SAPIENZA
UNIVERSITÀ DI ROMA

Control methods and applicative scenarios for next-generation cellular telecommunication networks

Scuola di Scienza e Tecnologia dell'Informazione e delle Comunicazioni
Dottorato di Ricerca in AUTOMATICA, BIOINGEGNERIA E
RICERCA OPERATIVA (ABRO) - Curriculum in Automatica (XXXV cycle)

Emanuele De Santis
ID number 1664777

Advisor
Prof. Francesco Delli Priscoli

Co-Advisors
Prof. Alessandro Di Giorgio
Prof. Antonio Pietrabissa

Director
Prof. Giuseppe Oriolo

Academic Year 2021/2022

Thesis defended on 25 January 2023
in front of a Board of Examiners composed by:

Prof.ssa Paola Cappanera (chairman)

Prof. Danilo Pani

Prof. Paolo Valigi

:

Prof. Alessandro Freddi

Dr. Andrea Michiorri

Control methods and applicative scenarios for next-generation cellular telecommunication networks

PhD Thesis. Sapienza University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: edesantis@diag.uniroma1.it

Acknowledgments

(intentionally left blank for the public version)

Abstract

The increasing demand for connectivity services to the mobile (cellular) network, together with the co-existence of former terrestrial mobile networks and the convergence of satellite telecommunication systems, lead the telecommunication operators to find solutions for enabling such convergence, by also making use of former cellular networks present on the field. Moreover, the introduction of next-generation cellular telecommunication systems leads to the need for telco operators to find new industrial use cases, which may enable new kinds of services both at the industrial and at the network level. This thesis discusses some control methodologies to be applied to facilitate the convergence of terrestrial and non-terrestrial networks in a multi-Radio Access Technology (RAT) environment, where different Radio Access Technologies are available at the same time, providing multi-connectivity services at increasing bandwidth and reduced latency, by also considering the users' perceived Quality of Experience (QoE). Such control techniques are mainly model-free and are based on Game Theory arguments and on Reinforcement Learning, and address two different problems: the network selection (i.e., deciding the best Access Point (AP) to serve a User Equipment (UE) request) and the dynamic traffic splitting and steering (i.e., in a multi-connectivity context, the problem of deciding the quantity of traffic of each UE to be sent to each of the connected AP). Moreover, an applicative scenario of 5G network for smart grid control, and in particular for the provisioning of Frequency Regulation services using charging plug-in electric vehicles (PEVs) has been proposed in this thesis, by analyzing the regulatory framework, the technical feasibility, and the economic feasibility of the proposed approach. All the proposed approaches are provided with extensive simulations to validate the concepts and the proposed control algorithms. In addition, a 5G multi-RAT radio access network simulator has been developed in the context of the work carried out by the Candidate for this thesis, in order to validate the proposed approaches in a realistic environment. Some of the proposed algorithms have been / will be also tested in real environments in the context of the activities of the H2020 project 5G-ALLSTAR and 5G-Solutions, which partially supported the ideas, algorithms, and results proposed in this thesis.

Contents

Acronyms

aFRR Frequency Restoration Reserves with Automatic activation

AP Access Point

ATSSS Access Traffic Steering, Switching and Splitting

BER Bit Error Rate

BRP Balance Responsible Party

BSP Balance Service Provider

BS Base Station

CAC Connection Admission Controller

CN Core Network

CPO Charging Point Operator

CP Capacity-constrained Problem

CS charging station

CU Centralized Unit

DeepRL Deep Reinforcement Learning

DQN Deep Q-Network

DSM Demand Side Managment

DSO Distribution System Operator

DU Distributed Unit

EVSE Electric Vehicle Supply Equipment

FSPL Free Space Path Loss

GEO Geostationary Earth Orbit

HAPS High Altitude Platform Station

ICT information and communication technologies

LAPS Low Altitude Platform Station

LEO Low Earth Orbit

LSTM Long Short-Term Memory

LTE Long Term Evolution

MADM Multiple Attribute Decision Making

MDP Markov Decision Process

MEC Multi-Access Edge Computing

NFV Network Function Virtualization

NR New Radio

NTN Non-Terrestial Network

OFDM Orthogonal Frequency Division Multiplexing

PCF Policy Control Function

PDU Protocol Data Unit

PEV plug-in electric vehicle

PRB Physical Resource Block

QoE Quality of Experience

QoS Quality of Service

RAN Radio Access Network

RAT Radio Access Technology

RBUR Resource Blocks Utilization Ratio

RES Renewable Energy Source

RL Reinforcement Learning

RSRP Reference Signal Received Power

SDN Software Defined Networking

SINR Signal over Interference plus Noise Ratio

SNR Signal to Noise Ratio

SOC state-of-charge

TDM Time Division Multiplexing

TSO Transmission System Operator

UAV Unmanned Aerial Vehicle

UE User Equipment

UPF User Plane Function

URLLC Ultra-Reliable Low-Latency Communication

UVAM Mixed Enabled Virtual Unit

List of Figures

List of Tables

Chapter 1

Introduction

This thesis contains the main activities and results the Candidate has carried out during his 3-year PhD, which has been developed in the context of the activities carried out by the Network Control Laboratory at the Department of Computer, Control and Management Engineering "Antonio Ruberti" of Sapienza University of Rome, and by the no-profit research consortium CRAT (Consortium for the Research in Automation and Telecommunication), which is composed by Sapienza University of Rome, University of Sannio, Politechnic University of Bari, Thales Alenia Space Italy and TopNetwork.

The proposed solutions, developed by the Candidate during his PhD, are partially supported by the EU-Korea H2020 project 5G-ALLSTAR (grant n. 815323) and by the EU H2020 project 5G-Solutions (grant n. 856691).

This thesis, as well as the work of the Candidate during his PhD, is focused on the control of next-generation cellular networks, with particular reference to 5G and 6G multi-RAT networks. At the same time, the candidate deepened into some applicative scenarios of such next-generation cellular networks, with a particular focus on the smart grids use case.

1.1 Motivation

By the beginning of the digital era, an ever-increasing demand for bandwidth, low latency, and, in general, connectivity has driven the evolution of fixed and mobile networks and their requirements. Mobile networks, which were traditionally used only to provide phone services and a few other services (like SMS, MMS, etc), became of crucial importance since the introduction of smartphones in the market. Indeed, the final users started to require more and more powerful connectivity services from such networks (e.g., broadband services [4, 5], low-latency [6, 7], fast mobility [8,9], etc.), driving the evolution of cellular networks to sustain such kind of

requirements from the users and the market. In the very latest years, 5G networks began to appear in the cities, providing enhanced broadband and very-low latency, which should sustain new kinds of applications (e.g., Virtual Reality, Augmented Reality, High-Resolution Video Streaming, etc.). Moreover, given the presence of multiple terrestrial RATs available at the same time (e.g., 4G, 3G) and with the evolution of satellite telecommunication systems (which now sustain capacities comparable to terrestrial network ones [10, 11]), it is now possible to better balance the users' requests among the available RATs, based on the QoE/QoS requested by the specific applications. In addition, the convergence of satellite communication systems with terrestrial 5G networks may provide further benefits in terms of radio coverage (e.g., in rural areas) and traffic offloading. Finally, with the introduction of Unmanned Aerial Vehicles (UAVs), High Altitude Platform Stations (HAPSs), and Low Altitude Platform Station (LAPS), it is possible to develop new 6G networks that exploit such solutions to provide 3D connectivity services and to provide intelligence to the network to reconfigure itself to mitigate or reduce congestion at the radio level.

Given such a context, it is fundamental to control physical resources at the radio level to provide the required performance to the end users, both for the access to the radio access network (i.e., choosing the best AP to connect to in a multi-RAT environment) and for the handling of the connections (i.e., dynamically adapt the data rate - and so the required physical resources- based on the condition of the network).

At the same time, it is fundamental to define new industrial use cases for next-generation radio access networks, to identify new business opportunities both for telco operators and for industry, which may drive the design of 6G networks and beyond to the definition of new requirements and services.

1.2 Objectives, contributions and thesis structure

The objective of this thesis is to provide a set of control methodologies for next-generation cellular radio access network control, and in particular for 5G and 6G networks composed by multiple RATs. Moreover, an applicative scenario of such networks in the field of smart grid control, exploiting the new capabilities offered by 5G and foreseen by 6G, has been proposed. The aim of this thesis is to highlight the potentiality of next-generation cellular networks both from a control perspective, by applying model-free techniques like Reinforcement Learning, and from an industrial one, and in particular in a field (the power systems one) which is considered not to be a killer application for 5G/6G networking services, but instead may reap great

benefit from such new cellular telecommunication systems.

This thesis is structured as follows:

- Part I deals with several control techniques for next-generation radio access networks control and is divided in:
  - Chapter 2, where a Game Theory based control methodology for 5G traffic splitting and steering in multi-RAT networks is detailed, making use of the concept of Wardrop equilibria to equalize a generic latency function among the different APs;
  - Chapter 3, where a Multi-Agent Reinforcement Learning technique is applied for 5G multi-RAT network selection, considering the various APs as competing entities in serving the users' requests;
  - Chapter 4, where a Deep Reinforcement Learning technique is proposed for 5G multi-RAT network selection, this time considering the various APs as coordinated entities (controlled by a Radio Access Network (RAN) controller) in allocating the requests from the UEs;
  - Chapter 5, where a vision on 6G multi-RAT networks, exploiting 3D connectivity and mobile APs is provided, together with a proposed system architecture and a proof-of-concept of the advantages brought by mobile APs;
- Part II deals with an applicative scenario of next-generation radio access networks, and in particular the one of Frequency Regulation for smart grids by using charging PEVs:
  - Chapter 6 provides the fundamentals of Frequency Regulation services, by providing the relevant mathematical formulation for the dynamics of the electricity network frequency;
  - Chapter 7 details the idea of using charging PEVs for the provisioning of Frequency Regulation services, by providing a detailed control architecture, a precise delay budget to make the service feasible, and some possible approaches for the computation of the power-frequency curves for each PEV participating in the service;
  - Chapter 8 describes a proposed controller based on a linear optimization problem to compute and assign local power-frequency curves to the PEVs, differentiating the effort of each PEV in the provisioning of the Frequency Regulation service based on both user constraints (desired state-of-charge (SOC) and remaining dwelling time), and on network

constraints (global droop curve agreed by the Transmission System Operator (TSO)).

- Chapter 9 draws the conclusion of this thesis and the future works that will be carried out by the Candidate.

1.3 Contributing publications

The contents of this thesis have been disseminated in the following publications:

- F. Delli Priscoli, E. De Santis, A. Giuseppi, and A. Pietrabissa, "Capacity-constrained wardrop equilibria and application to multi-connectivity in 5g networks," *Journal of the Franklin Institute*, vol. 358, no. 17, pp. 9364–9384, Nov. 2021. [Online]. Available: `https://doi.org/10.1016/j.jfranklin.2021.09.025`.

- A. Giuseppi, E. De Santis, F. Delli Priscoli, S. H. Won, T. Choi, and A. Pietrabissa, "Network selection in 5g networks based on markov games and friend-or-foe reinforcement learning," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, Apr. 2020. [Online]. Available: `https://doi.org/10.1109/wcncw48565.2020.9124723`.

- E. De Santis, A. Giuseppi, A. Pietrabissa, M. Capponi, and F. Delli Priscoli, "Satellite integration into 5g: Deep reinforcement learning for network selection," *Machine Intelligence Research*, vol. 19, no. 2, pp. 127–137, Apr. 2022. [Online]. Available: `https://doi.org/10.1007/s11633-022-1326-3`.

- E. Calvanese Strinati, S. Barbarossa, T. Choi, A. Pietrabissa, A. Giuseppi, E. De Santis, J. Vidal, Z. Becvar, T. Haustein, N. Cassiau, F. Costanzo, J. Kim, and I. Kim, "6g in the sky: On-demand intelligence at the edge of 3d networks (invited paper)," *ETRI Journal*, vol. 42, no. 5, pp. 643–657, Oct. 2020. [Online]. Available: `https://doi.org/10.4218/etrij.2020-0205`.

- R. Germanà, E. D. Santis, F. Liberati, and A. Di Giorgio, "On the participation of charging point operators to the frequency regulation service using plug-in electric vehicles and 5g communications," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*. IEEE, Sep. 2021. [Online]. Available: `https://doi.org/10.1109/eeeic/icpseurope51590.2021.9584495`.

- R. Germanà, F. Liberati, E. De Santis, A. Giuseppi, F. Delli Priscoli, and A. Di Giorgio, "Optimal control of plug-in electric vehicles charging for composition of frequency regulation services," *Energies*, vol. 14, no. 23, p. 7879, Nov. 2021. [Online]. Available: `https://doi.org/10.3390/en14237879`

Part I

Techniques for cellular network control

Chapter 2

Wardrop equilibria for traffic splitting and steering in 5G heterogeneous networks

This chapter details a first work in the field of radio access control for cellular networks, based on Game Theory and, in particular, on the concept of Wardrop equilibria. In this work, a set of downlink data flows (commodities in the following) coming from a set of UE is split/steered to a set of AP, exploiting the concept of multi-connectivity envisaged in 5G radio access networks, so that the same data flow can be divided into sub-flows associated to different APs to equalize a certain *latency function* that can be chosen by the Network Manager so to balance a certain parameter of the radio access network among the flows. In this work, the latency function used aims at balancing the number of Physical Resource Blocks (PRBs) used by each flow in each AP, and each flow (initially split among several APs) is then steered between the available APs to equalize the number of PRBs used by each subflow, and so at the end equalize the number of PRBs allocated in each AP. The algorithm presented in this chapter has been proved to converge to a Wardrop user equilibrium and the simulations presented in Section 2.3.2. validate the proposed controller using a 5G New Radio (NR) radio access network simulator developed ad-hoc by the Candidate for validating control algorithms on realistic radio access network scenarios. The same approach could be applied for uplink data flows without loss of generality from a mathematical and technical point of view.

2.1 Introduction

Load Balancing is a classic problem of network control and can be interpreted as a particular case of traffic routing with providers representing unitary paths and latency functions describing the performance of each provider. In adversarial (or selfish) routing, the control algorithms are aimed at leading the network into convenient equilibrium states without the cooperation of its agents. One of such states is known in mean-field game theory as Wardrop equilibrium (which can be regarded as a Nash equilibrium for infinite players [12]): in such state, the latencies experienced by the agents that constitute the traffic flows are equalized over all their available routes, and, as a consequence, no agent may improve its routing unilaterally. In this work, a particular case of selfish capacitated load balancing, in which the capacities of the service providers are limited has been studied. Therefore, as it will be discussed, the proposed control law objective will be to equalize the latencies of all the providers which are not saturated. This network state is a generalization of the Wardrop equilibrium in capacitated networks and is known in the literature as the Beckmann user equilibrium [13].

Multi-connectivity is an emerging challenge in the heterogeneous network scenario envisaged by 5G, where multiple RATs, such as LTE, 5G, and Satellite networks, are available to connect the network users to the core network [14]. According to the multi-connectivity paradigm, each UE may be able to be served by several of the various AP of the available RATs, potentially at the same time. The problem, referred to in the 5G literature as *multi-connectivity*, consists in dynamically choosing which acpAP shall serve each UE and deciding how much traffic relevant to each UE shall be routed through each of the serving APs. This work focuses on the downlink direction, i.e., it refers to the traffic transmitted from the core network to the UEs via the APs; nevertheless, similar considerations apply when considering the uplink direction.

In this work, the performance of the network APs is measured in terms of *latency functions* that capture the amount of resources (in terms of resource blocks) required from each AP to serve the various *commodities*. In the considered 5G scenario, such commodities consist in the so-called QoS-Flows, which are streams of data toward a UE that are characterized by standardized QoS requirements (e.g., bit error rate, maximum tolerated delay...). In general, the latency functions may account for different connection-specific performance indexes (e.g., amount of network resources utilized on a given AP, power consumption, service reliability), and may include additional factors, such as operator preferences or different usage tariffs.

Overall, the objective of the proposed control law for load balancing is to dynamically *steer* the downlink traffic in such a way that the values of the latency

functions are equalized.

The described scenario is typical in adversarial routing and load balancing problems, as the various connections are not concerned with the overall network state and aim at optimizing their own, individual, performances. The two main problems in the algorithm development are i) the fact that the latency functions are not known apriori, but can be only measured; ii) the fact that a distributed approach is needed since a centralized approach would require too much control traffic to exchange information among the potentially thousands of UEs. In this work, a distributed, non-cooperative and dynamic load balancing algorithm is consequently developed in the context of adversarial network equilibria; specifically, the algorithm considers every single packet included in a QoS-Flow as an *agent*, able to decide the AP it is assigned to. Such decisions are based on the measurements of the latency functions, obtained starting from the observation of the resource blocks allocated on the APs over which the commodity is routed to sustain the connection, and are made unilaterally in an adversarial framework, with no concern for the overall system performance.

The main motivations behind this work are then (i) to design a dynamic adversarial capacitated load balancing algorithm and to prove, using Lyapunov and Invariance Principle arguments, how the difference equation governing the global state of the system converges to an approximated Beckmann equilibrium, and (ii) to show the effectiveness of such an approach through its application to the multi-connectivity problem in a simulated 5G network scenario.

The work presented in this chapter was carried out within the H2020 5G-ALLSTAR project (www.5g-allstar.eu), aimed at the seamless, reliable, and ubiquitous provision of broadband services over heterogeneous 5G networks. However, it is possible to note that, since the algorithm is developed within the research framework of selfish routing, it can be applied to several problems and scenarios other than the one considered here.

2.1.1 Multi-Connectivity and Traffic Steering in 5G Networks

This work addresses the problem of traffic steering, i.e., of selecting which APs a QoS-Flow shall utilize to connect the UEs with the core network by modeling it as a load-balancing problem.

This vision is compliant with the latest developments of the 5G architecture (see Fig. 2.1), as designed by 5GPPP in [1]. Multi-connectivity comprises the concept of dynamic traffic steering, which envisages the ability to dynamically steer the traffic, partitioned into QoS-Flows among the various available APs of the RATs, based on feedback on the current AP performances. In this framework, QoS-Flows may be

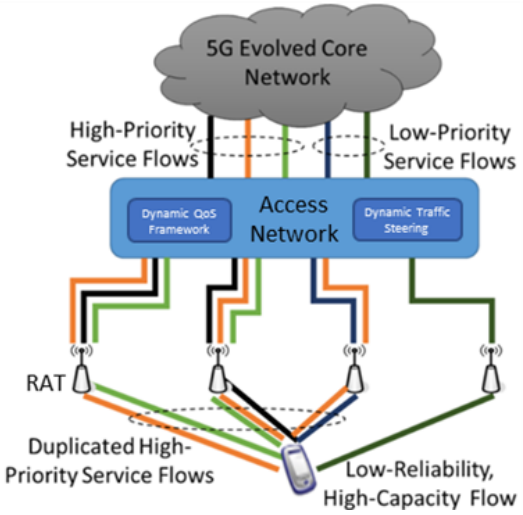


Figure 2.1. Dynamic Traffic Steering framework from [1]

duplicated over different APs to increase their resiliency, while other ones may be split over multiple RATs to increase their throughput or to better meet their QoS requirements.

Within the 5G architecture, the traffic steering problem is solved in three different ways: (i) with a User-Centric approach, where each UE decides its connection preferences according to local measures of some performance indicator; (ii) in a RAN-Assisted fashion, in which the decision is still made by the UEs but the RAN provides them with additional information on the network state; (iii) with a RAN-Controlled approach, where all decisions are made by the RAN, which is a centralized unit by nature, or delegated to the distributed control units that govern the single APs.

Several works study the problem of multi-connectivity in the heterogeneous network framework proposed by 5G, from both architectural [15], [16] and algorithmic [17–19] points of view. Multi-connectivity enables the problem of optimally steering the network traffic over the available APs, in such a way that the QoS requirements of the various QoS-Flows are met [20], [21]. The problem of access network selection has been studied utilizing several different approaches, spacing from fuzzy-logic control to multiple-attribute decision-making and combinatorial optimization [17]. Common solutions utilize the concept of utility and latency functions, as in this work, to capture the network performances [17, 22, 23]. Several works in the literature also employ game-theoretic approaches for the AP selection, typically in adversarial frameworks, as [17, 24, 25], leading the networks to Nash equilibrium states.

