "Spectrum allocation and mode selection for overlay d2d using carrier sensing threshold," in *2014 9th International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM)*, June 2014, pp. 26 – 31.

[24] K. Doppler, C. Yu, C. B. Ribeiro, and P. Janis, "Mode selection for device-to-device communication underlaying an lte-advanced network," in *2010 IEEE Wireless Communication and Networking Conference*, April 2010, pp. 1 – 6.

[25] K. Akkarajitsakul, P. Phunchongharn, E. Hossain, and V. K. Bhargava, "Mode selection for energy-efficient d2d communications in lte-advanced networks: A coalitional game approach," 2012.

[26] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, X. Cheng, and B. Jiao, "Efficiency resource allocation for device-to-device underlay communication systems: A reverse iterative combinatorial auction based approach," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 9, pp. 348–358, September 2013.

[27] R. Yin, G. Yu, C. Zhong, and Z. Zhang, "Distributed resource allocation for d2d communication underlaying cellular networks," in *2013 IEEE International Conference on Communications Workshops (ICC)*, June 2013, pp. 138–143.

[28] X. Ma, R. Yin, G. Yu, and Z. Zhang, "A distributed relay selection method for relay assisted device-to-device communication system," in *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC)*, Sept 2012, pp. 1020–1024.

[29] L. Wang, T. Peng, Y. Yang, and W. Wang, "Interference constrained relay selection of d2d communication for relay purpose underlaying cellular networks," in *2012 8th International Conference on Wireless Communications, Networking and Mobile Computing*, Sept 2012, pp. 1–5.

[30] R. M. Loynes, "The stability of a queue with non-independent inter-arrival and service times," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 58, no. 3, p. 497–520, 1962.

[31] M. J. Neely, *Stochastic network optimization with application to communication and queueing systems*. Morgan & Claypool Publishers, 2010, vol. 3, no. 1.

[32] C. P. C. Chanel, "Planning and perception in an uncertain environment mission: Application to the detection and recognition of targets by an autonomous helicopter," Theses, Institut Supérieur de l'Aéronautique et de l'Espace (ISAE), Apr. 2013. [Online]. Available: https://hal.archives-ouvertes.fr/tel-01296741

[33] R. D. Smallwood and E. Sondik, "The optimal control of partially observable markov processes over a finite horizon," *Operations research*, vol. 21, no. 5, pp. 1071–1088, 1973.

[34] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of markov decision processes," *Mathematics of Operations Research*, vol. 12, no. 3, pp. 441–450, 1987. [Online]. Available: http://www.jstor.org/stable/3689975

[35] O. Madani, S. Hanks, and A. Condon, "On the undecidability of probabilistic planning and infinite-horizon partially observable markov decision problems," in *Proceedings of the Sixteenth National Conference on Artificial Intelligence and the Eleventh Innovative Applications of Artificial Intelligence Conference Innovative Applications of Artificial Intelligence*, ser. AAAI '99/IAAI '99. Menlo Park, CA, USA: American Association for Artificial Intelligence, 1999, pp. 541–548. [Online]. Available: http://dl.acm.org/citation.cfm?id=315149.315395

[36] J. Pineau, G. Gordon, and S. Thrun, "Point-based value iteration: An anytime algorithm for pomdps," in *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, ser. IJCAI'03. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2003, pp. 1025–1030. [Online]. Available: http://dl.acm.org/citation.cfm?id=1630659.1630806

[37] A. R. Cassandra, "Exact and approximate algorithms for partially observable markov decision processes," 1998.

[38] T. Smith and R. Simmons, "Heuristic search value iteration for pomdps," in *Proceedings of the 20th Conference on Uncertainty in Artificial Intelligence*, ser. UAI '04. Arlington, Virginia, United States: AUAI Press, 2004, pp. 520–527. [Online]. Available: http://dl.acm.org/citation.cfm?id=1036843.1036906

[39] A. Cassandra, M. L. Littman, and N. L. Zhang, "Incremental pruning: A simple, fast, exact method for partially observable markov decision processes," in *In Proceedings of the Thirteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers, 1997, pp. 54–61.