Regarding game-theoretic solutions, one possible modeling choice is to have an

adversarial game between the users, as in [24, 26] that envisage a setup similar to the one used in this work. In such scenarios, the users compete to attain the best connection quality while eventually also minimizing their costs. An alternative approach is to set up a game between the various network operators, each controlling a set of APs as in [23, 25], and focusing on their economic performances.

The algorithm proposed in this work utilizes *differential game theory*, a branch of game theory that studies dynamical systems, and shares some of the characteristics of the previously mentioned works, such as the adversarial nature of its equilibrium. The control algorithm designed in this work will be proven to drive the communication network state to a convenient equilibrium state, and this convergence will be attained by following an explicit discrete-time control law, with no need for round-games or price/cost bidding auctions. Contrary to optimization-based works, the proposed control law is also suitable to steer the traffic flows in real-time, and, being a distributed decision process, it does not require any significant control traffic overhead.

The previous aspects, together with the explicit inclusion of constraints on the available transmission capacity, make the proposed approach a suitable candidate for deployment in 5G scenarios implementing network slicing [18], in which the APs provide a limited quantity of resources to the QoS-Flows of a given service type or managed by third party tenants (e.g., video streaming, autonomous guidance, voice...). Concerning the mentioned Dynamic Traffic Steering framework [5], the algorithm can be implemented in the RAN-Assisted and in the RAN-Controlled configurations: in the former case, the algorithm would run in the UEs based on the information received by the RAN; in the latter case, the algorithm would run directly in the RAN and, in particular, for Non-StandAlone 5G systems (5G-NSA), in either the Centralized Unit (CU) or the Distributed Unit (DU) [19] of the next-generation-Node-Bs (gNodeBs or gNBs) [20] that govern the various APs.

2.1.2 Adversarial Load Balancing in 5G Networks and Beckmann Equilibria

The problem of optimally distributing the flow is one of the most fundamental and challenging aspects of any network operation. In the framework of selfish routing, the network flow is formed by a stream of infinitely-many decision-making agents [27] that compete for attaining the best performance, without consideration for the congestion, and consequent performance degradation, that their decisions cause to the other agents. Wardrop equilibria [28] were then introduced to describe a network state in which no single agent can unilaterally improve its performances (e.g., in terms of travel time, as in the original Wardrop formulation). Being an

adversarial kind of equilibria, the overall network performance is not optimized and the performance loss is referred to as the *price of anarchy* in the literature [29]. The concept of Wardrop equilibrium has been extended to various families of networks, among which the capacitated ones [13], [30–32], and problems, as the load balancing one [33–35]. Even if Wardrop equilibria can be computed by centralized algorithms in polynomial time [36], for the low connection latency promised by 5G – and the consequent agile and fast traffic steering requirements – distributed approaches are more suitable, motivating the development of a dynamic algorithm.

Based on a simple representation of the network dynamics in terms of difference equations derived from the flow conservations laws, this work proposes a load balancing solution over the nodes of a dynamical network that represents the 5G infrastructure [37, 38], consisting in the connections between several APs and their users with the core network. In doing so, the algorithm takes into account that the amount of traffic each AP can support is limited, or capacitated, due to transmission power constraints and, in general, resource scarcity as in network slicing scenarios. This limitation implies that the user equilibrium to which the network will converge may not be in principle the Wardrop equilibrium [32], which is defined for unconstrained networks. Several works [13], [30–32] extended the original formulation of the Wardrop user equilibrium, which corresponds to a situation in which all the latencies of each commodity are equalized, to deal with capacitated networks. The resulting equilibrium, known as Beckman user equilibrium, is such that the latencies of all the unsaturated APs of each commodity are equalized. Differently from [13], [30–32], this work proposes a dynamic algorithm that will be proven to converge to a Beckmann equilibrium.

Regarding dynamic load balancing solutions for Wardrop equilibria in the literature, several works utilize the concepts of learning and exploration to cope with the limited feedback information that the decision-making agents have access to. To attain a better knowledge of the system state and dynamics, the agents sample different flow distribution strategies and then exploit the learned system characteristics to converge to optimal states. The authors of [39] present an asynchronous and distributed algorithm that employs reinforcement learning to update transmission probabilities, based on an estimation of the network edges' latencies. In [40], an iterative and distributed learning solution is proven to converge to a Wardrop equilibrium state using Lyapunov arguments, as in this work.

An important contribution has been given by Fischer et al. in [41–43]. In [41] and [43], a round-based algorithm is developed to solve a game among the various commodities, aimed at redistributing the traffic flow and reaching an approximated Wardrop equilibrium. In [42], a similar setup is analyzed assuming that the infor-

mation available to the agents may be stale. In [44], a dynamic discrete-time load-balancing algorithm, later extended to the time-delayed case in [45], is presented in the context of Virtual Private Networks, which converges to an approximate Wardrop equilibrium.

The present work extends the results of previous works, starting from the algorithm in [44], mainly in two directions:

i) the convergence properties of the algorithm are studied in the multi-commodity case (a requirement for application in the 5G framework), which was not explicitly discussed in the cited works;

ii) the algorithm analysis and design are extended to the case of capacitated networks, not dealt with by the dynamic algorithms in the literature, enabling the application of the solution to more realistic case studies in several domains.

2.2 Proposed Wardrop Load Balancing Algorithm

Subsection 2.2.1 describes the basic definitions needed for the algorithm analysis; subsection 2.2.2 presents the load balancing algorithm and the convergence proof; subsection 2.2.3 models the 5G traffic steering problem as a load balancing one.

2.2.1 Preliminaries on Wardrop and Beckmann Equilibria and Lyapunov Stability

As anticipated in Section 2.1, this work further develops a well-known model for selfish routing [41], where an infinite population of agents carries an infinitesimal amount of load each and builds on the previous work [44] concerning distributed load balancing algorithms. The proposed control scheme relies on common assumptions on the latency functions. The considered network consists of a set of $\mathcal{P}$ providers, which serve a set of $\mathcal{I}$ of commodities. Each commodity $i \in \mathcal{I}$ is characterized by a fixed flow demand λ^i and is served by a subset of providers $\mathcal{P}^i \subset \mathcal{P}$. Each commodity i using provider p is characterized by a latency function l_p^i and each provider p is characterized by a capacity c_p.

Assumption 2.2.1. The latency function $l_p^i(\xi)$ are positive, non decreasing and Lipschitz continuous with constant β_p^i, for $\xi \in [0, c_p]$, where c_p is the capacity of provider p, for all $p \in \mathcal{P}$. Furthermore, the maximum Lipschitz constant of all the l_p^i's is denoted as $\bar{\beta} = \max_{p\in\mathcal{P}, i\in\mathcal{I}} \beta_p^i$.

The assumption is not restrictive in real use-cases since the provider performances decrease with their load and poses a very mild design constraint on the function classes choices.

In non-capacitated algorithms, if x_p^i indicates the amount of the flow of commodity i allocated on the provider p, the set of feasible states is defined as

$$\mathcal{X} = \left\{ \boldsymbol{x} = (x_p)_{p\in\mathcal{P}} \quad | \quad x_p = \sum_{i\in\mathcal{I}} x_p^i, x_p^i \geq 0, \forall p \in \mathcal{P}^i, \quad \sum_{p\in\mathcal{P}^i} x_p^i = \lambda^i, \forall i \in \mathcal{I} \right\}, \tag{2.1}$$

and a flow $\boldsymbol{x} \in \mathcal{X}$ is at a Wardrop equilibrium if, for each commodity $i \in \mathcal{I}$, the latencies of the loaded providers are equalized, i.e., if $l_p^i(x_p^i) \leq l_q^i(x_q^i)$ for all $p \in \mathcal{P}^i$ such that $x_p^i > 0$, for all $q \in \mathcal{P}^i$ and for all $i \in \mathcal{I}$.

By defining the Beckmann-McGuire-Winsten potential

$$\Phi(\boldsymbol{x}) = \sum_{i\in\mathcal{I}} \sum_{p\in\mathcal{P}^i} \int_0^{x_p^i} l_p^i(\xi) d\xi \tag{2.2}$$

the Wardrop equilibria are the solutions of the optimization problem

$$\min_{\boldsymbol{x}\in\mathcal{X}} \Phi(\boldsymbol{x}). \tag{2.3}$$

Capacity-constrained networks are characterized by the additional capacity constraints

$$x_p \leq c_p \quad \forall p \in \mathcal{P}. \tag{2.4}$$

A flow $\boldsymbol{x} \in \mathcal{X}$ is feasible if constraints (2.4) hold, and the set of feasible states is defined as

$$\mathcal{X}_{CP} = \left\{ \boldsymbol{x} \in \mathcal{X} \quad | \quad x_p \leq c_p, \forall p \in \mathcal{P} \right\}. \tag{2.5}$$

Considering a flow $\boldsymbol{x} \in \mathcal{X}_{CP}$, provider $p \in \mathcal{P}$ is defined as *capacity-constrained* or *saturated* if $x_p = c_p$.

A flow $\boldsymbol{x} \in \mathcal{X}_{CP}$ is at Beckmann user equilibrium if, for each commodity, the latencies of the loaded and unconstrained providers are equalized, i.e., mode precisely:

Definition 2.2.1. [13] A flow $\boldsymbol{x} \in \mathcal{X}_{CP}$ is at Beckmann user equilibrium if $l_p^i(x_p^i) \leq l_q^i(x_q^i)$ for all $p \in \mathcal{P}^i$ such that $x_p^i > 0$, for all $q \in \mathcal{P}^i$ such that $x_q < c_q$ and for all $i \in \mathcal{I}$.

The set of equilibria is then

$$\mathcal{X}_{eq} = \Big\{ \boldsymbol{x} \in \mathcal{X}_{CP} \quad | \quad l_p^i(\boldsymbol{x}) \leq l_q^i(\boldsymbol{x}), \forall p \in \mathcal{P}^i, \text{ s.t. } x_p^i > 0, \forall q \in \mathcal{P}^i, \text{ s.t. } x_q < c_q, \forall i \in \mathcal{I} \Big\}. \tag{2.6}$$

The minimization problem (2.3) with constraints (2.4) will hereinafter referred to as Capacity-constrained Problem (CP). The Beckmann user equilibria [31] are the optimal solutions of the CP.

Property 2.2.1. [13] If the set of feasible solutions $\mathcal{X}_{CP}$ of the CP is nonempty, the optimization problem consists in minimizing a convex function over a nonempty polytope and, thus, the set of optimal flows $\mathcal{X}_{eq}$ is nonempty and convex.

The algorithm convergence proof of section 2.2.2 relies on LaSalle invariance principle for discrete-time nonlinear systems [46] [47].

Definition 2.2.2. $\mathcal{L} : \mathcal{X} \to \mathbf{R}$ is a candidate Lyapunov function for a discrete-time nonlinear system $\boldsymbol{x}[k+1] = f(\boldsymbol{x}[k])$ if:

1. $\mathcal{L} \in \mathcal{C}^1$ and is bounded from below;
2. If $\boldsymbol{x}_{eq} \in \mathcal{X}_{eq}$, where $\mathcal{X}_{eq}$ is the set of equilibrium points, $\mathcal{L}(\boldsymbol{x}_{eq}) = 0$ and $\mathcal{L}(\boldsymbol{x}) > 0$ if $\boldsymbol{x} \notin \mathcal{X}_{eq}$;
3. Along forward trajectories, $\mathcal{L}$ satisfies $\Delta\mathcal{L}(\boldsymbol{x}[k]) := \mathcal{L}\big(f(\boldsymbol{x}[k])\big) - \mathcal{L}(\boldsymbol{x}[k]) \leq 0, k = 0, 1, 2, ...$

Theorem 2.2.1. [46] Let $\mathcal{L}(\boldsymbol{x})$ be a candidate Lyapunov function for the discrete-time nonlinear system $\boldsymbol{x}[k+1] = f(\boldsymbol{x}[k])$. Then, any bounded trajectory tends to the largest invariant subset $\mathcal{M}$ contained in the set of points defined by $\Delta\mathcal{L}(\boldsymbol{x}) = 0$.

2.2.2 Capacitated Load Balancing Algorithm and Convergence Proof

For each commodity $i \in \mathcal{I}$, the control action consists in the decision, at time k, of *migrating* part of the flow mapped onto a given provider p to another provider q, with $p, q \in \mathcal{P}^i$. By denoting the rate of such migration with $r_{pq}^i[k]$, the system dynamics is written as

$$\boldsymbol{x}[k+1] = f(\boldsymbol{x}[k]), k = 0, 1, 2, ... \tag{2.7}$$

with

$$x_p[k] = \sum_{i \in \mathcal{I}} x_p^i[k], \tag{2.8}$$

$$x_p^i[k+1] = x_p^i[k] + \tau \sum_{q \in \mathcal{P}^i} \big(r_{qp}^i[k] - r_{pq}^i[k]\big), \tag{2.9}$$

and with feasible initial conditions

$$\boldsymbol{x}[0] \in \mathcal{X}_{CP} \tag{2.10}$$

forall $p, q \in \mathcal{P}^i$ and $i \in \mathcal{I}$.

The proposed controller builds on the dynamic algorithm in [44], which expresses the migration rate as

$$r^i_{pq}[k] = x^i_p[k]\sigma^i\mu^i_{pq}[k], \tag{2.11}$$

where σ^i is a positive migration gain and $\mu^i_{pq}[k]$ is the migration policy, representing the decision of whether (if it is positive) or not (if it is equal to zero) migrate some flow from provider p to provider q.

As in [44] for the Wardrop equilibria, approximated Beckmann user equilibria are defined.

Definition 2.2.3. The set of ε-Beckmann user equilibria is defined as

$$\begin{aligned}\mathcal{X}^{\varepsilon}_{eq} = \Big\{ \boldsymbol{x} \in \mathcal{X}_{CP} \quad | \quad & l^i_p(x^i_p) \leq l^i_q(x^i_q) + \varepsilon, \forall p \in \mathcal{P}^i \text{ s.t. } x^i_p > 0, \\ & \forall q \in \mathcal{P}^i \text{ s.t. } x_q \leq c_q - \frac{\varepsilon}{2\overline{\beta}}, \forall i \in \mathcal{I} \Big\}\end{aligned} \tag{2.12}$$

where $\varepsilon \geq 0$ represents a maximum tolerated latency mismatch.

Remark 2.2.1. The defined sets are such that $\mathcal{X}^{\varepsilon}_{eq} \to_{\varepsilon \to 0} \mathcal{X}_{eq}$ and $\mathcal{X}_{eq} \subseteq \mathcal{X}^{\varepsilon}_{eq} \subseteq \mathcal{X}_{CP}$: the objective of the controller is then, starting from a physically admissible state in $\mathcal{X}_{CP}$, to reach an approximated equilibrium state in $\mathcal{X}^{\varepsilon}_{eq}$, whose degree of approximation with respect to the equilibrium state in $\mathcal{X}_{eq}$ reduces with ε.

The tolerance ε is introduced since the kind of migration rates of equation 2.11 cannot guarantee convergence in the discrete-time case, however small the sampling period [42]. A flow $\boldsymbol{x} \in \mathcal{X}_{CP}$ is then at ε-Beckmann equilibrium if, for each commodity i, the latencies of the loaded and ε-unconstrained providers are equalized, where a provider $p \in \mathcal{P}^i$ is defined to be ε-unconstrained if $x_p < c_p - \frac{\varepsilon}{2\overline{\beta}}$.

In the proposed algorithm, the migration decision is defined as

$$\mu^i_{pq}[k] = \begin{cases} 0, \text{ if } l^i_p(x^i_p[k]) - l^i_q(x^i_q[k]) \leq \varepsilon \text{ or if } x_q[k] \geq c_q - \frac{\varepsilon}{2\overline{\beta}} \\ 1, \text{ otherwise} \end{cases} . \tag{2.13}$$

The controller system dynamics, hereafter denoted as load balancing dynamics,

is then expressed by equations (2.9), (2.11) and (2.13), with control gain set as

$$\sigma^i = \frac{\varepsilon}{2\tau\bar{\beta}\lambda^i(|\mathcal{P}^i|-1)|\mathcal{I}|}, \tag{2.14}$$

and with the tolerance set as

$$0 < \varepsilon \leq \min_{i\in\mathcal{I}} \bar{\beta}\lambda^i|\mathcal{I}|. \tag{2.15}$$

Remark 2.2.2. The approximated capacity-constrained user equilibria are such that, for each commodity, the latencies of the loaded and ε-unconstrained providers are equalized within the tolerance ε. Then, for a given equilibrium flow $\boldsymbol{x} \in \mathcal{X}^{\varepsilon}_{eq}$ and for each commodity $i \in \mathcal{I}$, three classes of providers exist: the unloaded providers $p \in \mathcal{P}^i$ such that $x^i_p = 0$; the ε-constrained providers $p \in \mathcal{P}^i$ such that $x_p > c_p - \frac{\varepsilon}{2\bar{\beta}}$; the ε-unconstrained providers, whose latencies are equalized.

The convergence property of the algorithm relies on the following 3 lemmas:

Lemma 2.2.2. Under Assumption 2.2.1, considering the LB dynamics, the latency variation of a provider $p \in \mathcal{P}^i$ in one time-step is bounded by

$$\left| l^i_p(x^i_p[k+1]) - l^i_p(x^i_p[k]) \right| \leq \frac{\varepsilon}{2|\mathcal{I}|} \tag{2.16}$$

Proof. Considering the generic commodity $i \in \mathcal{I}$, provider $p \in \mathcal{P}^i$ and time k, the maximum latency decrease occurs when no commodities migrate their population from the other providers to provider p:

$$\begin{aligned} &l^i_p(x^i_p[k+1]) = \\ &= l^i_p\Big(x^i_p[k] + \tau \sum_{q\in\mathcal{P}^i} (r^i_{qp}[k] - r^i_{pq}[k])\Big) \leq \\ &\leq l^i_p\Big(x^i_p[k] - \tau \sum_{q\in\mathcal{P}^i} r^i_{pq}[k]\Big) \end{aligned} \tag{2.17}$$

Since β^i_p is Lipschitz constant of the function $l^i_p(\cdot)$ between 0 and c_p, it follows that

$$l^i_p(x^i_p[k+1]) \geq l^i_p(x^i_p[k]) - \tau\beta^i_p \sum_{q\in\mathcal{P}^i} r^i_{pq}[k]. \tag{2.18}$$

Considering equations (2.11) and (2.14), the last term of equation (2.18) is written as

$$
\begin{aligned}
&\tau\beta_p^i \sum_{q\in\mathcal{P}^i} r_{pq}^i[k] = \\
&= \tau\beta_p^i \sum_{q\in\mathcal{P}^i} x_p^i[k]\sigma^i\mu_{pq}^i[k] = \\
&= \tau\beta_p^i x_p^i[k]\sigma^i \sum_{q\in\mathcal{P}^i} \mu_{pq}^i[k] = \\
&= \tau\beta_p^i x_p^i[k] \frac{\varepsilon}{2\tau\bar{\beta}\lambda^i(|\mathcal{P}^i|-1)|\mathcal{I}|} \sum_{q\in\mathcal{P}^i} \mu_{pq}^i[k] \leq \\
&\leq \frac{\varepsilon}{2|\mathcal{I}|}
\end{aligned}
\tag{2.19}
$$

where the inequality holds since $x_p^i[k] \leq \lambda^i$, $\beta_p^i \leq \bar{\beta}$ and since, recalling equation (2.13), there are at most $(|\mathcal{P}|-1)$ terms equal to 1 in $\sum_{q\in\mathcal{P}} \mu_{pq}^j[k]$. It follows that

$$
l_p^i(x_p^i[k+1]) \geq l_p^i(x_p^i[k]) - \frac{\varepsilon}{2|\mathcal{I}|}. \tag{2.20}
$$

Similarly, the maximum latency increase occurs when no commodities migrate their populations from provider p to other providers:

$$
l_p^i(x_p^i[k+1]) \leq l_p^i(x_p^i[k]) + \tau \sum_{q\in\mathcal{P}} r_{qp}^i[k], \tag{2.21}
$$

which yields

$$
l_p^i(x_p^i[k+1]) \leq l_p^i(x_p^i[k]) + \frac{\varepsilon}{2|\mathcal{I}|}. \tag{2.22}
$$

□

Lemma 2.2.3. $\mathcal{X}_{CP}$ is a positively invariant set for the LB dynamics.

Proof. The aim of this proof is to show, for all $k \geq 0$, for all $p \in \mathcal{P}^i$ and for all $i \in \mathcal{I}$:

i) $\sum_{p\in\mathcal{P}^i} x_p^i[k] = \lambda^i$,

ii) $x_p^i[k] \geq 0$,

iii) $x_p[k] \leq c_p$.

Here below are reported the proofs for the three statements:

i) Considering that $x[0] \in \mathcal{X}_{CP}$, equations (2.9), (2.11) and (2.8) yield that the population remains constant, since

$$\begin{aligned} & x_p^i[k+1] - x_p^i[k] = \\ & = \sum_{p\in\mathcal{P}^i} \sum_{q\in\mathcal{P}^i} (r_{qp}^i[k] - r_{pq}^i[k]) = \\ & = \sum_{p\in\mathcal{P}^i} \sum_{q\in\mathcal{P}^i} r_{qp}^i[k] - \sum_{q\in\mathcal{P}^i} \sum_{p\in\mathcal{P}^i} r_{qp}^i[k] = 0, \end{aligned} \tag{2.23}$$

and thus that $\sum_{p\in\mathcal{P}^i} x_p^i[k] = \sum_{p\in\mathcal{P}^i} x_p^i[0] = \lambda^i$, $\forall k \geq 0$.

ii) Given that $x_p^i[0] \geq 0$, it is proven below by induction that $x_p^i[k] \geq 0$, $\forall k \geq 0$. Assuming that $x_p^i[k] \geq 0$, for a given k, it is sufficient to prove that

$$x_p^i[k+1] = x_p^i[k] + \tau \sum_{q\in\mathcal{P}^i} (r_{qp}^i[k] - r_{pq}^i[k]) \geq 0, \quad \forall p \in \mathcal{P}^i. \tag{2.24}$$

If $x_p^i[k] = 0$, it follows that $r_{pq}^i[k] = 0$ and thus the equation (2.24) yields $x_p^i[k+1] \geq 0$.

If $x_p^i[k] > 0$, from equation (2.11) it follows that $r_{pq}^i[k] \geq 0$. Thus, the following inequality holds (in the worst case, no providers migrate part of their population to a provider p):

$$x_p^i[k+1] \geq x_p^i[k] - \tau \sum_{q\in\mathcal{P}^i} r_{pq}^i[k]. \tag{2.25}$$

A sufficient condition for inequality (2.24) to hold is then

$$x_p^i[k] - \tau \sum_{q\in\mathcal{P}^i} r_{pq}^i[k] \geq 0. \tag{2.26}$$

Recalling equations (2.11) and (2.13), eq. (2.26) is written as

$$\begin{aligned} & x_p^i[k] - \tau \sum_{q\in\mathcal{P}^i} r_{pq}^i[k] = \\ & = x_p^i[k] - \tau \sum_{q\in\mathcal{P}^i} x_p^i[k]\sigma^i\mu_{pq}^i[k] = \\ & = x_p^i[k]\Big(1 - \tau\sigma^i \sum_{q\in\mathcal{P}^i} \mu_{pq}^i[k]\Big) \geq \\ & \geq x_p^i[k]\Big(1 - \tau\sigma^i(|\mathcal{P}^i| - 1)\Big), \end{aligned} \tag{2.27}$$

where the inequality holds since the summation has at most $(|\mathcal{P}^i| - 1)$ terms equal to 1. In the case $x_p^i[k] > 0$, equations (2.14) and (2.15) are sufficient for equation (2.27) to be non-negative;

iii) Given that $x_p[0] \leq c_p$, it is proven below by induction that $x_p[k] \leq c_p$, $\forall k \geq 0$. Assuming that $x_p[k] \leq c_p$, for a given k, it is sufficient to prove that

$$x_p[k+1] = x_p[k] + \tau \sum_{i \in \mathcal{I}} \sum_{q \in \mathcal{P}^i} (r_{qp}^i[k] - r_{pq}^i[k]) \leq c_p, \quad \forall p \in \mathcal{P}^i. \tag{2.28}$$

If $x_p[k] \geq c_p - \frac{\varepsilon}{2\bar{\beta}}$ equation (2.13) entails that $r_{qp}^i[k] = 0$ for all $q \in \mathcal{P}^i$ and $i \in \mathcal{I}$ and, thus, from equation (2.9), that $x_p[k+1] \leq x_p[k]$.

Otherwise, if $x_p[k] < c_p - \frac{\varepsilon}{2\bar{\beta}}$, it is possible to consider that

$$\begin{aligned} & x_p[k+1] \leq x_p[k] + \tau \sum_{i \in \mathcal{I}} \sum_{q \in \mathcal{P}^i} r_{qp}^i[k] = \\ & = x_p[k] + \tau \sum_{i \in \mathcal{I}} x^i[k] \sigma^i \sum_{q \in \mathcal{P}^i} \mu_{qp}^i[k] \leq \\ & \leq x_p[k] + \sum_{i \in \mathcal{I}} \frac{\varepsilon}{2\bar{\beta}|\mathcal{I}|} = x_p[k] + \frac{\varepsilon}{2\bar{\beta}} < c_p \end{aligned} \tag{2.29}$$

□

Lemma 2.2.4. $\mathcal{L}(\boldsymbol{x}) := \Phi(\boldsymbol{x}) - \Phi_{min}$, where Φ_{min} is the minimum value of $\Phi(\boldsymbol{x})$ for all the minimizers of the CP, is a candidate Lyapunov function for the LB dynamics

Proof. For the definition of Φ_{min}, the function $\mathcal{L}(\boldsymbol{x})$ is positive definite in $\mathcal{X}_{CP}$.

Let $\Delta\mathcal{L}(\boldsymbol{x}[k])$ denote the difference of the Lyapunov function $\mathcal{L}(\boldsymbol{x})$ along the solutions of the controlled system:

$$\begin{aligned} & \Delta\mathcal{L}(\boldsymbol{x}[k]) = \mathcal{L}(\boldsymbol{x}[k+1]) - \mathcal{L}(\boldsymbol{x}[k]) = \\ & = \sum_{p \in \mathcal{P}} \int_{x_p[k]}^{x_p[k+1]} l_p(\xi) d\xi \leq \\ & \leq \sum_{p \in \mathcal{P}} \Big(x_p[k+1] - x_p[k] \Big) l_p(x_p[k+1]) = \\ & = \tau \sum_{p \in \mathcal{P}} \sum_{i \in \mathcal{I}} \Big(\sum_{q \in \mathcal{P}} r_{qp}^i[k] - \sum_{q \in \mathcal{P}} r_{pq}^i[k] \Big) l_p(x_p[k+1]) = \\ & = \tau \sum_{p \in \mathcal{P}} \sum_{i \in \mathcal{I}} \sum_{q \in \mathcal{P}} r_{pq}^i[k] \Big(l_q(x_q[k+1]) - l_p(x_p[k+1]) \Big), \end{aligned} \tag{2.30}$$

where the inequality holds from geometric considerations: if $x_p[k+1] > x_p[k]$, recalling that the l_p's are nondecreasing functions, the definite integral $\int_{x_p[k]}^{x_p[k+1]} l_p(\xi)d\xi$ is smaller that the quantity $(x_p[k+1] - x_p[k])l_p(x_p[k+1])$; conversely, if $x_p[k+1] < x_p[k]$, the integral $\int_{x_p[k]}^{x_p[k+1]} l_p(\xi)d\xi$ is larger than the quantity $(x_p[k+1] - x_p[k])l_p(x_p[k+1])$.

Analysing each term of the inner summation, two cases hold: if $r_{pq}^i(t) = 0$ the term is null, otherwise, if $r_{pq}^i(t) > 0$, the term is negative. In fact, it is shown below that, if $r_{pq}^i[k] > 0$, it holds that $l_p(x_p[k+1]) - l_q(x_q[k+1]) > 0$.

Lemma 2.2.2 states that

$$\begin{aligned} & l_p(x_p[k+1]) - l_q(x_q[k+1]) \geq \\ & \geq \left(l_p(x_p[k]) - \frac{\varepsilon}{2}\right) - \left(l_q(x_q[k]) + \frac{\varepsilon}{2}\right) = \\ & = l_p(x_p[k]) - l_q(x_q[k]) - \varepsilon > 0, \end{aligned} \tag{2.31}$$

where the inequality holds since a necessary condition for $r_{pq}^i[k] > 0$ is that $l_p(x_p[k]) - l_q(x_q[k]) > \varepsilon$ (see equation (2.13)). □

Finally, the following theorem proves the convergence towards an approximated Beckmann user equilibrium.

Theorem 2.2.5. The trajectories of the LB dynamics asymptotically tend to the set of equilibria $\mathcal{X}_{eq}^{\varepsilon}$.

Proof. Given that Lemma 2.2.4 states that $\mathcal{L}(\boldsymbol{x})$ is a candidate Lyapunov function for the LB dynamics, the proof relies on the LaSalle invariance principle of Theorem 2.2.1, i.e., on showing that $\mathcal{X}_{eq}^{\varepsilon}$ is the maximum invariant set where $\Delta\mathcal{L} = 0$.

Let $\boldsymbol{x} \in \mathcal{X}_{eq}^{\varepsilon}$ and $\boldsymbol{x}[0] = \boldsymbol{x}$. By comparing definition (2.6) and equation (2.13), it holds that $r_{pq}^i[k] = 0$ for all $p, q \in \mathcal{P}^i$ and $i \in \mathcal{I}$, which entails i) that $\boldsymbol{x}[k] = \boldsymbol{x}[0] = \boldsymbol{x}_{eq} \in \mathcal{X}_{eq}^{\varepsilon}$ for all $k > 0$, i.e., that $\mathcal{X}_{eq}^{\varepsilon}$ is a positively invariant set, and ii) that $\Delta\mathcal{L}(\boldsymbol{x}[k]) = 0$ in $\mathcal{X}_{eq}$ (see equation (2.30)).

To show that $\mathcal{X}_{eq}^{\varepsilon}$ is the maximum set where $\Delta\mathcal{L}(\boldsymbol{x}[k]) = 0$, it is proven below that $\Delta\mathcal{L}(\boldsymbol{x}[k]) < 0$ if $\boldsymbol{x}[k] = \boldsymbol{x}$ with $\boldsymbol{x} \notin \mathcal{X}_{eq}$. In fact by definition (2.12), in this case there exist at least one pair of providers $p, q \in \mathcal{P}^i$ and a commodity $i \in \mathcal{I}$ such that $l_p(x_p[k]) - l_q(x_q[k]) > \varepsilon$, with $x_p^i[k] > 0$ and $x_q[k] < c_p - \frac{\varepsilon}{2\beta}$, which, in turn, yields $r_{pq}^i[k] > 0$ (see equations (2.11), (2.14) and (2.13)). Having established that $r_{pq}^i[k] > 0$ with $l_p(x_p[k]) - l_q(x_q[k]) > \varepsilon$, it follows that the corresponding term of the inner summation of equation (2.30) is negative, which is a sufficient condition for $\Delta\mathcal{L}(\boldsymbol{x}[k]) < 0$ (recalling that, in the proof of Lemma 2.2.4, it is shown that the terms of equation (2.30) are non-positive). □

2.2.3 5G Traffic Steering as a dynamic load-balancing problem

In the dynamic multi-connectivity framework of 5G networks [1], each UE selects the serving APs for its QoS-Flows. The network resources (capacity) are hence provided by the APs, and their efficient usage guides the design of traffic steering controllers. As introduced, in 5G systems, the dynamic management of such resources becomes of crucial importance in network slicing scenarios [48].

To model a multi-connectivity scenario in a network slicing environment as a dynamical network of the form (2.7)-(2.10) the AP p is regarded as a provider in the set of providers $\mathcal{P}$, the QoS-Flow associated with a UE as a commodity i in the set of commodities $\mathcal{I}$ and the amount of bitrate of the commodity i that is provided by the AP p at time k is associated to the state variable $x_p^i[k]$. The bitrate demand of the commodity i is then λ^i, which can be assumed, for a limited time window, to be constant.

In the following, a network slicing scenario in which the network operator dedicated a certain amount of bitrate c_p on each AP p to the control slice is considered.

Regarding latency functions, a natural choice is associating a different latency function l_p^i to the radio connection between the UE of commodity i and the AP $p \in \mathcal{P}^i$. This choice allows to capture quantities related to the specific connection performance, such as the resource blocks [3] usage, the power consumption of the single commodity i or its QoS degradation, but in turn implies that each commodity i is subject to a different latency from provider p, that may even depend only on x_p^i. It is worth mentioning that, in this kind of scenarios, in general, the network admits various equilibria characterized by different costs (latencies) [49]. Nevertheless, in the proposed framework depicted in Fig. 2.2, the considered network is characterized by parallel areas [49], implying that its equilibrium cost Φ_{min} is unique. In fact, with simple manipulations, the scenario of Fig. 2.2 can be shown to be equivalent to a network in which the latencies are associated with the depicted radio links, each of which can only be used by a single commodity. The scenario is then equivalent to a standard adversarial routing scenario with a unique equilibrium cost.

Regarding the mapping of the proposed control law onto the standard 5G architecture, Access Traffic Steering, Switching and Splitting (ATSSS) [50] decision rules for multi-connectivity are typically produced by a software module of the 5G core network, the Policy Control Function (PCF). The PCF configures the UEs and the User Plane Function (UPF), an entity directly connected to the gNodeBs of the RAN, to handle traffic steering based on local measurements, respectively for the uplink and the downlink phase. Such ATSSS rules may define the set of APs $\mathcal{P}^i$ available to the user i, depending on its contract with the provider, their priority, and in general, may define a control law to guide the steering of the QoS flows that

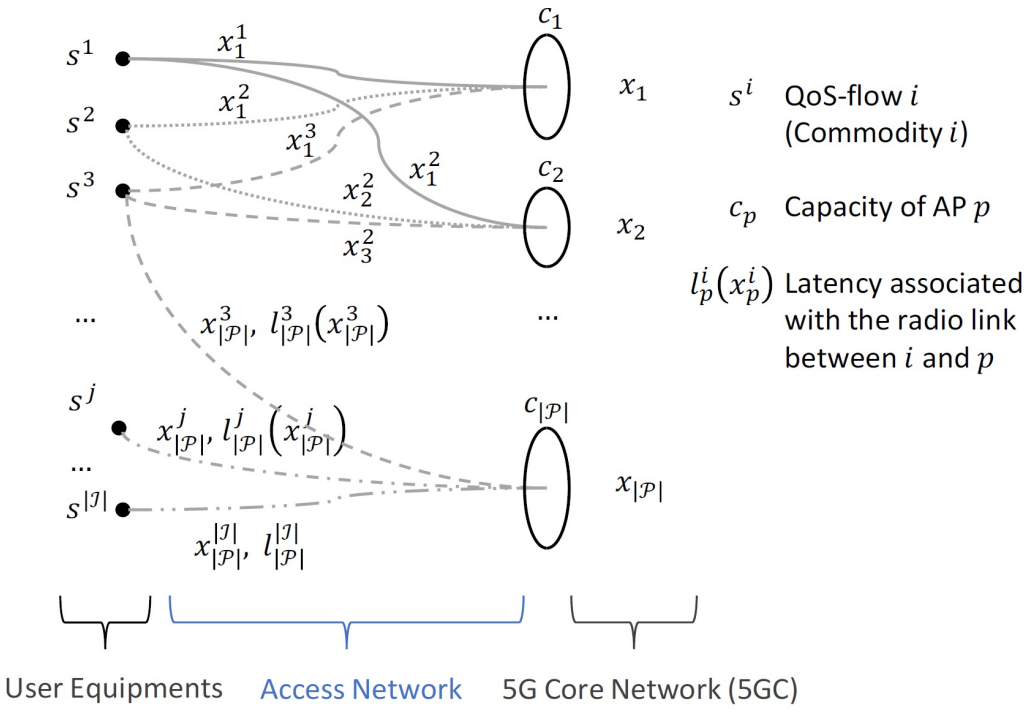


Figure 2.2. Load balancing graph. Solid line represents the QoS-flow of commodity 1, dotted line of commodity 2, dashed line of commodity 3 etc.

constitute the considered Protocol Data Unit (PDU) session. The dynamic traffic steering functionalities [1] are taken at RAN level, as depicted in Fig. 2.1 and so the proposed algorithm is designed to be deployed either in the DU of the gNodeBs that constitute the controlled RAN or in the UEs. The rules provided by the PCF can be included in the control logic by properly weighting or forbidding the various AP connections.

2.3 Numerical Simulation

This section reports the simulation setup and the results respectively in subsections 2.3.1 and 2.3.2.

2.3.1 Simulation Setup

For the validation of the proposed algorithm, in the scope of the 5G-ALLSTAR project, an open-source network simulator available in [51] has been developed, able to model different AP technologies, connection protocols and interference models in a multi-connectivity scenario. The network in Fig. 2.3 has been considered, consisting of a 4x4 Km area covered by a macro cell (provider BS1), a satellite (provider BS0), and six micro cells (BS2-BS7).

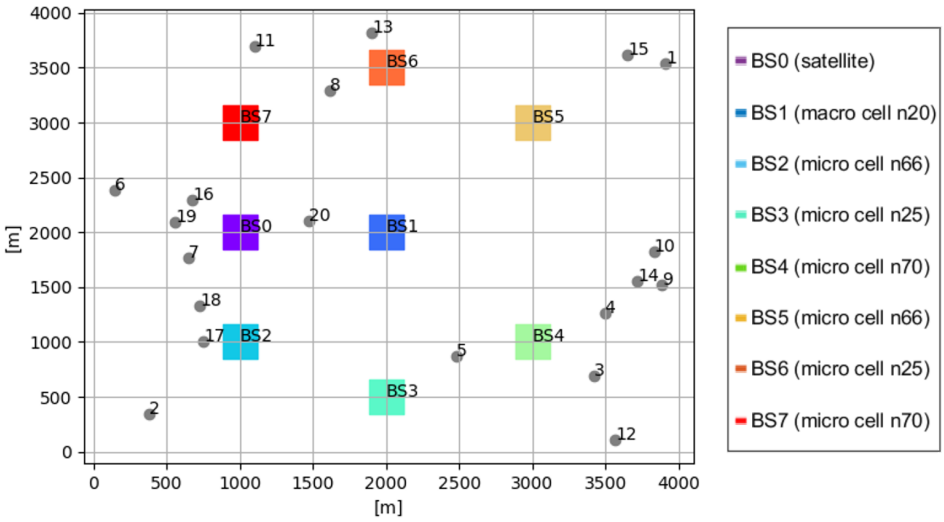


Figure 2.3. Network scenario

Operating band N°	Carrier frequency (GHz)	Bandwidth (MHz)
n20	0.8	20
n25	1.9	40
n66	1.7	40
n70	2	25

Table 2.1. Characteristic of micro and macro cells

A total of 20 UEs/commodities (grey dots in Fig. 2.3) were randomly distributed in the area, each requiring a constant load $\lambda^i = 50$ Mbps. The implemented interference model is taken from [52] and the frequency characteristics of the terrestrial APs are summarized in table 2.1 [53], [54].

Regarding the satellite AP, a Time Division Multiplexing (TDM) as in the example 6.6.2 of [55] is considered. The satellite parameters are adapted to have at least 1 bit per symbol with typical Signal to Noise Ratio (SNR) values [56], [57]. According to the TDM frame structure used, it is possible to allocate only blocks of 64 symbols ($1\mu s$). Moreover, each allocation must consider a header made of 288 symbols and a spacing between allocations of 64 symbols. Additional implementation details and updates can be found in [51].

It is considered as latency functions l_p^i the number of resource blocks utilized by the commodity i on the access point p. This particular choice will drive the network towards a state in which each connection equalizes the resource block usage over its available unsaturated APs $p \in \mathcal{P}^i$.

Assuming a stationary UE i (i.e., with constant path loss with all the access points p) and no interference, the amount of bitrate provided by a resource block on a given access point p is fixed. This implies that, in ideal conditions, l_p^i is almost linear, with a slope that depends on the utilized frequency bands, in line with Assumption 2.2.1. Note that several different choices could be made for the latency function, spacing from quantities that capture connection reliability, to transmission delay and user satisfaction, as the only requirements that such functions must satisfy are represented by Assumption 2.2.1, which open the possibility of considering a large family of functions (e.g., including polynomial or exponential ones). It is mentioned that in non-ideal cases, the value of $\bar{\beta}$ in (2.14) may need to be estimated by observing the growth of the selected latency functions during different network operative conditions. To allow a fair comparison with the terrestrial AP resource blocks, the assumptions made for the satellite imply that its latency function is equal to the number of its allocated symbols divided by 64. Additionally, each AP was associated with a multiplicative scaling factor for their latency functions to model different operating costs. In particular, the satellite was given the highest factor (0.5), the macro cell was given a medium value (0.2) and the lowest weight was associated with micro cells (0.1). Regarding the capacitated nature of the considered network, is assumed that the network operator dedicated to the controlled slice 200 Mbps on all micro cells, save for BS4, which was capacitated at 55 Mbps.