[40] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial intelligence*, vol. 101, no. 1-2, pp. 99–134, 1998.

[41] M. Hauskrecht, "Value-function approximations for partially observable markov decision processes," *J. Artif. Int. Res.*, vol. 13, no. 1, pp. 33–94, Aug. 2000. [Online]. Available: http://dl.acm.org/citation.cfm?id=1622262.1622264

[42] M. L. Littman, A. R. Cassandra, and L. P. Kaelbling, "Learning policies for partially observable environments: Scaling up," in *Machine Learning Proceedings 1995*. Elsevier, 1995, pp. 362–370.

[43] W. S. Lovejoy, "Computationally feasible bounds for partially observed markov decision processes," *Operations Research*, vol. 39, no. 1, pp. 162–175, 1991. [Online]. Available: http://www.jstor.org/stable/171496

[44] G. L. Nemhauser, L. A. Wolsey, and M. L. Fisher, "An analysis of approximations for maximizing submodular set functions—i," *Mathematical programming*, vol. 14, no. 1, pp. 265–294, 1978.

[45] R. K. Iyer and J. A. Bilmes, "Submodular optimization with submodular cover and submodular knapsack constraints," in *Advances in Neural Information Processing Systems 26*, C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2013, pp. 2436–2444. [Online]. Available: http://papers.nips.cc/paper/4911-submodular-optimization-with-submodular-cover-and-submodular-knapsack-constraints.pdf

[46] M. Sviridenko, "A note on maximizing a submodular set function subject to a knapsack constraint," *Oper. Res. Lett.*, vol. 32, pp. 41–43, 2004.

[47] Y. Liu, R. Wang, and Z. Han, "Interference-constrained pricing for d2d networks," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 475–486, Jan 2017.

[48] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, and B. Jiao, "Interference-aware resource allocation for device-to-device communications as an underlay using sequential second price auction," in *Communications (ICC), 2012 IEEE International Conference on*. IEEE, 2012, pp. 445–449.

[49] N. Giatsoglou, A. Antonopoulos, E. Kartsakli, J. Vardakas, and C. Verikoukis, "Transmission policies for interference management in full-duplex d2d communication," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[50] N. Lee, X. Lin, J. G. Andrews, and R. W. Heath, "Power control for d2d underlaid cellular networks: Modeling, algorithms, and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 1, pp. 1–13, Jan 2015.

[51] H. ElSawy, E. Hossain, and M. S. Alouini, "Analytical modeling of mode selection and power control for underlay d2d communication in cellular networks," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 4147–4161, Nov 2014.

[52] X. Lin, R. Ratasuk, A. Ghosh, and J. G. Andrews, "Modeling, analysis, and optimization of multicast device-to-device transmissions," *IEEE Transactions on Wireless Communications*, vol. 13, no. 8, pp. 4346–4359, Aug 2014.

[53] C. Vitale, V. Mancuso, and G. Rizzo, "Modelling D2D communications in cellular access networks via coupled processors," Jan 2015, pp. 1–8.

[54] L. Lei, Y. Zhang, X. S. Shen, C. Lin, and Z. Zhong, "Performance analysis of device-to-device communications with dynamic interference using stochastic petri nets," *IEEE Transactions on Wireless Communications*, vol. 12, no. 12, pp. 6121–6141, December 2013.

[55] L. Su, Y. Ji, P. Wang, and F. Liu, "Resource allocation using particle swarm optimization for d2d communication underlay of cellular networks," pp. 129–133, April 2013.

[56] M. H. Han, B. G. Kim, and J. W. Lee, "Subchannel and transmission mode scheduling for d2d communication in ofdma networks," pp. 1–5, Sept 2012.

[57] M. Jung, K. Hwang, and S. Choi, "Joint mode selection and power allocation scheme for power-efficient device-to-device (d2d) communication," pp. 1–5, May 2012.

[58] S. C. Spinella, G. Araniti, A. Iera, and A. Molinaro, "Integration of ad-hoc networks with infrastructured systems for multicast services provisioning," pp. 1–6, 2009.