Concerning the parameters of the controller, the choice of latency functions leads to an experimentally determined value $\bar{\beta} = 2.44$ (found by try and error), the latency tolerance is selected as $\varepsilon = 0.5$ and the sampling time as $\tau = 10^{-3}$s. The

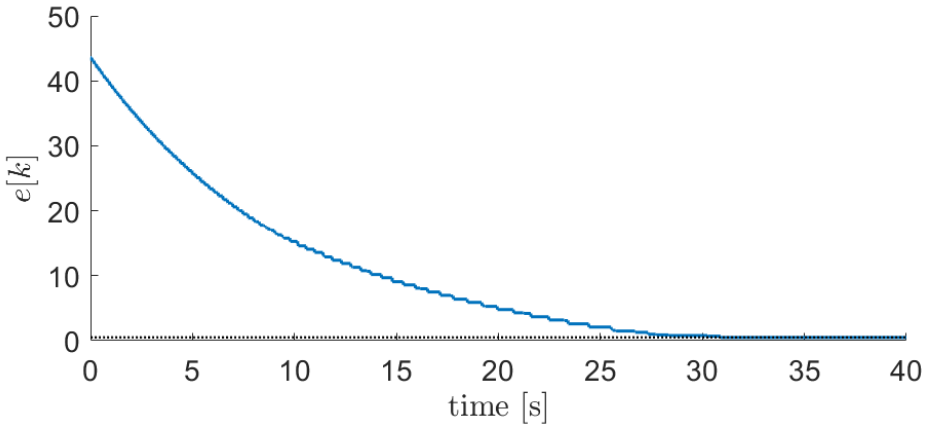


Figure 2.4. Maximum latency mismatch during the simulation (dotted line: tolerance ε)

resulting values for σ^i are in the range $[0.02, 0.05]$.

2.3.2 Simulation Results

Simulation runs were initialized by distributing uniformly the load of the commodities over $|\mathcal{P}^i| - 1$ of their available APs, selected randomly.

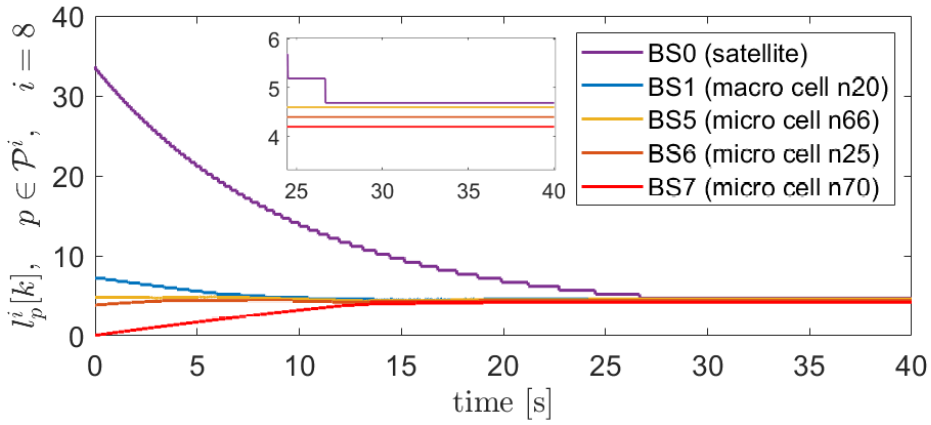
The reported simulations showed a convergence time to an ε-Beckmann user equilibrium in the order of 30 seconds, averaged over 25 runs. It is worth remarking that such convergence time is not related to the 5G QoS requirements, as it is assumed that the various access points can provide the proper QoS level (e.g., connection latency, average Bit Error Rate (BER), reliability level,...) if their capacities are not violated.

Fig. 2.4 shows, for an example run, how the maximum latency mismatch over all the commodities, defined as

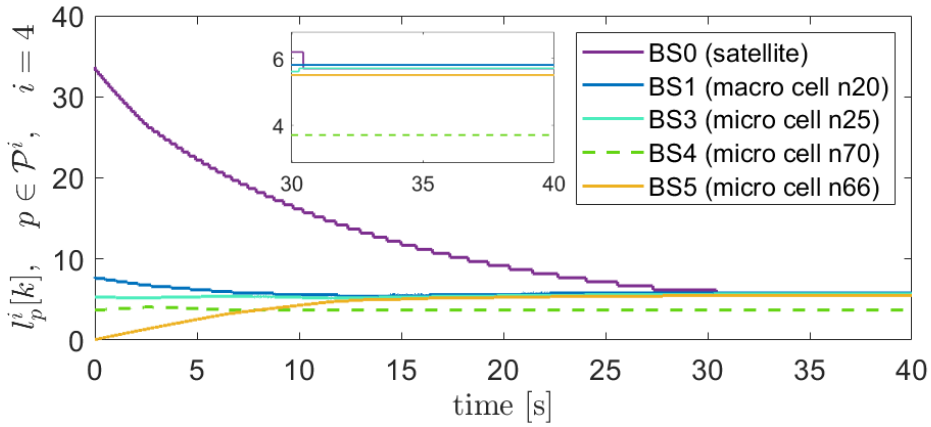
$$e[k] = \max_{i \in \mathcal{I}} \left\{ \max_{p \in \mathcal{P}^i | x_p^i[k] > 0} l_p(x_p[k]) - \min_{q \in \mathcal{P}^i | x_q[k] < c_q - \frac{\varepsilon}{2\beta}} l_q(x_q[k]) \right\},$$

decreases with time and, even if the initial conditions are quite unbalanced, with $e[k] > 40$, after 30 seconds $e[k]$ is already below the threshold ε.

For the example run, Fig. 2.5 reports the evolution of the latencies that characterize the commodities 4 and 8, for all of their available APs. The upper plot shows that the latencies of the APs available to QoS-Flow 8 converge to a common value, as expected, within the threshold ε; in particular it is possible to notice how the commodity rapidly starts using the (initially unused) micro cell BS7 and rapidly discharges the satellite. The lower plot of Fig. 2.5 shows the latencies of the QoS-Flow 4 and highlights that the latency of micro-cell BS4 does not converge to the latencies of the other used APs: the reason is that the AP becomes ε-saturated after



(a)



(b)

Figure 2.5. Commodity latency examples during the simulation (solid lines: unconstrained providers used by the commodity; dashed line: constrained providers; zoomed sub-plots to show the convergence within the tolerance $\varepsilon = 0.5$

about 3 seconds (see Fig. 2.6) – thus, by definition, the population of QoS-Flow 4 still converges to an ε-Beckmann equilibrium. Note that the latency associated with BS4 starts higher than its final value, as the commodity migrates towards BS5, but remains the lowest latency for the commodity 4 from 10 seconds onwards, as the other QoS flows already ε-saturated BS4 (i.e., no bitrate can be migrated to it).

Finally, Fig. 2.6 shows the population dynamics over the APs, highlighting how the macro cell is the most utilized AP, while all the micro cells allocate a similar amount of bitrate. The satellite, whose latency was the most penalized as it is the most costly connection technology, is rapidly discharged.

For the sake of comparison, in Fig. 2.7 the proposed controller has been benchmarked against two classic examples of load balancing solutions in heterogeneous networks. The figure reports the latency functions values experienced by the com-

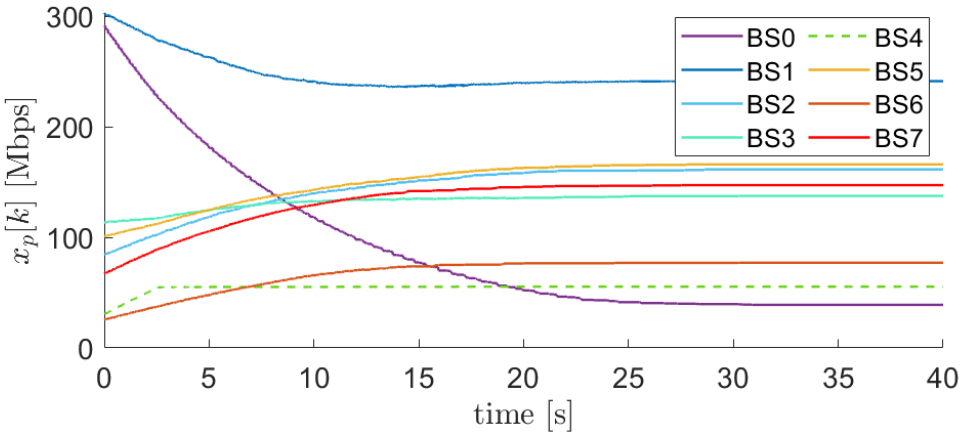


Figure 2.6. Network state in terms of total bitrate allocated on the various APs (solid lines: unconstrained providers, dashed line: constrained provider)

modity $i = 20$ over eight APs in the set $\mathcal{P}^2 0$. The choice of commodity 20 was related to the fact that it is one of the closest to the center of the considered area, as depicted in 2.3, and can hence be served by any AP. The first benchmarking algorithm ("uniform") uniformly distributes the bitrate demand over the various available and unsaturated APs. Due to the capacitated nature of the network, is assumed that this controller distributes the load of the commodities according to their index i, so that the traffic of commodity $i = 1$ is the first one to be allocated while the one of commodity $i = 20$ is the last.

The second algorithm ("weighted") distributes the bitrate considering the scaling factors associated with the latencies of the APs (0.5 for the satellite, 0.2 for the macro cell, and 0.1 for micro cells) so that for every unit of traffic allocated on the satellite 5 are allocated on the micro cells and 2.5 on the macro cell.

From the analysis of the figure, one can note that the proposed controller – in the figure, the values are the ones achieved after convergence (∼30 seconds) - successfully equalizes the latencies up to the threshold $\varepsilon = 0.5$. On the contrary, the other two controllers fail to allocate any bitrate on BS4, as it was already saturated by the other commodities that were prioritized. The uniform distribution causes the first controller to experience a very high latency on the satellite (BS0), while the distance and consequent low signal-to-noise-ratio causes the weighted controller to allocate too much bitrate on BS5 (this behavior is further amplified by the fact that BS5 is a micro cell associated to a scaling factor of 0.1), requiring a significant amount of resource blocks.

Overall, it is possible to conclude that the proposed controller better balances

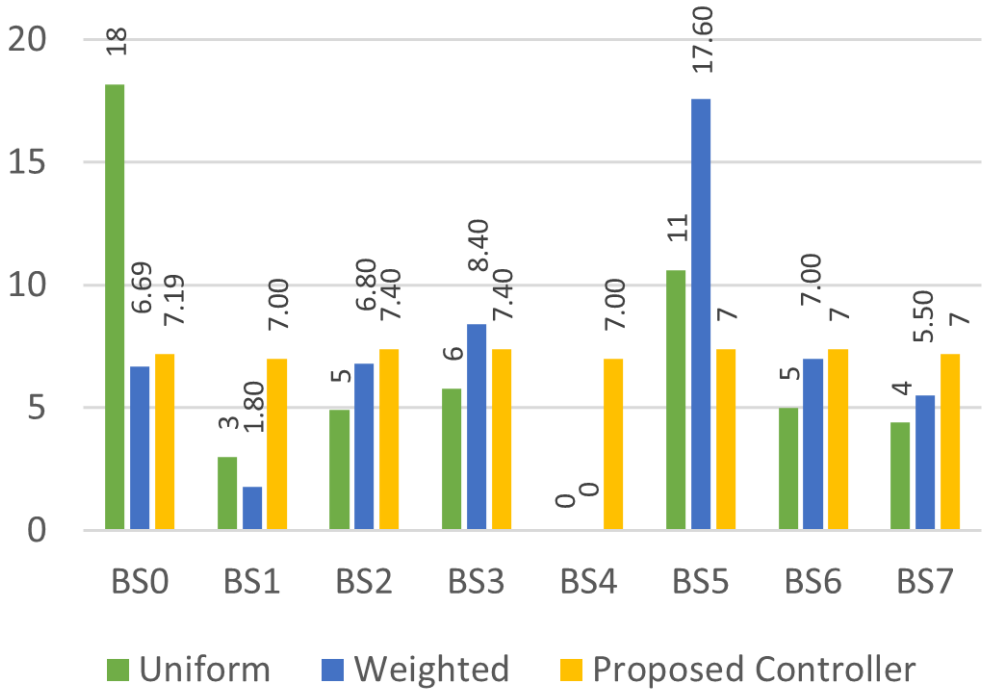


Figure 2.7. Comparison for the three considered algorithms of the latencies over APs for commodity $i = 20$

the usage of network resources because it is a feedback-based solution that steers the traffic flow based on online measurements of the latency functions. The main limitation of the proposed approach is related to the availability of the measurements needed to compute the steering decisions (i.e., the latency values in terms of assigned resource blocks), whose impact on the control traffic overhead is to be evaluated considering the control traffic already necessary for the different access technologies, and the estimation of $\bar{\beta}$ which, however, can be performed starting from the channel models and the expected traffic that the network is designed to support. Regarding the complexity of the algorithm, the computation overhead is negligible since the control law (2.11) only involves basic operations (summations, multiplications, and comparison between real numbers) that remain limited in number even for RANs with a high number of APs.

To conclude, it is worth mentioning that the two benchmarking algorithms discussed above could be used to initialize the network resource allocation, speeding up the convergence time.

2.4 Conclusion

This work develops a distributed, non-cooperative and dynamic load balancing algorithm in the framework of adversarial selfish routing with link capacities. Each provider is associated with a latency function which represents its performance as a function of the provider's load. By using Lyapunov arguments, the proposed algorithm is proved to converge to an approximate Beckmann user equilibrium, in which the latencies of the non-saturated providers are equalized up to a tolerated latency mismatch.

The algorithm is then applied to the problem of multi-connectivity, one of the key features of 5G networks, which enables the user equipment to simultaneously transmit/receive traffic flows over different access networks, to increase the transmission rate and/or improve the transmission reliability. In multi-connectivity, the traffic steering functionality is in charge of distributing the traffic load of each flow over the different access networks. This work models the traffic steering problem as a capacitated load-balancing problem by associating a latency function to each access point/user equipment radio link. The problem is then solved using the developed algorithm. An open-source simulation environment was proposed, and some numerical simulation results validate the approach.

Besides the modeling of the 5G Multi-connectivity problem as a dynamic load-balancing one, this work presents, up to the authors' knowledge, the first multi-commodity, dynamic, and adversarial load-balancing algorithm which explicitly considers capacitated providers.

Future work is aimed (i) at introducing latency constraints in the problem formulation to model more Quality-of-Service constraints of the 5G services and (ii) at considering time-varying loads.

Chapter 3

Friend-or-foe Q-Learning for network selection in 5G multi-RAT network

In this chapter, the problem of Network Selection in 5G heterogeneous networks has been faced. Differently from the work presented in the previous chapter, which focuses on how to split/steer the traffic among the APs for which a UE has an already-established connection, this work focuses on the selection of the best AP to handle a new request from a UE. In this sense, this work is complementary to the traffic splitting/steering approach of the previous chapter.

This work is based on Game Theory, and in particular on the concept of Nash equilibria. Optimal policies (i.e., an adversarial Nash equilibrium for the Markov Game) for the agents, that in this case are the APs, upon receiving a request from a UE are learned through a Multi-Agent Reinforcement Learning technique named *Friend-or-Foe Q-Learning*, which is detailed in Section 3.2.2: This technique is considered in its adversarial variant, i.e., each AP competes with the other APs to allocate the request from a UE. This assumption leads to a more realistic scenario with no centralized entity to coordinate the APs in choosing how to allocate the UE request.

3.1 Introduction

The problem of Network Selection arises in the framework of the so-called "heterogeneous networks", modern communication scenarios in which several different RATs are available to connect a user with the Core Network (CN). In such networks, when a new connection is established, a decision regarding which AP to utilize shall

be taken by either the UE or the network itself, based on a feedback-based analysis of the network state.

Different criteria (e.g., congestion state, power efficiency, reliability) may be utilized for the selection, and, based on the scope of the information gathered for the analysis, it is possible to identify three different classes of approaches [58]:

- *User-Centric approach*: in which the UE monitors the APs state and takes its connection decisions based on some thresholds-based performance parameters (e.g. SNR) measurable locally. In advanced scenarios, the UE could consider other RATs characteristics (e.g. coverage, user preferences, etc.) to better satisfy the application and user needs.

- *RAN-Assisted approach*: in this approach, information exchange is done between the AP and the UE, so that the latter can select the one it prefers based on broader feedback that may capture aspects not locally measurable, such the congestion level on the specific RATs, their expected resource allocation and their predicted/historical connection reliability.

- *RAN-Controlled approach*: the previous approaches were user-centric by nature, and consequently could only attain a sub-optimal solution to the network selection problem, in this approach the decision is taken directly by the RAN, a controller that oversees the functioning of the various RATs that constitute the Access Network. The decision taken by the RAN can either be centralized or distributed, as the RAN itself may have some functionalities distributed over the various RATs. In this approach, the UEs may be configured to report radio measurements on their local radio environment to integrate the feedback available to the centralized network controller. This solution is the one adopted by 3GPP for addressing dual-connectivity issues.

The solution presented in this work can be classified as a control strategy of the RAN-Controlled category, characterized by the distribution of the control logic over the controller of the various RATs controllers that regulate the APs connection admittance logics, so that the network resources available for the connection are optimally exploited.

From a methodological point of view, several approaches were investigated in the literature for the network selection problem, spacing from solutions based on Multiple Attribute Decision Making (MADM) [59] [20], to Fuzzy Logic control systems [60], and Game Theory-based approaches [24] [25]. Additionally, Markov Decision Processes (MDPs) and Reinforcement Learning (RL) were tested, among the others, in [61] and [62]. The proposed approach utilizes results from both RL and

Game Theory in a multi-agent framework. The problem will be modeled in such a way that the distributed RAT controllers will compete with each other for being selected to serve the connection requests. The overall goal of the control strategy will be the optimal usage of network resources, without relying on centralized control approaches – as, for example, with a common least loaded allocation logic that assigns the upcoming connections to the RAT with the lowest resource usage. In this regard, the present work employs the so-called "friend-or-foe Q-learning" algorithm to govern the network according to an adversarial Nash strategy.

3.2 Preliminaries on Learning Markov Games

3.2.1 Markov Games and Nash Equilibria

A Markov Game among N players is defined as the tuple $(\mathcal{S}, \mathcal{A}, T, \mathcal{R}, \gamma)$, where:

- $\mathcal{S}$ is the finite state space
- $\mathcal{A} = \{A_i, i = 1, ..., N\}$ is the collection of the action sets available to the various players i.
- $T(s, a_1, a_2, ..., a_N, s') : \mathcal{S} \times A_1 \times A_2 \times ... \times A_N \times \mathcal{S} \to [0, 1]$ is the state transition function, which describes the transaction probability between the two states s and s' when the agents take the actions $a_1, a_2, ..., a_N$
- $\mathcal{R} = \{R_i(s, a_1, a_2, ..., a_N) : \mathcal{S} \times A_1 \times A_2 \times ... \times A_N \to \mathbb{R}, \quad A_i = \{a_i\}\}$ is the collection of reward functions that attribute a reward to each agent when they take actions $a_1, a_2, ..., a_N$ and the system is in state s.
- $0 \leq \gamma < 1$ is the discount factor that captures the trade-off between short-term and long-term performances sought by the agents.

In this section the so-called *general sum games* are considered, meaning that no assumption is made on the cumulative reward attained by the agents, contrary to zero-sum games.

A policy $\pi_i(s) : \mathcal{S} \to \mathbb{R}^{\#(A_i)}$ is a function that maps the state of the system into a probability distribution over the actions of player i. Each player is associated with a (state,action)-value function Q_i defined as

$$\begin{aligned} &Q_i(s, a_1, a_2, ..., a_N) = \\ &\quad R_i(s, a_1, a_2, ..., a_N) + \gamma \sum_{s'} T(s, a_1, a_2, ..., a_N, s') Q_i(s', \pi_1, \pi_2, ..., \pi_N), \end{aligned} \tag{3.1}$$

in which $Q_i(s', \pi_1, \pi_2, ..., \pi_N)$ is the weighted sum of the Q_i's according to the policies π_i's. For their definition, the (state,action)-value functions represent the expected discounted reward attained over time by the players starting from state s, taking actions $a_1, a_2, .., a_N$ and following the policies $\pi_1, \pi_2, ..., \pi_N$ from there on. The goal of the controllers that will determine the policy of each agent is the one of maximizing its own value function *unilaterally* (i.e., without cooperation).

An important concept to introduce in the framework of Markov Games is the one of adversarial Nash equilibria, which are a set of policies $\pi_1, ..., \pi_N$ characterized by the following two properties:

- no player can improve unilaterally its policy, i.e., $R_i(s, \pi_1, ..., \pi_N) \geq R_i(s, \pi_1, ..., \pi_{i-1}, \pi'_i, \pi_{i+1}, ..., \pi_N)$ [1]
- no player sees its reward lowered by a change in the policies of the other players, i.e., $R_i(s, \pi_1, ..., \pi_N) \leq R_i(s, \pi'_1, ..., \pi'_{i-1}, \pi_i, \pi'_{i+1}, ..., \pi'_N)$.

3.2.2 Multi-Agent Reinforcement Learning

In scenarios in which the agents are not provided with a complete and accurate model of the system, model-free control solutions as RL have to be implemented to attain the desired system behavior.

The attractiveness of RL in Multi-agent scenarios is due to the fact that it allows the agent i behavior, described by policy π_i, to adapt to the strategy employed by the other agents. This capability becomes of crucial importance when the various agents compete one against each other, as each agent has no incentive to share information regarding its own configuration with the others. Nevertheless, RL also allows the agent to learn about the environment characteristics by directly interacting with it, meaning that no explicit knowledge of the functions T and R_j, $j = 1, ..., N$ is assumed or necessary for reaching an optimal strategy.

In this section, the Friend-or-Foe Q-Learning algorithm from [63] is employed in its adversarial variant. The additional degree of information that agent i requires other than the feedback observation of the tuple $(s, a_1, ..., a_N, s', r_i)$ is the classification of the other players as either friends (cooperating agents that try to maximize their rewards jointly) or foes (competing agents that try to maximize their own rewards unilaterally and, consequently, to minimize player i's reward).

In this algorithm, each agent i learns its (state-action)-value function Q_i according to the following rule:

[1] With a small abuse of notation $R_i(s, \pi_1, ..., \pi_N)$ is here denoted as the expected payoff obtained by agent i in state s in case all agents $j \in \{1, ..., N\}$ choose actions $(a_1, a_2, ..., a_N)$ according to their policies $(\pi_1, ..., \pi_N)$

$$\begin{aligned} Q_i(s, a_1, a_2, .., a_N) =& (1 - \alpha(t))Q_i(s, a_1, a_2, .., a_N) + \\ &+ \alpha(t)(r_i + \gamma Nash_i(s, Q_1, Q_2, .., Q_N)) \end{aligned} \tag{3.2}$$

where $Nash_i(s, Q_1, Q_2, .., Q_N)$ is computed as

$$\begin{aligned} Nash_i(s, Q_1, Q_2, .., Q_N) = & \max_{\pi \in \Pi(A_1 \times A_k)} \min_{[a_{k+1}, ..., a_N] \in (A_{k+1}, ..., A_N)} \\ & \sum_{[a_1, ..., a_k] \in (A_1, ..., A_k)} \pi(a_1) \cdot ... \cdot \pi(a_k) Q(s, a_1, ..., a_N). \end{aligned} \tag{3.3}$$

In (3.3), it is assumed, without loss of generality, that the first k players, i.e., players $1, ..., k$, cooperate with agent i and the remaining players, i.e., players $k + 1, ..., N$, are its foes. The sequence $0 \leq \alpha(t) < 1$ represents the evolution over time of the *learning rate* of the agents. Under the hypothesis that $\sum_t \alpha(t) = +\infty$ and $\sum_t \alpha(t)^2 < +\infty$ [64], then Theorem 6 in [63] proves that Foe Q-Learning (i.e., in the case in which all agents are foes, which is the one considered in the following) converges to an adversarial equilibrium, provided that such equilibrium exists.

3.3 Problem modelling

As already introduced, the network selection problem will be modeled as a Markov Game in which each AP is a competing player.

3.3.1 State Space

The state of the network can be represented by the percentage of occupied resources on each of the APs. To have a finite number of states, a possible solution is to quantize the percentage of resources with a factor q .The set of states is then defined as:

$$\mathcal{S} = \left\{ [s_1, s_2, ..., s_N], s_i \in \left\{ 0, \frac{1}{q}, \frac{2}{q}, ..., \frac{q-1}{q}, 1 \right\}, 0 \leq s_i \leq 1 \right\}, \tag{3.4}$$

meaning that there are $(q + 1)^N$ different states.

3.3.2 Action Space

The actions available to each of the agents regard the decision of whether to accept or decline the allocation of the incoming connection. Assuming that m different service classes are available to network users, a total of 2^m actions are required to model all the possible different choices. Note that some actions might be unavailable

since the APs could decide to accept only services of certain classes. The action set of user i is then defined as:

$$A_i = \left\{ [a^1, a^2, .., a^m], a^j \in \{0, 1\} \right\}. \tag{3.5}$$

3.3.3 Reward Functions

The reward that is given to each agent i for successfully allocating a service of class j depends on the service characteristics and the amount of resources involved in the allocation.

Assuming that the service requires t_j resources, it is possible to model the reward as:

$$r_i = \sigma_{ij} t_j \frac{B_i + t_j}{C_i}, \tag{3.6}$$

where B_i and C_i represent respectively the amount of resources occupied on the AP i before the new allocation and the total capacity of the AP i. The factor σ_{ij} serves the purpose of prioritizing certain services over other ones and/or modeling the fact that some APs are more appropriate, in terms of Quality of Service, for certain services. In this sense, an higher value of σ_{ij} incentivizes users i of class j to be allocated by the algorithm.

The structure of the reward allows incentivizing the agent to allocate all of their resources, while also dedicating them to the most prioritized services.

When a new service request arrives at the agents, each of them selects its action and consequently takes the decision of being available for the allocation or not. One agent is sampled randomly from the list of available ones, and the allocation procedure proceeds. The agents that offered their availability to allocate the incoming service but were not selected for the actual allocation receive a small negative reward to disincentivize the behavior of always offering the allocation availability. Furthermore, an agent that offered the allocation but was not able to fulfill it due to a scarcity of resources is given a highly negative reward to penalize its behavior and the connection is discarded.

To avoid all agents rejecting the less rewarding services, a negative reward is also given to all the agents if no agent offers its availability for the new allocation.

3.3.4 ε-Greedy Policy Selection

A fundamental concept in RL is the trade-off between knowledge exploitation and environment exploration. The update of the Q_i tables (3.2) and the solution of the *maxmin* problem (3.3) represent, respectively, the process of learning from experi-

ence, or knowledge acquiring, and its exploitation to derive a proper strategy for the player. To provide the players with an adequate degree of exploration, the action selection is subject to the following rule, known as ε-greedy selection:

$$a_i = \begin{cases} \arg\max_{a_i}(Nash_i), \text{with probability } 1-\varepsilon \\ \text{randomly chosen in the set } A_i, \text{with probability } \varepsilon \end{cases} \tag{3.7}$$

where $Nash_i$ is the operator described in (3.3), in the case in which all the APs are assumed to compete one with each other and, hence, no friend player cooperates with the agent i. The exploration vs exploitation is tuned by the parameter ε, which can incentivize exploration (random action) if high, or exploitation of the acquired knowledge ($\arg\max_{a_i}(Nash_i)$) if low. A possible refinement to (3.7) is to consider a decreasing sequence of values for ε, modeling the fact that the agent benefits more from the exploration process at the beginning, while knowledge exploitation becomes more effective as the agent experienced the system evolution and its possible states several times.

3.3.5 Maxmin Linear Programming Formulation

In general, a *maxmin* optimization problem takes the following form [65]:

$$\begin{aligned} &\max \min_{j=1,\dots,n} J(x_j) = c_j x_j \\ &A_{eq}x = g \\ &A_{ub}x \leq b, \end{aligned} \tag{3.8}$$

where $c_j \geq 0$ are scalars, $A_e q$, $A_u b$ are matrices and g and b are vectors of appropriate dimensions.

It is well known that such a formulation is equivalent to the following LP problem [65]:

$$\begin{aligned} &\max_{z\in\mathbb{R}} J(x_j) = c_j x_j \\ &A_{eq}x = g \\ &A_{ub}x \leq b \\ &z \leq c_j x_j, j = 1, ..., n \end{aligned} \tag{3.9}$$

where z is an unknown scalar that is bounded by the smallest value $c_j x_j$ through the additional third constraint.

In the context of Foe-Q-Learning, the *maxmin* problem for agent i that appears

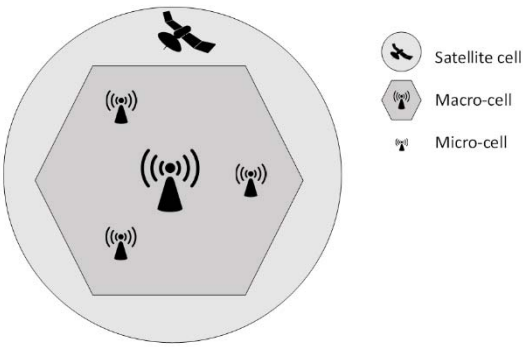


Figure 3.1. Connection Area covered by 3 different Radio Access Technologies

in (3.3) becomes:

$$\begin{aligned} &\max_{\pi\in\Pi(A_i)} \min_{[a_1,...,a_{i-1},a_{i+1},...,a_N]\in(A_1,...,A_{i-1},A_{i+1},...,A_N)} \sum_{a_i\in A_i} \pi(a_i)Q_i(s,a_1,...,a_N) \\ &\quad \pi(a_i) \geq 0 \quad \forall a_i \in A_i \\ &\quad \sum_{a_i\in A_i} \pi(a_i) = 1 \end{aligned} \tag{3.10}$$

which, by introducing the additional variables h (that are fixed equal to the objective function terms) and z (that is bounded by the smaller h_i), leads to an equivalent LP formulation of the form:

$$\begin{aligned} &\max_{\pi\in\Pi(A_i)} z \\ &\quad h_i = \sum_{a_i\in A_i} \pi(a_i)Q_i(s,a_1,...,a_N) \\ &\qquad \forall[a_1,...,a_{i-1},a_{i+1},...,a_N] \in (A_1,...,A_{i-1},A_{i+1},...,A_N) \\ &\quad \pi(a_i) \geq 0 \quad \forall a_i \in A_i \\ &\quad \sum_{a_i\in A_i} \pi(a_i) = 1 \\ &\quad z \leq h_i \quad \forall[a_1,...,a_{i-1},a_{i+1},...,a_N] \in (A_1,...,A_{i-1},A_{i+1},...,A_N) \end{aligned} \tag{3.11}$$

Algorithm 1 reports the pseudocode for the Network Selection Algorithm.

Algorithm 1 An algorithm with caption

$Q_i(s,a) \leftarrow$ arbitrary value, $i = 1, ..., N$
for connection request from a user of class j **do**

- Each player observes the state of the system and the class j of the upcoming connection
- Each player selects its action a_i with an ε-greedy policy based on their $Nash_i$ function

if $a^j = 1$ for at least one agent **then**

- The connection is allocated using the resources of one of the agents that selected an action with $a^j = 1$.

else

- One agent is selected randomly to allocate the incoming request

end if

- Each player i receive a reward r_i as described in (3.6)
- Each player i updates its Q_i table according to (3.2)

end for

3.4 Simulations

3.4.1 Simulation setup

The scenario considered in the simulation is reported in Fig. 3.1, and consists of an area covered by three different RATs. The number of agents considered is then $N = 3$. The resource considered for the connection is throughput, and each RAT had a maximum capacity of 1000 Mbps. Two service classes were modelled, the first characterised by a resource request of $t_1 = 1$ Mbps and the latter by $t_2 = 5$ Mbps. The parameter σ_{ij} was set differently for each simulation.

A total of 2000 user requests were generated, where each request had a 0.8 probability of being a new connection and 0.2 of being the end of a connection, with consequent resource deallocation. The connection requests were uniformly distributed over the two service classes.

Regarding the RL-based controller parameters, $\alpha(t)$ was set as $\alpha(t) = 1/(1 + \lfloor t/10 \rfloor)$, where the $\lfloor\rfloor$ represents the lower-integer operator, and $\varepsilon(t)$ halved every 100 iterations starting from $\varepsilon(0) = 0.6$. Finally, the discount factor γ was set to 0.9.

The simulative scenario considered is a simplified one, but maintains the dimensioning and the key characteristics of the test cases that were developed in the scope of the 5G-ALLSTAR project [66].

To validate the proposed approach, different simulations will be represented in the following considering a baseline approach and two different versions of the

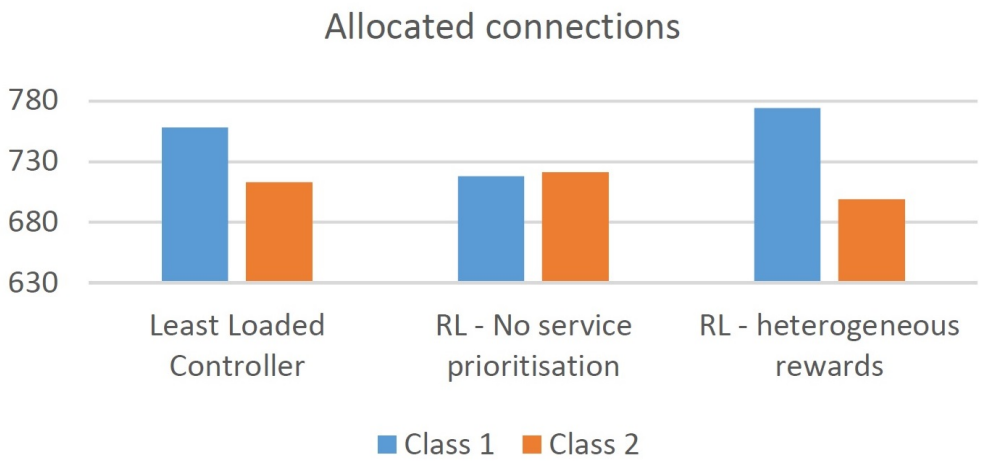


Figure 3.2. Number of allocated connections for the three controllers, divided by service class

proposed controller:

- the baseline controller considered as a benchmark follows a least-loaded AP logic, as it assigns the upcoming connections to the APs with the lowest relative resource usage. Such a controller is centralized by nature, as it requires complete knowledge of the state of the system
- a version of the proposed Foe-Q-Learning where $\sigma_{ij} = 1 \ \forall i, j$, meaning that the reward of the agents depends only on the amount of allocated resources and no priority is given to any of the two service classes
- a version of the proposed Foe-Q-Learning that has been trained with parameters $\sigma_{i,1} = 2$ and $\sigma_{i,2} = 0.2$, $\forall i$. In this way the proposed controller is trained to receive a higher per-bps reward for the first service class rather than for the second one, thus prioritizing requests coming from the users belonging to the first service class.

3.4.2 Simulation results

From the analysis of Fig. 3.2 and Fig. 3.3, it is possible to note how the two RL agents behave differently from the centralised Least Loaded controller. In particular, the controller trained with $\sigma_{ij} = 1 \ \forall i, j$, tends to uniformly accept the two services, in line with the fact that they were characterized by the same amount of reward per-Mbps, while the second RL controller (simulation three) favors the allocation of services of the first class. Overall, as it is possible to notice from Fig. 3.4,

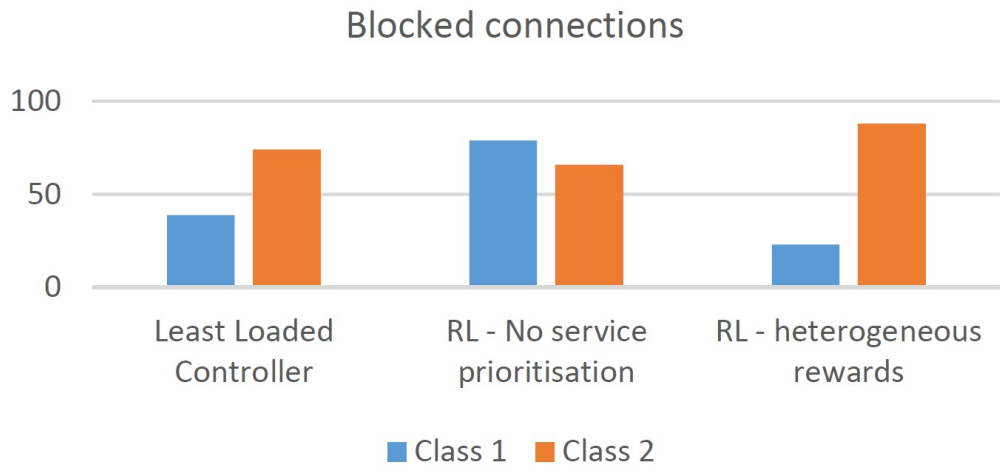


Figure 3.3. Number of blocked connections for the three controllers, divided by service class

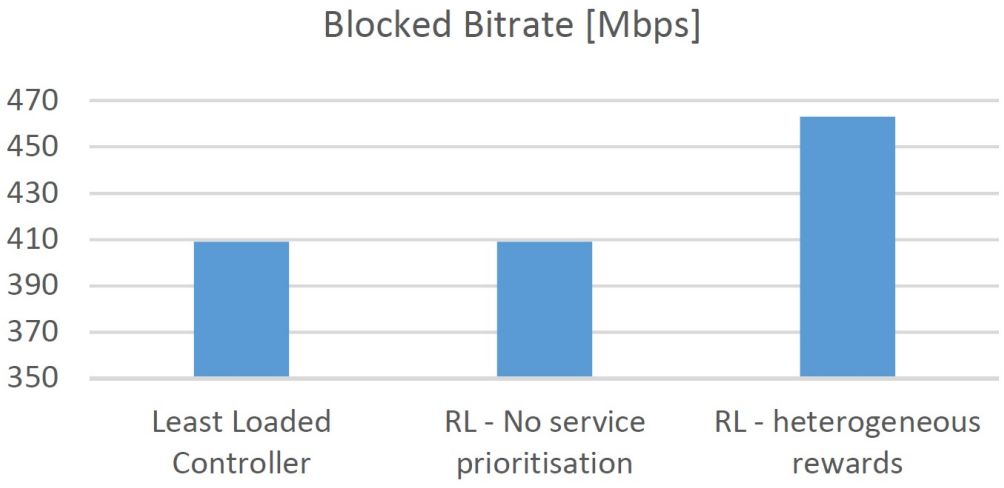


Figure 3.4. Total amount of bitrate of the blocked connections

both the Least Loaded controller and the first version of Foe-Q-Learning controller blocked a total of 409 Mbps, meaning that the first RL solution fully exploits its available resources. The second Foe-Q-Learning controller, on the contrary, allocates a slightly lower amount of throughput, blocking a total of 463 Mbps. This different behavior is since the agents obtain, for the same amount of resources, a different pay-off depending on the service class. In fact, due to the choice of the parameters σ_{ij}, the services of the first class provide ten times the amount of reward per Mbps with respect to the other. Even if the second RL controller blocked more Mbps, this translated into an improvement in performances, measured in terms of its cumulative total reward, of approximately 10%.

3.5 Conclusion

The work presented a distributed control approach for the problem of network selection. The proposed solution was based on Friend-or-Foe Q-learning, a multi-agent distributed Reinforcement Learning approach to solve Markov Games. The problem was modeled as a standard multi-agent Markov Decision Problem, and an adversarial game was formulated. The preliminary simulations presented validated the concept of the approach, while future testing on more realistic scenarios will be carried out within the scope of the H2020 5G-ALLSTAR project [66].

Chapter 4

Deep Reinforcement Learning for network selection in 5G heterogeneous networks

As in the previous chapter, the problem of Network Selection in 5G heterogeneous networks is considered again. In this chapter, the Network Selection of a UE request is made by a RAN controller that governs the various APs (even belonging to different RATs) in the RAN. This work is still based on learning the optimal policy (this time for the RAN controller instead of the APs), but, differently from the work described in the previous chapter, with a Deep Reinforcement Learning technique named DQN, which is described in detail in Section 4.3. The advantages with respect to the approach proposed in the previous chapter are reduced computation time and complexity (in *Friend-or-Foe Q-Learning* each agent has to compute a linear optimization problem at each learning step), and better results since the APs are not competing for each other to allocate the current UE request (i.e., this approach does not pay the *price of anarchy*).