[59] TR36814, "Evolved universal terrestrial radio access (e-utra); further advancements for e-utra physical layer aspects release 9 v9.2.0 2017," 2017.

[60] G. Standardization, "Terrestrial radio access (e-utra); further advancements for e-utra physical layer aspects," Release 12- 2014.

[61] Q. Incorporated, "3GPP contribution R1-142940: Discussion on ProSe provisioning parameter," 2014.

[62] D. J. Love, R. W. Heath, W. Santipach, and M. L. Honig, "What is the value of limited feedback for MIMO channels?" *IEEE Communications Magazine*, vol. 42, no. 10, pp. 54–59, Oct 2004.

[63] L. Sivridis and J. He, "A strategy to reduce the signaling requirements of CQI feedback schemes," *Wireless Personal Communications*, vol. 70, no. 1, pp. 85–98, May 2013. [Online]. Available: https://doi.org/10.1007/s11277-012-0680-9

[64] H. Kim and Y. Han, "An opportunistic channel quality feedback scheme for proportional fair scheduling," *IEEE Communications Letters*, vol. 11, no. 6, pp. 501–503, June 2007.

[65] J. Jeon, K. Son, H. W. Lee, and S. Chong, "Feedback reduction for multiuser OFDM systems," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 160–169, Jan 2010.

[66] D. J. Love, R. W. Heath, V. K. N. Lau, D. Gesbert, B. D. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, pp. 1341–1365, October 2008.

[67] Q. Ye, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "Resource optimization in device-to-device cellular systems using time-frequency hopping," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5467–5480, Oct 2014.

[68] S. Maghsudi and S. Stańczak, "Hybrid centralized and distributed resource allocation for device-to-device communication underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 4, pp. 2481–2495, April 2016.

[69] R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference-aware graph based resource sharing for device-to-device communications underlaying cellular networks," April 2013, pp. 140–145.

[70] A. Aijaz, M. Tshangini, M. R. Nakhai, X. Chu, and A. Aghvami, "Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence under statistical qos guarantees," *Communications, IEEE Transactions on*, vol. 62, pp. 2353–2365, 07 2014.

[71] M. Zulhasnine, C. Huang, and A. Srinivasan, "Efficient resource allocation for device-to-device communication underlaying LTE network," *IEEE 6th International Conference on Wireless and Mobile Computing, Networking and Communications*, pp. 368 – 375, Oct 2010.

[72] T. Han, R. Yin, Y. Xu, and G. Yu, "Uplink channel reusing selection optimization for device-to-device communication underlaying cellular networks," *2012 IEEE 23rd International Symposium on Personal, Indoor and Mobile Radio Communications - (PIMRC)*, pp. 559 – 564, Sept 2012.

[73] L. Su, Y. Ji, P. Wang, and F. Liu, "Resource allocation using particle swarm optimization for D2D communication underlay of cellular networks," *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 129 – 133, April 2013.

[74] G. Yu, L. Xu, D. Feng, R. Yin, G. Y. Li, and Y. Jiang, "Joint mode selection and resource allocation for device-to-device communications," *IEEE Transactions on Communications*, vol. 62, no. 11, pp. 3814 – 3824, Nov 2014.

[75] D. Feng, L. Lu, Y. W. Yi, G. Y. Li, G. Feng, and S. Li, "QoS-aware resource allocation for device-to-device communications with channel uncertainty," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, Aug 2016.

[76] H. Tang and Z. Ding, "Monotonic optimization for power control of d2d underlay with partial csi," *2016 IEEE International Conference on Communications (ICC)*, pp. 1 – 6, May 2016.

[77] A. Liu, V. Lau, F. Zhuang, and J. Chen, "Mixed timescale cross-layer optimization for multi-antenna D2D networks," *2015 IEEE Global Communications Conference (GLOBECOM)*, pp. 1 – 6, Dec 2015.

[78] Q. Ye, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "Distributed resource allocation in device-to-device enhanced cellular networks," *IEEE Transactions on Communications*, vol. 63, no. 2, pp. 441 – 454, Feb 2015.