This work also introduced the concept of QoE, since each UE request is assigned to a certain QoS flow, for which a reward for the DQN agent for successfully allocating such QoS flow is modeled according to the QoE perceived by the user.

Moreover, the approach described in this chapter is tested against the overmentioned 5G NR multi-RAT radio access network simulator developed by the Candidate in the context of his PhD, giving more realistic results.

4.1 Introduction

The exponential increase in bandwidth, coverage, and data rate demands, along with the diversification of use cases that are planning to use cellular RANs to provide connectivity, has prompted the development of the fifth-generation (5G) RAT. Through the support for higher mobile bandwidths complemented with low latency and more reliable communications, the 5G RAT is expected to address the significant increase in data rate demands that network operators are expecting and to support the diversification of services required by UE during the coming years. Moreover, the 5G specifications, starting with [67], will include other RATs in the 5G environment, such as 4G Long Term Evolution (LTE) and satellite AP. In this system, where the connection demand continues to increase, appropriate network resources management is required since an optimal allocation of those resources will guarantee better performances and will help to ensure user requirements in terms of QoE without overloading the network.

In this work, a network selection technique relying on MDPs and DQN algorithm [68] has been studied. A centralized controller will take care of allocating requests in the best way coming from UE analyzing the network state in terms of APs load and UE perceived transmission power. The goal of this study is to show the effectiveness of the proposed deep reinforcement learning approach by simulations with a realistic multi-RAT (5G/4G/Satellite) network scenario. Moreover, several classes of user requests have been modeled in order to represent different connection service requirements in terms of downlink bitrate, QoS requirements, and QoE profiles.

Network selection plays a fundamental role in providing stable connections with an adequate level of QoS. Hence, network operators and providers commonly exploit several advanced techniques to select the best AP to allocate new connections. Among the various techniques proposed in the literature, MADM proved to be one of the most flexible solutions to capture user preferences and QoE-related aspects in the decision process [69–73]. In MADM solutions, the information characterizing the decision-making is made by the so-called attribute values and attribute weights: The first ones describe characteristics, qualities, and performances of different alternatives, whereas the latter ones are used to measure the relevance of attributes.

By modeling the network selection problem as a MADM, it is then possible to decide the trade-off among service QoS requirements, user preferences, and overall network congestion.

A similar approach is followed in this work, in which a different QoE profile is associated with the various connections, depending on its specific service character-

istics.

Among other solutions, it is possible to mention fuzzy logic approaches [60, 74–76], a methodology that allows fast decision-making, but relies heavily on the operator's knowledge and best practices, and game theory [25, 77–80]. In game theory based approaches, the problem is modeled as a set of players/agents coupled with a set of network states and possible agent actions, commonly utilizing the MDP framework [64]. The main idea behind this method is that the player's actions are influenced by the choices and actions of the other players. The interaction among the players can either be adversarial, i.e., each agent tries to maximize its performance, or cooperative, when agents share a common objective.

The approaches mentioned so far are typically employed in scenarios in which the controller is provided with a model of the network and user behavior, such as a statistical distribution of the incoming connection requests and QoE profiles, like in [81,82], where the authors studied how to maximize QoE/QoS for specific services (e.g., video streaming applications. On the contrary, this work employs RL [64], a model-free control methodology that allows the network controller to automatically acquire knowledge of the system by interacting with it and experiencing its response to different control policies.

RL has been extensively applied in the network control domain [83–88] and has become particularly appealing over the last few years due to the innovations bought by its deep-learning-based variant, namely Deep Reinforcement Learning (DeepRL) [68], that allowed RL-based controllers to address previously challenging problems due to their complexity and high dimensionality [89]. DeepRL has also been used for network selection and radio resource assignment, respectively [84,90]. This work differs from these other two because it aims at maximizing the user's perceived QoE in a multi-RAT environment, where multiple radio access technologies are available at the same time. For multi-RAT network control, deep learning approaches (e.g., using Long Short-Term Memory (LSTM)) have been used in [91], which focuses on the cloud-edge computation offloading in satellite-UAV-served 6G networks. The main contributions of this work are:

1. The design of a two-step network control algorithm based on deep reinforcement learning for the problem of network selection and optimal resource management in the heterogeneous 5G networks setting also envisages the presence of satellite communication systems.

2. The inclusion in such a control framework of QoE maximization by considering three different service types with different QoS-QoE relations.

3. The development of an open-source network simulator [51] able to model sev-

eral different radio access technologies, including satellite systems, in terms of network resource usage.

4.2 Sketch of the control algorithm

The algorithm designed in this work is a 2-step process: First, the controller that governs the RAN receives a connection request and determines which available AP it should be allocated. The AP reserves the allocation of the network resources needed to satisfy the connection minimum QoS requirements to guarantee service provision. Then, the distributed controllers that oversee the various APs distribute the remaining network resources to the connections they sustain to improve the QoE of their users. Fig. 4.1 reports a functional diagram of the proposed control scheme, highlighting the flowchart of the algorithm and the related data flow.

The first part of the proposed control algorithm will be based on a deep reinforcement learning agent, whereas the network resource allocation will distribute the available resources over the various connections according to their priority.

Section 4.3 provides the reader with the necessary background information on MDP and DeepRL.

4.3 Markov decision process, Q-learning, and deep-Q-network

An MDP is defined as the tuple $\{\mathcal{S}, \mathcal{A}, T, R, \Sigma, \gamma\}$ where $\mathcal{S}$ and $\mathcal{A}$ are the (continuous or discrete) finite state and action set, respectively, T is the transition probability function $T : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$, with $T(s, a, s')$ denoting the probability that the next state is s' when the current state is s and the chosen action is a. and with $\sum_{s' \in \mathcal{S}} T(s, a, s') = 1$, R is the one-step reward function $R : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to \mathbb{R}$, Σ is the initial state distribution and $\gamma \in (0, 1)$ is the discount factor that weights future rewards against immediate ones. The set of actions might be state-dependent as not all the actions might be available at each state, the set of actions available at a given state $s \in \mathcal{S}$ will be denoted by $\mathcal{A}(s) \subseteq \mathcal{A}$.

A deterministic policy $\pi : \mathcal{S} \to \mathcal{A}$ selects one action for each state. Let Π be the set of feasible policies π such that $\pi(s) \in \mathcal{A}(s)$ for all $s \in \mathcal{S}$. The expected discounted reward obtained by starting from state s and following policy π thereafter is represented by the state-value function, defined as

$$V_\pi(s) = \mathbb{E}_\pi\Big(\sum_t \gamma^t R(s_t, a_t, s_{t+1}) | s_0 = s \Big) \tag{4.1}$$

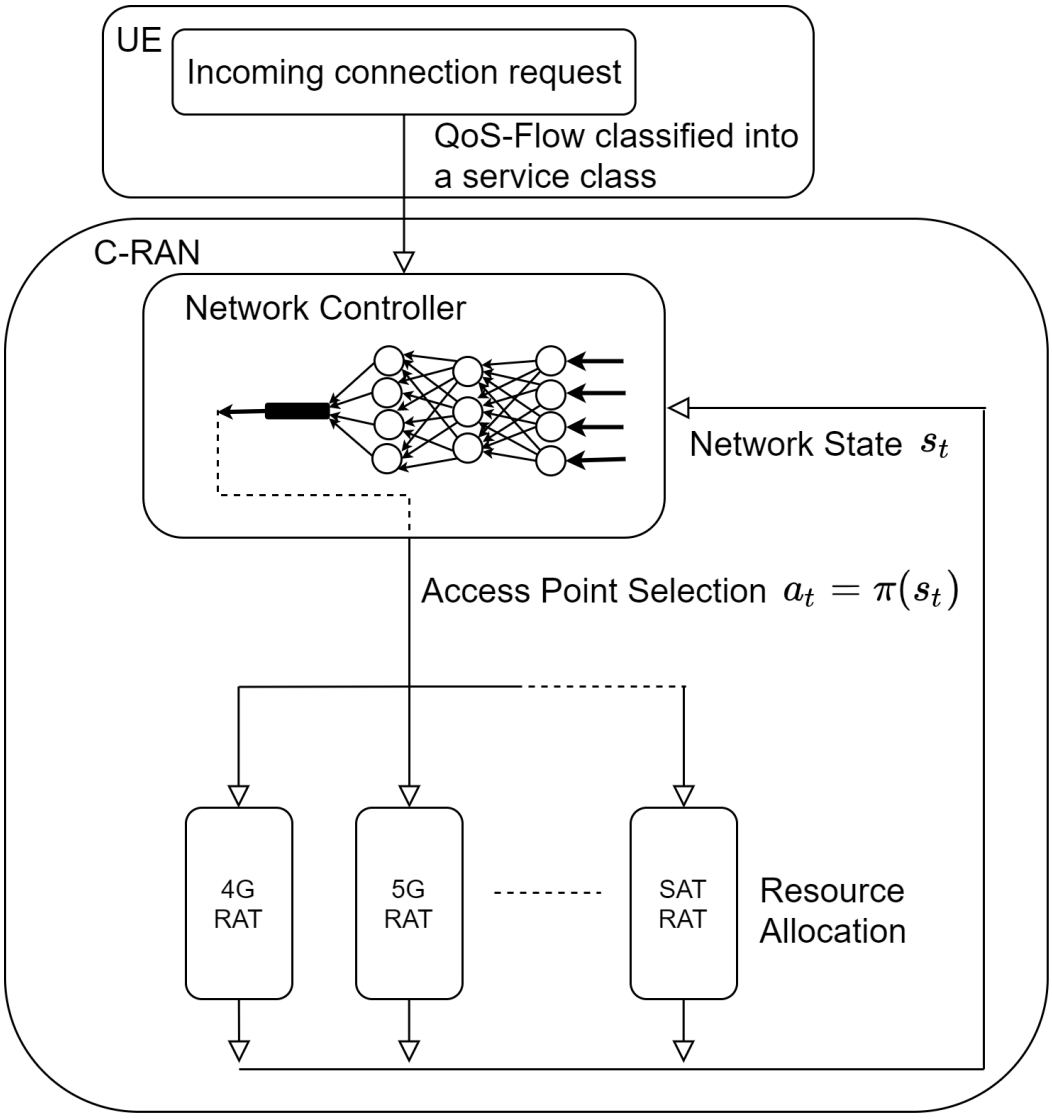


Figure 4.1. Flow-chart of the control algorithm

where $\mathbb{E}_\pi$ is the expected value under policy π and s_t and a_t represent the state and action at time t. Similarly the state-action-value function

$$Q_\pi(s,a) = \mathbb{E}_\pi\Big(\sum_t \gamma^t R(s_t,a_t,s_{t+1})|s_0 = s, a_0 = a\Big) \tag{4.2}$$

represents the expected discounted reward obtained by following the policy π when starting from state s and taking action $a \in \mathcal{A}(s)$.

Solving the MDP means to find the optimal policy π^* that maximizes the expected cumulative discounted reward, i.e., $\pi^* = \arg\max_{\pi\in\Pi} V_\pi(s)$. Dynamic programming approaches [64] can be used to determine exactly π^*. However, they typically require the complete knowledge of the MDP dynamics - in particular of T and R - and their computing time exponentially increases with the dimensions of state and action sets.

Conversely, RL algorithms, such as Q-learning, aim to obtain an estimate of the optimal state-action-value function Q_{π^*} based on the experience the controller gathers by interacting with the environment.

The standard update rule for Q-learning is

$$Q(s_t,a_t) = (1-\alpha_t)Q(s_t,a_t) + \alpha_t(r_t + \gamma \max_{a\in\mathcal{A}(s_{t+1})} Q(s_{t+1},a)) \tag{4.3}$$

where $r_t = R(s_t,a_t,s_{t+1})$ is the measured reward obtained at time t and $\alpha_t > 0$ is the learning rate, which, in order to assure convergence, is subject to the conditions $\sum_{t=1}^{\infty}\alpha_t = \infty$ and $\sum_{t=1}^{\infty}\alpha_t^2 < \infty$.

The balancing between exploration and exploitation is controller by the parameter $\varepsilon_t \in [0,1]$ in the so-called ε-greedy policies: at any time t, the agent chooses a random action with probability ε_t, whereas it chooses the action that maximizes the state-action-value function (i.e., $\arg\max_{a\in\mathcal{A}(s)} Q(s,a)$) with probability $(1-\varepsilon_t)$.

It is worth noting that in standard RL approaches, the Q function is updated only for the visited state-action pairs. Thus, in order to have a complete estimation of the optimal Q function, it is needed to visit at least once every state-action pair. This implies that the state space $\mathcal{S}$ and the action space $\mathcal{A}$ must be finite and discrete, and if their dimensions increase, RL algorithms incur the so-called *curse of dimensionality.*

To address these issues, the DQN algorithm was proposed in [68] as a deep learning solution for function approximation-based Q-learning [64]. DQN approximates the Q function by means of a deep neural network able to approximate high-dimensional functions with low-dimensional representations. The training process for the neural network is detailed in [68], and despite having included some technical solutions to address the neural network limitations, such as the target network and

memory buffers, conceptually it remains the same as in the standard Q-learning, with (4.3) replaced by the neural network training process and in particular by the weight updates.

The main advantage of using DQN is its ability to cope with continuous state spaces, and it proved capable of solving complex problems, such as playing video games. Note that DQN still considers discrete action sets, actor-critic solutions such as the deterministic deep policy gradient (DDPG) should be used when dealing with continuous actions.

4.4 Problem modelling

This section presents the modeling of the network selection problem as an MDP. In particular, sections 4.4.1-4.4.3 formulate the sets and functions required for the MDP formalism, while sections 4.4.4 and 4.4.5 detail the physical processes that allow the conversion of network resources into bitrate provision.

Let I be the set of UEs connected within a RAN constituted by a set P of APs. Each UE $i \in I$ is connected to an AP $p \in P$ of the RAN, characterized by a certain amount W_p of PRBs available. In addition, let $P^i \subseteq P$ be the set of the APs available at UE i, depending on its position and antennas. Moreover, let K be the set of different service types considered, each one characterized by a different minimum bitrate B_k, $k \in K$. Finally, let n_{pk} be the number of requests of type k allocated to an AP p.

Three different types of services are considered here, as in [92], namely: elastic, non-elastic, and multi-codec, each characterized by a different QoE profile.

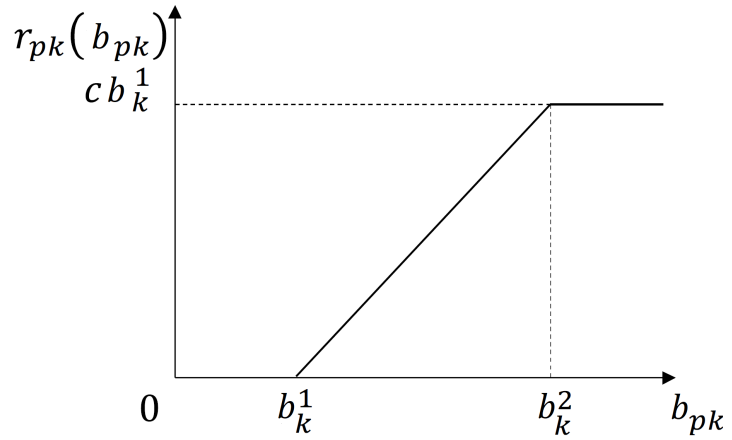
Let b^i_{pk} be the bitrate allocated on AP p for the service k requested by UE i. It is possible to model the three QoE profiles as the functions $r^i_{pk}(b^i_{pk})$ depicted in Fig. 4.2.

In particular:

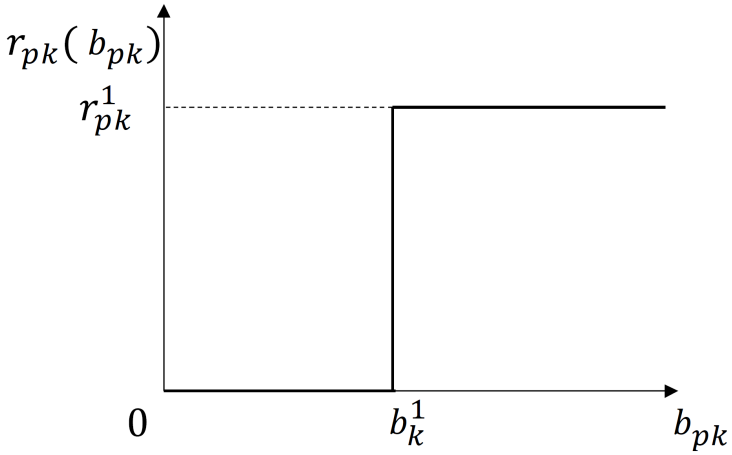
1. Elastic services have a linear QoE behavior with respect to the allocated bitrate, starting from a minimum level b^1_k up to a maximum bitrate b^2_k, where the perceived quality is saturated, as depicted in Fig. 4.2a. This service captures applications such as web surfing and file downloading.

2. Non-elastic services have a threshold-like behavior with respect to the allocated bitrate. Thus, if the bitrate is less than b^1_k, the perceived quality is 0; otherwise, it is maximal, as depicted in Fig. 4.2b. This service type represents well real-time applications with guaranteed bitrate requirements.

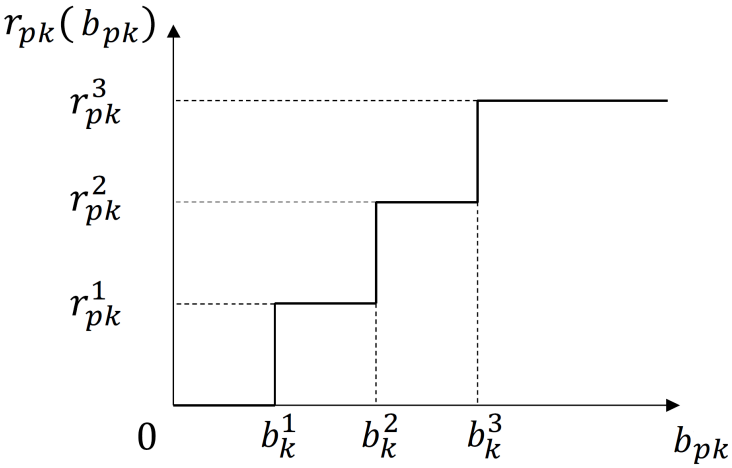
3. Multi-codec services have a stair-like QoE profile, as the perceived quality has



(a) QoE profile of elastic services ($k = 1$)



(b) QoE profile of non-elastic services ($k = 2$)



(c) QoE profile of multi-codec services ($k = 3$)

Figure 4.2. QoE profiles of the different service types

a different threshold corresponding to the utilized codec, depending on the amount of bitrate allocated b_k^1, b_k^2, b_k^3, as reported in Fig. 4.2c. This service type represents multi-codec video and audio streaming.

The proposed modeling of the services is compliant with the 5G standards, as the so-called QoS flows that constitute the various connections can be associated with one of the three service types introduced above depending on their QoS requirements and characteristics.

4.4.1 State space definition

As already introduced, each AP is characterized by the number of its physical resources available for allocation, denoted as W_p, $p \in P$.

To allow the controller to take an optimal decision on the allocation of a new incoming connection request from a given UE, the state of the network should contain information regarding:

i) the congestion level of the physical resources over the various APs;

ii) the coverage quality that the APs provide to the UE;

iii) the service class,

to infer its associated QoE profile and its bitrate requirements.

In this sense, the minimum quantity of physical resources that need to be allocated to sustain a single QoS-flow i of type k on a given access point p is denoted as w_{pk}^i, with $i \in I_{pk}$ is defined as the set of QoS-flows of type k related to AP p. Note that, referring to figures 4.2a-4.2c, this quantity represents the number of resources needed to provide the UE with a connection with an associated bitrate b_k^1.

Let $\eta_p^1(t)$ denote the number of resources allocated at time t to sustain the allocated services (i.e., the number of physical resources required to support the ongoing QoS-flows at their minimum bitrate level). By definition,

$$\eta_p^1(t) = \sum_{k \in K} \sum_{i \in I_{pk}} w_{pk}^i(t), \quad p \in P \tag{4.4}$$

Let $l_p(t)$ the load level of an AP p, defined as the allocated physical resources over the total available ones:

$$l_p(t) = \frac{\eta_p^1(t)}{W_p}, \quad p \in P. \tag{4.5}$$

Given a UE $i \in I$ requesting a service of type $k \in K$, the state space is then given by the following three quantities:

1. The load level related to each AP $p \in P$;
2. The Reference Signal Received Power (RSRP) value $\mathcal{P}_{ip}$ for each AP p, measured by the UE itself;
3. The minimum amount of bitrate required for the requested service class B_k (b_k^1 in figures 4.2a-4.2c).