[79] Y. Li, D. Jin, J. Yuan, and Z. Han, "Coalitional games for resource allocation in the device-to-device uplink underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3965 – 3977, July 2014.

[80] R. Yin, C. Zhong, G. Yu, Z. Zhang, K. Wong, and X. Chen, "Joint spectrum and power allocation for D2D communications underlaying cellular networks," *IEEE Transactions on Vehicular Technology*, 2015.

[81] X. Xiao and J. L. X. Tao and, "A qos-aware power optimization scheme in ofdma systems with integrated device-to-device (d2d) communications," in *2011 IEEE Vehicular Technology Conference (VTC Fall)*, Sept 2011, pp. 1 – 5.

[82] J. M. B. da Silva, G. Fodor, and T. F. Maciel, "Performance analysis of network-assisted two-hop D2D communications," *IEEE Globecom Workshops (GC Wkshps)*, pp. 1050 – 1056, Dec 2014.

[83] F. Wang, C. Xu, L. Song, and Z. Han, "Energy-efficient resource allocation for device-to-device underlay communication," *IEEE Transactions on Wireless Communications*, vol. 14, no. 4, April 2015.

[84] R. Ibrahim, M. Assaad, B. Sayrac, and A. Ephremides, "Overlay d2d vs. cellular communications: A stability region analysis," in *2017 International Symposium on Wireless Communication Systems (ISWCS)*, Aug 2017, pp. 378–383.

[85] T. 3rd Generation Partnership Project, "Technical specification 3GPP TS 36.213: Evolved universal terrestrial radio access (E-UTRA); physical layer procedures," 2009.

[86] J. Singh, Z. Pi, and H. Nguyen, "Low-complexity optimal CSI feedback in LTE," *IEEE 10th Consumer Communications and Networking Conference (CCNC)*, pp. 472–478, Jan 2013.

[87] Y. Bougurn, E. Hardouin, and F. Wolff, *LTE et les réseaux 4G*. Groupe Eyrolles, 2012.

[88] T. Chen, H. Kim, and Y. Yang, "Energy efficiency metrics for green wireless communications," in *2010 International Conference on Wireless Communications Signal Processing (WCSP)*, Oct 2010, pp. 1–6.

[89] Y. Chen, P. Martins, L. Decreusefond, F. Yan, and X. Lagrange, "Stochastic analysis of a cellular network with mobile relays," in *2014 IEEE Global Communications Conference*, Dec 2014, pp. 4758–4763.

[90] J. Liu and N. Kato, "Device-to-device communication overlaying two-hop multi-channel uplink cellular networks," in *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, ser. MobiHoc '15. New York, NY, USA: ACM, 2015, pp. 307–316. [Online]. Available: http://doi.acm.org/10.1145/2746285.2746311

[91] K. Vanganuru, S. Ferrante, and G. Sternberg, "System capacity and coverage of a cellular network with d2d mobile relays," in *MILCOM 2012 - 2012 IEEE Military Communications Conference*, Oct 2012, pp. 1–6.

[92] H. Lin, D. Gu, W. Wang, and H. Yang, "Capacity analysis of dedicated fixed and mobile relay in lte-advanced cellular networks," in *2009 IEEE International Conference on Communications Technology and Applications*, Oct 2009, pp. 354–359.

[93] L. Babun, A. . Yürekli, and . Güvenç, "Multi-hop and d2d communications for extending coverage in public safety scenarios," in *2015 IEEE 40th Local Computer Networks Conference Workshops (LCN Workshops)*, Oct 2015, pp. 912–919.

[94] R. Atat, L. Liu, N. Mastronarde, and Y. Yi, "Energy harvesting-based d2d-assisted machine-type communications," *IEEE Transactions on Communications*, vol. 65, no. 3, pp. 1289–1302, March 2017.

[95] M. Hasan and E. Hossain, "Distributed resource allocation for relay-aided device-to-device communication under channel uncertainties: A stable matching approach," *IEEE Transactions on Communications*, vol. 63, no. 10, pp. 3882–3897, Oct 2015.

[96] J. M. B. da Silva, G. Fodor, and T. F. Maciel, "Performance analysis of network-assisted two-hop d2d communications," in *2014 IEEE Globecom Workshops (GC Wkshps)*, Dec 2014, pp. 1050–1056.

[97] D. Niyato and P. Wang, "Optimization of the mobile router and traffic sources in vehicular delay-tolerant network," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 5095–5104, Nov 2009.

[98] Z. Zhang, X. Chen, Y. Zhang, C. Zhong, and W. Wang, "Energy-efficient opportunistic packet scheduling in mobile relay systems," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5327–5336, July 2016.

[99] Y. Wei, F. R. Yu, and M. Song, "Distributed optimal relay selection in wireless cooperative networks with finite-state markov channels," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 5, pp. 2149–2158, Jun 2010.

[100] S. Yun, K. Lee, S. Park, and J. K. Choi, "Energy efficient relay selection scheme with drx mechanism in 3gpp lte network," in *2013 International Conference on ICT Convergence (ICTC)*, Oct 2013, pp. 6–11.

[101] T. Kim and M. Dong, "An iterative hungarian method to joint relay selection and resource allocation for d2d communications," *IEEE Wireless Communications Letters*, vol. 3, no. 6, pp. 625–628, Dec 2014.

[102] R. Ma, Y. Chang, H. Chen, and C. Chiu, "On relay selection schemes for relay-assisted d2d communications in lte-a systems," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 9, pp. 8303–8314, Sept 2017.

[103] Y. Li, P. Wang, D. Niyato, and W. Zhuang, "A dynamic relay selection scheme for mobile users in wireless relay networks," in *2011 Proceedings IEEE INFOCOM*, April 2011, pp. 256–260.

[104] Y. Li, Y. Hou, Z. Huang, and Y. Wei, "Cooperative relay selection policy using partially observable markov decision process," in *2011 Seventh International Conference on Natural Computation*, vol. 1, July 2011, pp. 508–512.

[105] Y. Polyanskiy, H. V. Poor, and S. Verdu, "Channel coding rate in the finite blocklength regime," *IEEE Transactions on Information Theory*, vol. 56, no. 5, pp. 2307–2359, May 2010.

[106] J. D. Isom, S. P. Meyn, and R. D. Braatz, "Piecewise linear dynamic programming for constrained pomdps," in *AAAI*, vol. 1, 2008, pp. 291–296.

[107] D. Kim, J. Lee, K. E. Kim, and P. Poupart, "Point-based value iteration for constrained pomdps," in *IJCAI*, 2011, pp. 1968–1974.

[108] J. L. D. Lagrange, *Nouvelle méthode pour résoudre les équations littérales par le moyen des séries*. Chez Haude et Spener, Libraires de la Cour & de l'Académie royale, 1770.

[109]

[110] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource allocation and cross-layer control in wireless networks*, 2006.

[111] D. A. Levin and Y. Peres, *Markov chains and mixing times*. American Mathematical Soc., 2017, vol. 107.

[112] N. Abbas, "Design and performance of coordination schemes in cellular networks," Theses, Télécom ParisTech, Nov. 2016. [Online]. Available: https://pastel.archives-ouvertes.fr/tel-01618909

[113] TR36803, "Evolved universal terrestrial radio access (e-utra); user equipment (ue) radio transmission and reception v 1.1.0 2008," 2008.

[114] R. Ibrahim, M. Assaad, B. Sayrac, and A. Gati, "When distributed outperforms centralized scheduling in d2d-enabled cellular networks," in *Proceedings of the 21st ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, ser. MSWIM '18. New York, NY, USA: ACM, 2018, pp. 189–198. [Online]. Available: http://doi.acm.org/10.1145/3242102.3242115

université PARIS-SACLAY
ÉCOLE DOCTORALE
Sciences et technologies de l'information et de la communication (STIC)

Titre: Utilisation des communications Device-to-Device pour améliorer l'efficacité des réseaux cellulaires

Mots clés: communications directes entre mobiles D2D, Réseaux Cellulaires, sélection de mode de communication, allocation des ressources, sélection des relais, théorie des files d'attente, optimisation Lyapunov, processus de décision markovien partiellement observable POMDP.