The state set can then be defined as

$$\mathcal{S} = \left\{ s = \{(l_p)_{p \in P'} (\mathcal{P}_{ip})_{i \in I, p \in P'} (B_k)_{k \in K}\} \right\}. \tag{4.6}$$

The resulting state $s \in \mathcal{S}$ is a vector with $2|P| + 1$ elements. With little abuse of notation, $l_p(s)$, $\mathcal{P}_{ip}(s)$ and $B_k(s)$ represent respectively the load level op AP p, the RSRP value, and the minimum required amount of bitrate in state s.

4.4.2 Action space definition

When a new connection request arrives at the network controller, there are two possible outcomes:

1. The controller accepts the request and allocates it to (exactly) one AP p.
2. The connection is rejected as there are no APs that can handle it due to insufficient resources.

The RAN controller is then required to act as an advanced Connection Admission Controller (CAC).

The action set can be defined similarly to in [92]. Let δ_p be a vector with $2|P|+1$ values, i.e., the same dimension of the state vector $s \in \mathcal{S}$, where all the values are zeros, but the element associated with the AP p. The single non-zero element in δ_p represents the extra load that would be added to the access point p in case the new connection request is accepted. It follows that, in each state s, a request service may be allocated on AP p if and only if $s + \delta_p \in \mathcal{S}$, i.e., by allocating the new request to the AP p, the newly generated state still belongs to $\mathcal{S}$.

The action set available in a state $s \in \mathcal{S}$ is then defined as

$$\mathcal{A}(s) = \left\{ (\zeta_1, ...\zeta_{|P|}) \middle| \sum_{j \in 1,...,|P|} \zeta_j = 1, \zeta_j \in \{0,1\} \forall j \right\} \cup \mathbf{0}, \tag{4.7}$$

where $\mathbf{0}$ is a P-vector of zeroes, and the action is a vector whose only non-zero element is equal to one and indicates which AP has been selected for the allocation. The special case in which $a_t = \mathbf{0}$ represents a condition in which the connection

request must be rejected due to a lack of network resources, as no AP can allocate the incoming request assuring its minimum required bitrate.

In the simulation in section 4.5, the requests of service type $k \in K$ for each UE are assumed to arrive according to a Poisson distribution in time with mean value υ_k and a termination rate following an exponential distribution with mean μ_k.

4.4.3 Reward definition

In the presented definition of the states and actions, it was assumed that the network controller only allocates the network resources needed to satisfy the minimum amount of bitrate required by the various connections. As introduced in Section 4.2, the network control algorithm follows a two-step procedure: Firstly, it selects which AP will serve the incoming connection request. Then, each AP distributes its remaining resources $\eta_p^1(s)$ over its connections, according to some prioritization order that may take into account the user tariff or operator preferences.

In our simulations, the APs will firstly distribute their available resources uniformly to the multi-codec services, so each connection receives a bitrate up to b_3^3. Afterward, the remaining resources are uniformly distributed to the elastic services up to a bitrate of b_1^1. Non-elastic services, due to their threshold-like behavior, are always given a bitrate of b_2^1.

To define the reward function, it is needed to introduce S_{pi} as the amount of additional bitrate that the AP p is able to provide to the connection i using a share of its remaining resources. This quantity is directly linked to the QoE profile associated with the connection. As it is possible to notice from Fig. 4.2, the QoE obtained by the allocation depends on the minimum bitrate allocated by the DQN algorithm $b_k^1 + S_{pi}$, i.e., the total bitrate available to the service i of class k.

The reward function shall then capture three cases:

1. The connection request is rejected (i.e., no AP allocates the connection).
2. The connection is allocated on an AP with low resource usage.
3. The connection is allocated on an AP that is already providing several other connections.

To capture those three cases, the reward $r_t(s_t, a_t, s_{t+1})$ obtained by the controller with allocating a connection i of class k on AP p can be defined as

$$r_t(s_t, a_t, s_{t+1}) = \begin{cases} -r^0 < 0 & \text{if } a_t = \mathbf{0} \\ r_{pk}(b_k^1 + S_{pi}) & \text{if } l_p(t+1) \leq 0.5 \\ r_{pk}(b_k^1 + S_{pi}) - r^{sat} & \text{if } l_p(t+1) > 0.5. \end{cases} \tag{4.8}$$

The negative reward $-r^0$ represents a penalty given to the agent if the allocation is rejected to capture the cost incurred by the network operator in failing to provide a connection. The term $r_{pk}(b_k^1 + S_{pi})$ is a positive reward, shaped depending on k as in Fig. 4.2, that captures the QoE of the new user, and the term $-r^{sat}$ is a negative reward subtracted from $r_{pk}(b_k^1 + S_{pi})$ is case the new allocation is destined to an AP whose saturation level is higher than the desired threshold (50% in our case). The long-term maximization of this reward allows the network controller to maximize the overall QoE of its users while keeping the connection rejection rate minimized.

4.4.4 5G NR and 4G LTE resource allocation description

In order to relate the physical resources that appear in the state definition with the transmission bitrate needed by the reward function to estimate the QoE level, it is now necessary to detail their relationships and how one translates into the other for both terrestrial and satellite APs.

5G NR APs have a limited set of resources [3], both in terms of frequency bandwidth and time to allocate UE requests. The minimum allocation unit for a 5G NR AP is the PRB, each composed of 12 frequency subcarriers with a $2^\mu \times 15$ KHz bandwidth and a time duration of $2^{-\mu} \times 1$ ms, where $\mu \in \{0, 1, 2, 3, 4\}$ is the parameter called *numerology* defined by 5G NR standards [3].

For 4G LTE APs, the definition of PRB still stands, but the numerology parameter is constrained to $\mu = 0$, so there is no flexibility in using less/more subcarrier bandwidths and more/less time slot durations. Even if 4G LTE will likely to be replaced by 5G NR in the next few years, it has been considered in this work since it is currently the predominant radio access technology for mobile devices, and its seamless integration in the multi-connectivity framework allows for more stable and broadly available connectivity.

The receiving power, or RSRP, $\mathcal{P}_{ip}$ that appears in the states of (4.6) represents the transmission power measured by the UE $i \in I$ between itself, and the AP $p \in P$ is computed as follows:

$$\mathcal{P}_{ip} = \mathcal{P}_p \cdot G_p \cdot L_p \cdot L_{ip} \tag{4.9}$$

where $\mathcal{P}_p$ is the AP's antenna power, G_p is the AP's antenna gain, L_p is the AP's feeder losses, and L_{ip} is the path loss between UE i and AP p.

In the simulations in Section 4.5, the path loss L_{ip} is computed through the COST-HATA model [93], which is a statistical model that considers many factors as the building density (rural, suburban, urban), the carrier frequency used for the communications, and the relative heights of UE and AP.

In order to estimate the number of resource blocks to be allocated by the AP

$p \in P$ for the communication with the UE $i \in I$, the Signal over Interference plus Noise Ratio (SINR) has to be computed. The thermal noise part can be computed according to

$$\mathcal{N}_p = k_b T^{env} B_p \Theta_p \tag{4.10}$$

$$\Theta_p(t) = \frac{\sum_{\tau \in (t-T,t)} \sum_{j \in I \setminus i} C_{jp}(\tau) N_{jp}(\tau)}{T \cdot \#R_p} \tag{4.11}$$

where $\Theta_p(t)$ is the Resource Blocks Utilization Ratio (RBUR) of AP p at time t, k_b is the Boltzmann constant, T^{env} is the environmental temperature, B_p is the total bandwidth of the AP p, T is the length of the moving average, $C_{jp}(t)$ is equal to 1 if UE j is connected to AP p at time t and 0 otherwise, and $N_{jp}(t)$ is the number of PRB allocated to AP p to UE j, and $\#R_p$ is the total number of resource blocks of AP p.

The interference part is computed as follows:

$$\mathcal{I}_{ip} = \sum_{p' \in P, p' \neq p} F_{pp'} P_{ip'} \Theta'_p(t) \tag{4.12}$$

where $F_{pp'}$ is 1 if AP p and p' share the same carrier frequency and 0 otherwise.

Using (4.9), (4.10) and (4.12) it is possible to compute the SINR, and so it is possible to estimate the data rate that can be transmitted by allocating one PRB to UE i using the Shannon formula

$$r_{ip} = 2^{-\mu} 10^{-3} B_{\text{PRB}} \log_2(1 + \text{SINR}_{ip}) \tag{4.13}$$

where B_{PRB} is the bandwidth of a single PRB, and it can be computed as $B_{\text{PRB}} = 12 \cdot 2^{\mu} 15$ KHz.

Now, given a certain bitrate request b_k from UE i requesting service of class k, it is possible to compute the number of resource blocks to be allocated by AP p to satisfy the request: $n_{ip}^{\text{PRB}} = \lceil (b_k / r_{ip}) \rceil$.

4.4.5 Satellite resource allocation description

Contrary to ground APs, the satellite APs usually use TDM to serve multiple UEs at the same time. In this case, the minimum allocation unit is a block of symbols that occupies a certain time slot in the satellite time frame.

The receiving power P_{ip} can still be computed as (4.9), but in this case, the path loss function will be the Free Space Path Loss (FSPL):

$$L_{ip}^{\mathrm{FSPL}} = \left(\frac{4\pi d_{ip} f}{c}\right)^2 \tag{4.14}$$

where d_{ip} is the Euclidean distance between UE i and AP p, f is the carrier frequency used in the communications and c is the speed of light.

The thermal noise can be computed as in (4.10), and the interference can be computed as in (4.12). Using the Shannon formula (considering that in the satellite case the bandwidth used is the total bandwidth of the satellite AP since TDM utilizes all the bandwidth for a specified amount of time inside the satellite time frame), one has that the bitrate obtainable by a single block of symbols is:

$$r_{ip} = bB \log_2(1 + \mathrm{SINR}_{ip}) \tag{4.15}$$

where b is the ratio between the number of symbols in a single block of symbols and the total number of symbols of the satellite AP. The number of blocks to be allocated for a requested bitrate b_k from UE i requesting a service of class k is then computed as $n_{ip}^{\mathrm{block}} = \lceil (b_k / r_{ip}) \rceil$.

4.5 Simulation results and validation

In order to demonstrate the effectiveness of the proposed approach, a simulative environment has been built up according to the model definition previously introduced.

4.5.1 Scenario definition

As shown in Fig. 4.3, a scenario consisting of four terrestrial APs (NR1 and NR2 are 5G NR APs and the remaining two are 4G LTE APs) and a satellite AP in a 2.5×2.5 km^2 area has been developed to demonstrate the proposed work.

In particular, for 5G NR access points, a carrier frequency of 1.7 GHz (band n66) with numerology $\mu = 2$ has been considered, while for 4G LTE access points, a carrier frequency of 800 MHz (band 20) has been considered. All the terrestrial APs have 20 dB power, 16 dB antenna gain, and 3 dB feeder losses. For the satellite access point, the Inmarsat implementation of example 6.6.2 of [55] has been considered. A total of 100 UEs has been considered in the given area. Each of them follows, as described before, a Poisson distribution for requesting data with a certain service type and for the duration of such requests; the parameters for each service type are described in table 4.1. Moreover, the RL parameters have been set as follows: $\gamma = 0.9$, $\varepsilon_0 = 1$, $\varepsilon - \mathrm{decay} = 0.9995$ and $\varepsilon - \mathrm{min} = 0.01$. As for the DQN parameters, a replay buffer of 2000 tuples, a batch size of 64 tuples, and the update

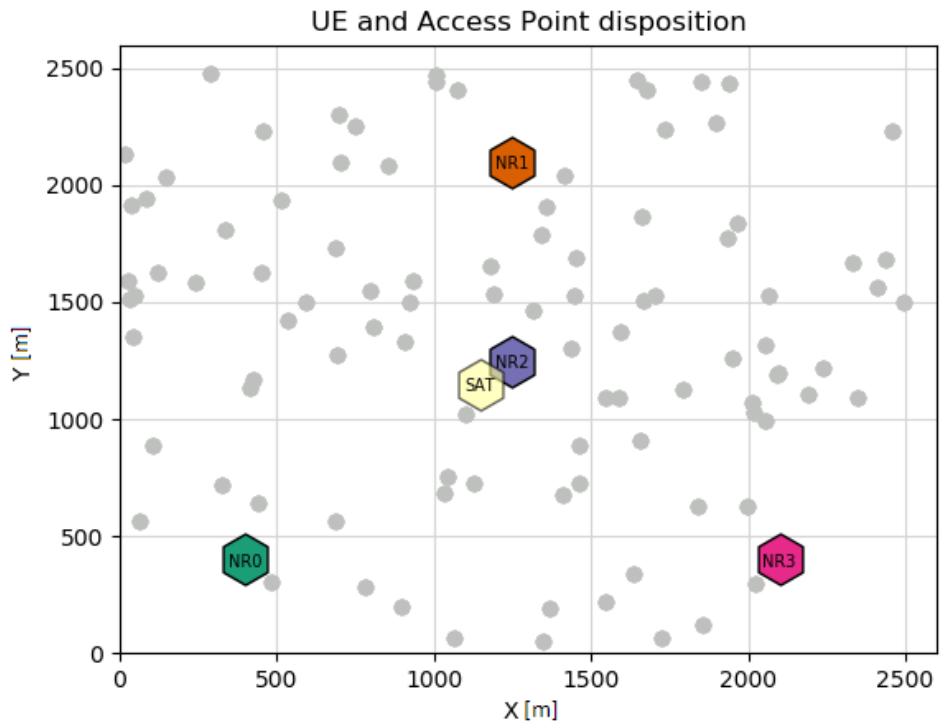


Figure 4.3. Considered network scenario

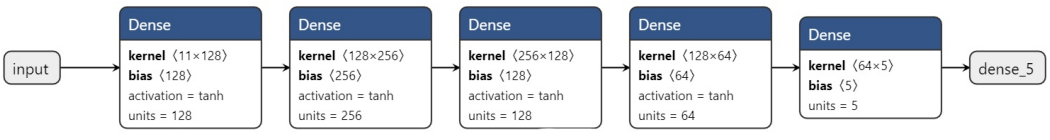


Figure 4.4. Deep Neural Network architecture used for the simulations

	Elastic	Non-elastic	Multi-codec
Bitrate (Mbps)	10	200	100
Arrival rate (s)	2	6	4
Dwelling time (s)	30	120	90

Table 4.1. Parameters of the various service types

of target network weights every 50 steps have been considered. Finally, the Deep Neural Network hidden layers, whose structure is represented in Fig. 4.4, have tanh activation function, a learning rate of 10^{-4}, and the network performs $4 \cdot 10^4$ training steps before finishing the training. The training process and its testing using the proposed radio access network simulator, run on an Intel Core i7 6700HQ machine with 16 GB RAM. No dedicated GPU has been used for the training process since the small size of the batches. Most of the computation complexity is, of course, in the training process of the four hidden layers of the DQN network; once trained, DQN has $\mathcal{O}(1)$ computational cost to compute the best action.

4.5.2 Simulation results

The results displayed in figures 4.5-4.11 will focus on the performance of the controller in terms of QoS-flows allocation and their management. In order to validate the results of the proposed DQN algorithm, a set of other approaches have been simulated. In particular, a classical, tabular, Q-learning (QL in figures 4.5-4.11 approach has been simulated, together with a least loaded (LL in figures 4.5-4.11) approach, where a new request will be allocated to the least loaded AP, and a Max-RSRP (MR) approach, where a new request will be allocated to the AP with the maximum receiving power. The Q-learning approach shares the same MDP representation as the one presented for the DQN, except for the fact that the state-space needed to be discretized so that the AP loads and the RSRP values contained in the states in (4.6) were uniformly quantized into four levels.

The various controllers have been tested on the same scenarios to obtain fair performance results. Moreover, to ensure more balanced experiments, the results are the average between ten different scenarios, each tested by all the different controllers. Finally, both the DQN and the QL controllers have been trained before executing the simulations. Several metrics are shown to understand better the performances of the controllers with respect to each other.

As it emerges from Fig. 4.5, the DQN controller outperforms the other controllers in terms of rejection rate, even if both the Max-RSRP (MR) one and the Q-learning (RL) one have similar results. This behavior is not surprising since the

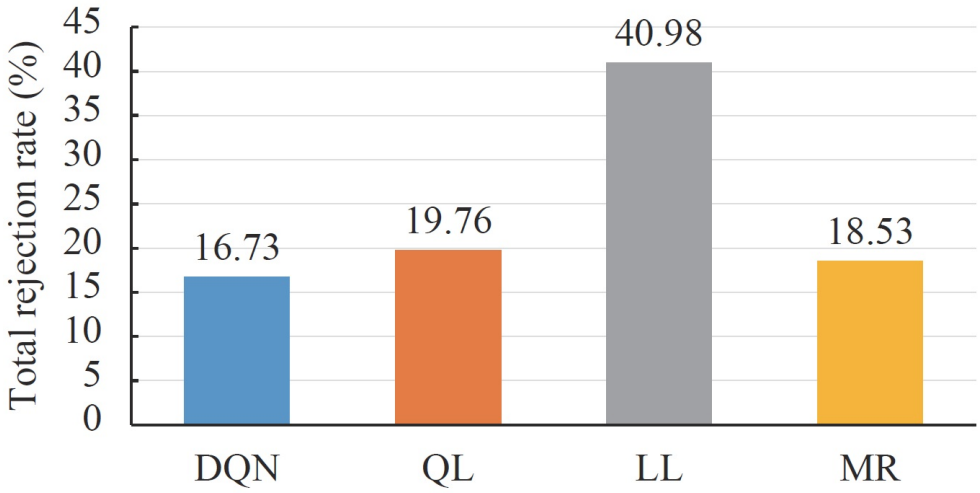


Figure 4.5. Overall rejection rate

Max-RSRP approach allocates requests to the AP with minimum path-loss, so the number of requested physical resources will be in general lower, and the Q-learning approach has a similar behavior w.r.t. the DQN approach, since the only difference is in its finite state space.

Fig. 4.6 reports the rejection rate of each controller divided by service type. In Fig. 4.6, it is possible to note how all controllers allocate a lower percentage of the non-elastic service requests, whereas the LL controller shows a significantly higher rejection rate for the elastic services.

In terms of bitrate, the DQN approach is found to be the best one, allocating around 48 Gbit over the accepted incoming requests.

Fig. 4.7 details the allocated bitrate percentage with respect to the total requested bitrate divided by service type.

The result demonstrates that DQN behaves almost in the same way as MR for what concerns the elastic services, while it allocates about 6% more than the other approaches for what regards non-elastic traffic and about 3% for what regards the multi-codec requests.

In addition, from Fig. 4.8, which represents the average percentage of successful allocations in each AP on all the requests made by the UEs, it is evident that the least loaded controller is the one that better balances the load between the APs. Despite its limited performances according to the other metrics presented, due to its definition, it allocates requests to the least used AP at the given time instant, resulting in an overall reasonable balance among the APs.