Résumé: Cette thèse étudie les communications directes entre les mobiles, appelées communications D2D, en tant que technique prometteuse pour améliorer les futurs réseaux cellulaires. Cette technologie permet une communication directe entre deux terminaux mobiles sans passer par la station de base. La modélisation, l'évaluation et l'optimisation des différents aspects des communications D2D constituent les objectifs fondamentaux de cette thèse et sont réalisés principalement à l'aide des outils mathématiques suivants: la théorie des files d'attente, l'optimisation de Lyapunov et les processus de décision markovien partiellement observable POMDP. Les résultats de cette étude sont présentés en trois parties. Dans la première partie, nous étudions un schéma de sélection entre mode cellulaire et mode D2D. Nous dérivons les régions de stabilité des scénarios suivants: réseaux cellulaires purs et réseaux cellulaires où les communications D2D sont activées. Une comparaison entre ces deux scénarios conduit à l'élaboration d'un algorithme de sélection entre le mode cellulaire et le mode D2D qui permet d'améliorer la capacité du réseau. Dans la deuxième partie, nous développons un algorithme d'allocation de ressources des communications D2D. Les utilisateurs D2D sont en mesure d'estimer leur propre qualité de canal, cependant la station de base a besoin de recevoir des messages de signalisation pour acquérir cette information. Sur la base de cette connaissance disponibles au niveau des utilisateurs D2D, une approche d'allocation des ressources est proposée afin d'améliorer l'efficacité énergétique des communications D2D. La version distribuée de cet algorithme s'avère plus performante que celle centralisée. Dans le schéma distribué des collisions peuvent se produire durant la transmission de l'état des canaux D2D ; ainsi un algorithme de réduction des collisions est élaboré. En outre, la mise en œuvre des algorithmes centralisé et distribué dans un réseau cellulaire, type LTE, est décrite en détails.Dans la troisième partie, nous étudions une politique de sélection des relais D2D mobiles. La mobilité des relais représente un des principaux défis que rencontre toute stratégie de sélection de relais. Le problème est modélisé par un processus contraint de décision markovien partiellement observable qui prend en compte le dynamisme des relais et vise à trouver la politique de sélection de relais qui optimise la performance du réseau cellulaire sous des contraintes de coût.

Title: Use of Device-to-Device communications for efficient cellular networks

Keywords: Device-to-Device (D2D) communications, cellular networks, mode selection, resource allocation, relay selection, queuing theory, Lyapunov optimization, Partially Observed Markov Decision Process (POMDP).

Abstract: This thesis considers Device-to-Device (D2D) communications as a promising technique for enhancing future cellular networks. Modeling, evaluating and optimizing D2D features are the fundamental goals of this thesis and are mainly achieved using the following mathematical tools: queuing theory, Lyapunov optimization and Partially Observed Markov Decision Process (POMDP). The findings of this study are presented in three parts. In the first part, we investigate a D2D mode selection scheme. We derive the queuing stability regions of both scenarios: pure cellular networks and D2D-enabled cellular networks. Comparing both scenarios leads us to elaborate a D2D vs cellular mode selection design that improves the capacity of the network. In the second part, we develop a D2D resource allocation algorithm. We observe that D2D users are able to estimate their local Channel State Information (CSI), however the base station needs some signaling exchange to acquire this information. Based on the D2D users' knowledge of their local CSI, we provide an energy efficient resource allocation framework that shows how distributed scheduling outperforms centralized one. In the distributed approach, collisions may occur between the different CSI reporting; thus, we propose a collision reduction algorithm. Moreover, we give a detailed description on how both centralized and distributed algorithms can be implemented in practice. In the third part, we propose a mobile relay selection policy in a D2D relay-aided network. Relays' mobility appears as a crucial challenge for defining the strategy of selecting the optimal D2D relays. The problem is formulated as a constrained POMDP which captures the dynamism of the relays and aims to find the optimal relay selection policy that maximizes the performance of the network under cost constraints.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Rue Joliot-Curie 91190 Saint-Aubin, France