The other controllers appear to be less balanced when allocating resources, with

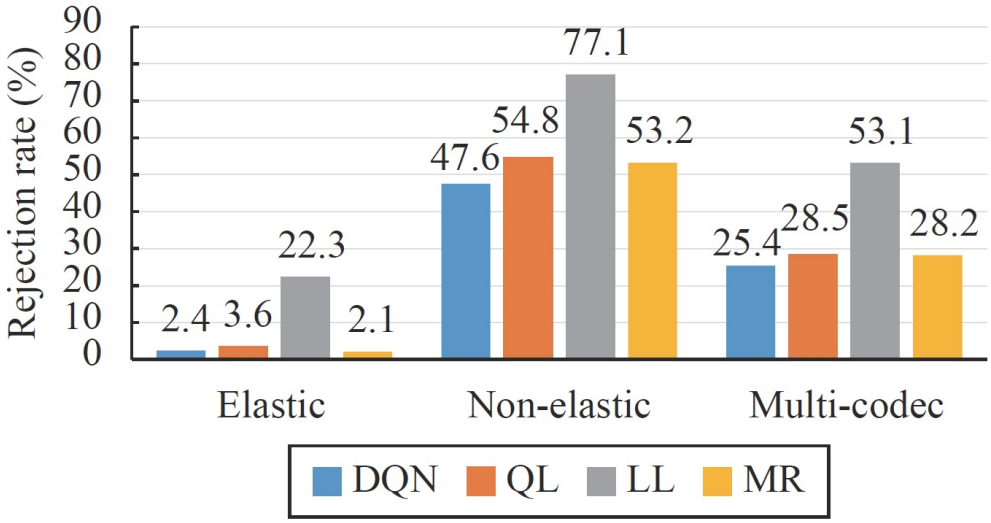


Figure 4.6. Rejection rates divided by service type

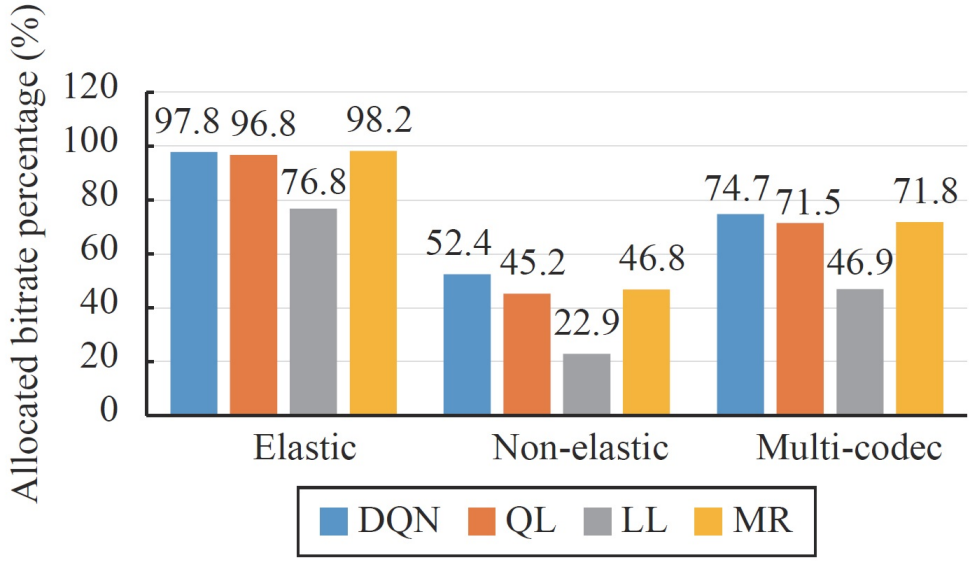


Figure 4.7. Allocated bitrate percentage divided by service type

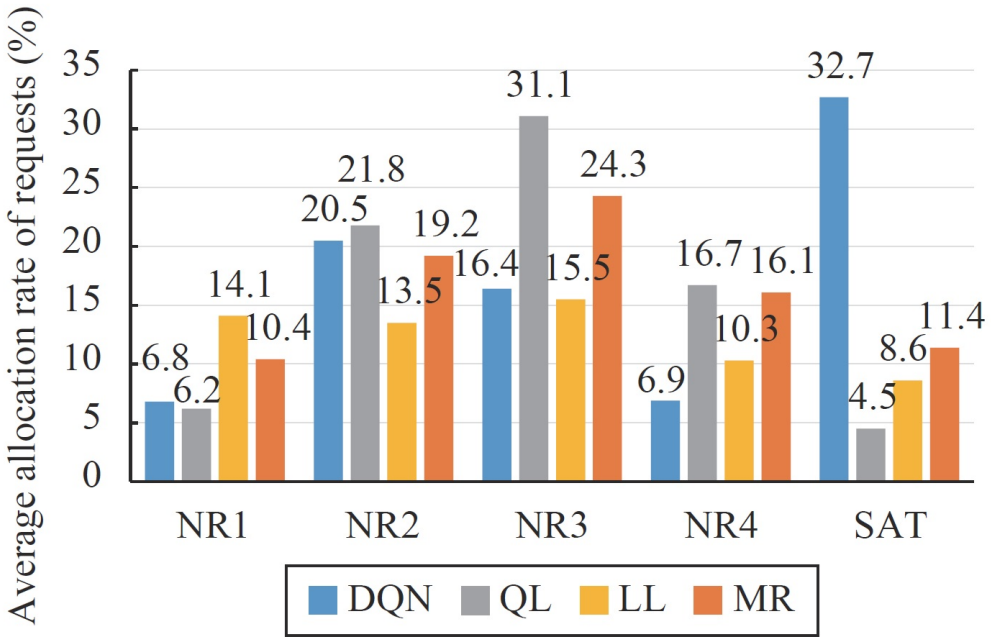


Figure 4.8. Load distribution among each AP

one or two base stations being exploited more than the others. In particular, the DQN controller relies heavily on the satellite base station to allocate incoming requests, allocating about 30% of requests to this AP. DQN is hence the only approach that fully exploits satellite resources, as the others tend to utilize mainly the NR base stations.

Fig. 4.9 represents the QoE collected by each of the controllers. The values for each controller are computed by summing the QoE gained by each request according to the QoE profiles defined in Section 4.2 and then normalized on the result obtained by the DQN controller. As expected, the Q-learning controller has similar performances with respect to the DQN one, reaching the highest QoE level. The performance gap increases when comparing a learning-based agent against the other approaches.

Finally, Fig. 4.10 represents the QoE collected by each controller in case the number of UEs is less than 100. From Fig. 4.10, it is possible to notice, as expected, that the DQN controller is able to achieve better performances compared to the competitor controllers when the number of UEs is smaller, while, if the number of UEs increases, the network is more likely saturated, making the DQN, Q-learning and MR approaches gain similar levels of QoE. In fact, the rejection rates of all approaches increase as the network overload increases. However, the DQN and the Q-learning controllers continue to prefer rejecting non-elastic traffic in favor of

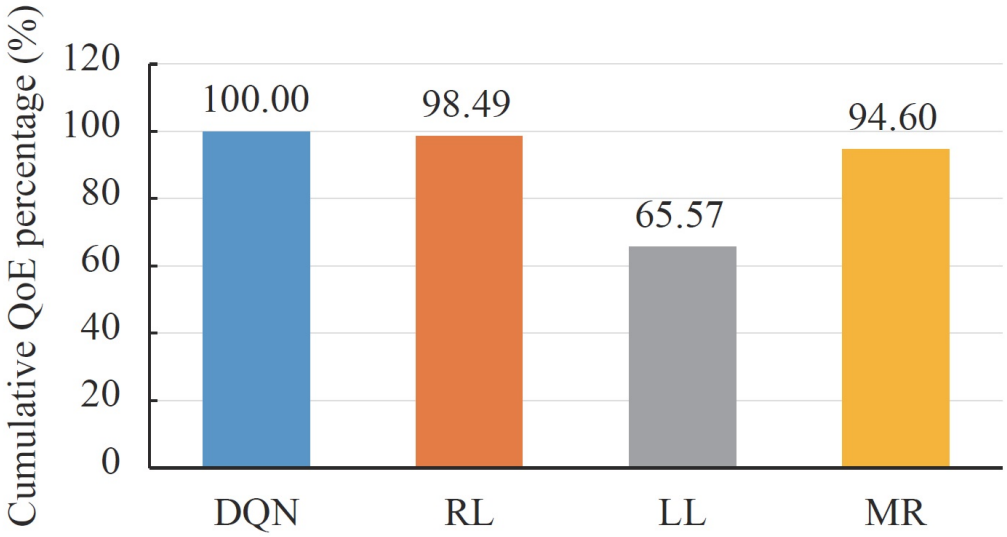


Figure 4.9. Cumulative QoE gained by each of the controllers with respect to DQN controller

multi-codec and elastic service classes.

As shown in Fig. 4.11, the cumulative QoE of the DQN approach, normalized with the QoE of the case with 100 UEs, still increases (in a sub-linear way) as the number of UEs increases. This is because, even if the network is going towards saturation, the controller is still able to allocate some more UEs w.r.t. the cases with fewer UEs.

4.6 Conclusion

The work proposed a network controller based on deep reinforcement learning to enable the integration of satellite systems into 5G heterogeneous networks. The proposed controller dealt with the problem of network selection by formulating it as a Markov decision process and was compared to several standard benchmark algorithms. The proposed solution proved to be able to cope with large-scale scenarios involving 100 different UEs.

For validation purposes, an open-source network simulator [51] that realistically captures the network resource usage of different radio technologies, including satellite connections, has been developed.

Overall, the proposed controller improved the performance of the network, increasing the connection-flow acceptance rate and providing better resource management compared to the other methods tested.

Future works are related to the introduction of other unmodeled complexities

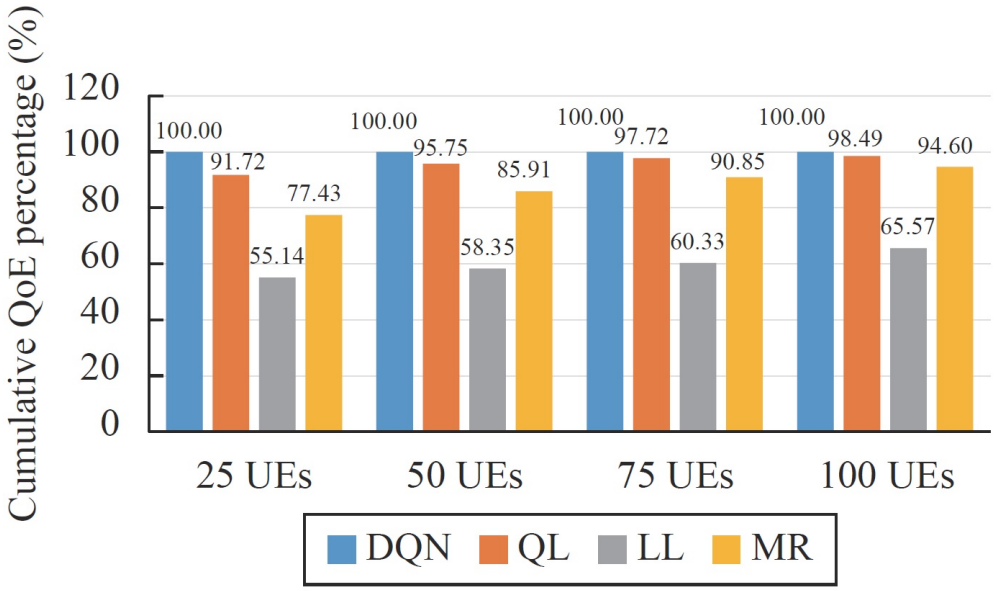


Figure 4.10. Cumulative QoE with different numbers of the user equipment, normalized on the corresponding DQN performance

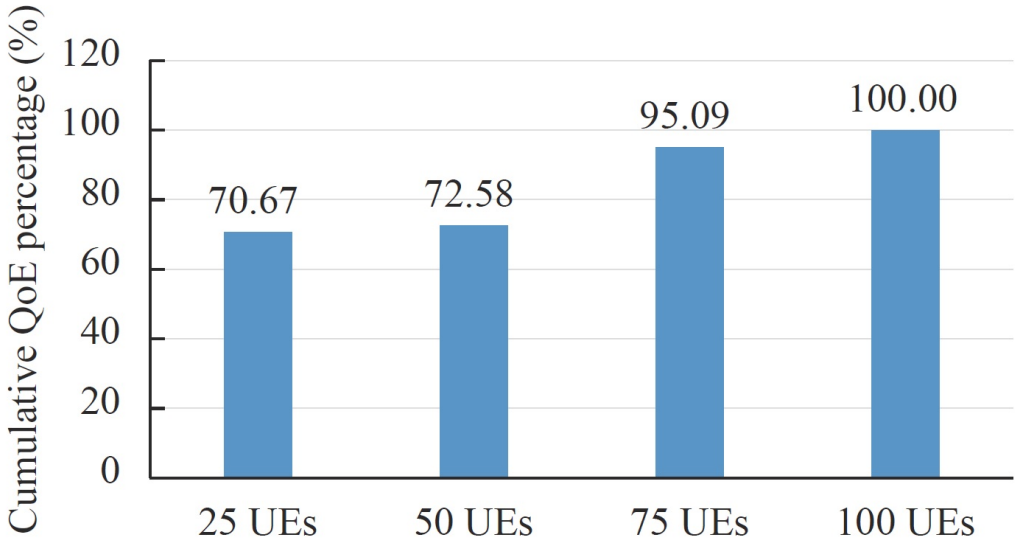


Figure 4.11. Cumulative QoE percentage for DQN approach with different numbers of user equipment, normalized on the DQN performance with 100 UEs

in the simulator, such as user and access point mobility. Actor-critic algorithms [89] will also be explored to enable the split of QoS-flows and multi-connectivity, allocating a single flow over different access points at the same time.

Chapter 5

Vision on 6G cellular networks exploiting 3D-connectivity and multi-RAT

This work gives a vision of 6G cellular networks architecture, which should comply with new KPI requirements of 6G, by considering new opportunities given by UAVs, HAPSs and LAPSs, that may act both as radio signal relay (i.e., just as signal amplifiers to augment the radio coverage of fixed base stations) or even as mobile base stations, that may cover remote areas or may change their position according to local network saturations at the radio access level. Moreover, by using UAVs and HAPSs/LAPSs it is possible to augment the current 2D radio coverage offered by 4G/5G to a 3D connectivity (and in particular a hierarchical 3D connectivity made up by Geostationary Earth Orbit (GEO)/Low Earth Orbit (LEO) satellites, HAPSs/LAPSs, UAVs, and terrestrial APs).

These new ideas have been tested against the multi-RAT radio access network simulator described in previous chapters to demonstrate the feasibility of the proposed approach, which is strongly encouraged by both standardization bodies (e.g., 3GPP) and literature on this research field.

5.1 Introduction

Coverage is a critical key performance indicator (KPI) when deploying wireless networks. Up to 4G networks, most of the efforts have been focused on increasing link capacity while ensuring sufficient coverage in the two-dimensional (2D) plane. 5G with its multi-dimensional requirements adds more stringent constraints for, e.g., mission-critical services with requirements on low latency and high reliability

(URLLC), massive amount of devices (eMMB), range extensions, and on Operational Costs (OPEX) of the communication infrastructure. 5G allows exploiting new opportunities by sharing the underlying infrastructure among isolated and self-contained networks through the concept of network slicing. Moreover, starting from the 4G-LTE all-IP architecture, the network offers communication coverage and integration of cloud support. Nevertheless, services were offered on a 2D-plane, and cloud services were conceived for data fetching/storage (over significant distances between data centers and the users connected) and to provide services (e.g., social media or instant messaging) to mobile internet users. Newly emerging 5G services ask for solutions going beyond this framework, including ubiquitous coverage/capacity availability and service scalability adapted to new use cases, application scenarios, and traffic conditions, which would be a tough challenge for the one-network-fits-all 4G-LTE architecture.

While the availability of good terrestrial coverage has become common in densely populated areas and regions, the underlying business model based on a flat fee per user does not scale well in sparsely populated regions or areas with difficult orography (e.g., islands, rugged mountainous terrain or off-shore). Worldwide mobile network operators provide usually no, poor, or at best low-quality connectivity in those cases, while the potentials of these regions can only be fully exploited when providing connectivity for the digitization of their economic activities, e.g., smart agriculture or mining. Relevant KPIs in this context are ubiquitous connectivity, scalability, and affordability. Moving from 2D to 3D-coverage is an enabling solution, the third dimension resulting from placing network elements up into the sky and space.

5.1.1 Cooperation among Terrestrial and aerial/spatial networks

Many recent research projects investigate the cooperation between terrestrial and LEO satellite networks for 5G NR. Within the 3GPP framework, use cases and associated system requirements for the satellite integration in the 5G eco-system are specified and continuously updated by the working group SA1 in [94]. Standardization impact to the NR specification are studied in [95] [67], considering Non-Terrestial Network (NTN) as an integral part of NR. The successful outcome from these studies led to normative work in Release 16 specifying extensions to NR for UAVs [96], HAPS and satellites based on well-defined channel models, deployment scenarios, and system parameters. Likewise, future 3GPP releases will focus on solutions for RAN protocols and architecture.

The 5G AgiLe and fLexible integration of SaTellite And cellulaR (5G-ALLSTAR) H2020 project [97] investigates multi-connectivity technologies that integrate cellu-

lar and satellite networks to provide reliable, ubiquitous, and broadband services for 5G NR. This is the follow-up of the first investigations on terrestrial with non-terrestrial communication integration in the 5G CHAMPION project [98]. Multi-connectivity requires significant innovations in the integration of millimeter-wave (mmWave) 5G-NR-based cellular system with a NR-based satellite system, as well as the adoption of spectrum sharing and interference management techniques. The H2020 project VITAL addresses the terrestrial and satellite networks by enabling Software Defined Networking (SDN) based, federated resources management in hybrid satellite-terrestrial networks. The H2020 project SANSA aims at enhancing the capacity and resilience of wireless backhauling through the cooperation of terrestrial-satellite networks. In these projects load balancing, efficient spectrum usage, improved coverage, and link performance are sought.

HAPS [99] are unmanned aircrafts positioned above 20 km altitude, in the stratosphere, for very long-duration flights counted in years. Since the 1990s, many initiatives have been launched worldwide to explore potential applications, including telecommunications services. HAPS offers wide area coverage with advantages compared to satellites in terms of cost, ease of deployment/reuse and large payloads, lower delays, and signal attenuation. Recently, Google's Loon project has been deploying a network of high-altitude solar-powered balloons that move using wind jets. They embark regenerative payloads and inter-balloon communication links and their network coexists with terrestrial LTE networks providing service to rural mobile broadband users in areas where terrestrial coverage does not exist. Some other operational HAPS with higher payload capacity (like Thales-Alenia's Stratobus dirigible) are expected by 2021-2023.

At a lower altitude, drones are UAV that have the capacity of dynamically providing radio on-demand coverage exploiting embarked light base stations [100, 101]. UAV's and HAPS have received considerable attention [102] in terms of data traffic management [103], network coverage enhancement [104, 105], improving quality of service [106, 107], propulsion and transmission powers [108], latency minimization [109], or exploitation of network access [110].

5.1.2 Hierarchical BS fleets for providing computing and intelligence functionalities

Several works in the literature such as [108, 109, 111] and [105] propose different architectures and mathematical models for 3D networks comprising multiple UAVs, focusing in particular on the communication aspects such as data backhauling and reduced latency, whereas the architecture that will be presented in this work focuses on joint communication, computation, and caching capabilities, which are

considered as components of a single 3D system. Extending the use of UAV's to provide not only radio access, but also mobile computing functionalities is actually considered a promising paradigm to satisfy *on demand* communication and computation requests, and deliver *context-aware* cloud services to mobile users. The first attempt to host cloudlet processors on the UAV, is addressed in [112]. The target is to minimize the energy at the UEs while optimizing transmission data rates, jointly with the UAV's trajectory under latency constraints. In [113] the authors include an edge computing scenario with aerial platforms and heterogeneous IoT devices. A dynamic formulation appears in [114], where computation offloading is handled with Stochastic Optimization tools having energy consumption as a goal while optimizing the trajectory of UAVs. In [115], a dynamic online strategy jointly allocates communication and computation resources, while selecting the vehicle's altitude, with the aim of minimizing the system energy and satisfying latency constraints. The work in [116] introduces Fog Computing into a swarm of drones, with the aim of handling computation-intensive offloading of tasks. In [117] the sum power consumption is minimized for a multi UAV-enabled Multi-Access Edge Computing (MEC) network.

According to this novel vision, research is needed to investigate solutions in realistic scenarios in which 3D services are supported by a hierarchical fleet BS embarked in UAVs, HAPS, and LEO satellites, each having its own specific features in terms of payload, flight autonomy, mobility, service coverage time, altitude, revisit time, computation, storage, coverage area, link power budget, etc. In such a challenging context, ensuring end-to-end service continuity for ground users or users moving in the 3D space entails rethinking the mobility management mechanisms incorporating proactive allocation of the content, smart proactive caching of recurrent computational results [118], instantiation of virtual machines, interference management and joint handover between radio access points and mobile edge computing hosts. This will require the development of a fast live migration of the light virtual machines, e.g., dockers, and an extension of Network Function Virtualization (NFV)/SDN orchestration schemes to make them more inclusive with respect to the types of network nodes and also faster to support the mobility of both user terminals and network elements.

Artificial intelligence (AI) can help to solve these issues. The last decade has witnessed rapid progress in the field, driven by the increased computational capacity of computers and the wide availability of data sets. In end-to-end communications, ETSI Experiential Network Intelligence (ENI) group investigates how 5G networks can leverage AI to achieve autonomous, and thus cost-effective, slice management and orchestration. Inspired by the success of AI in solving complicated control

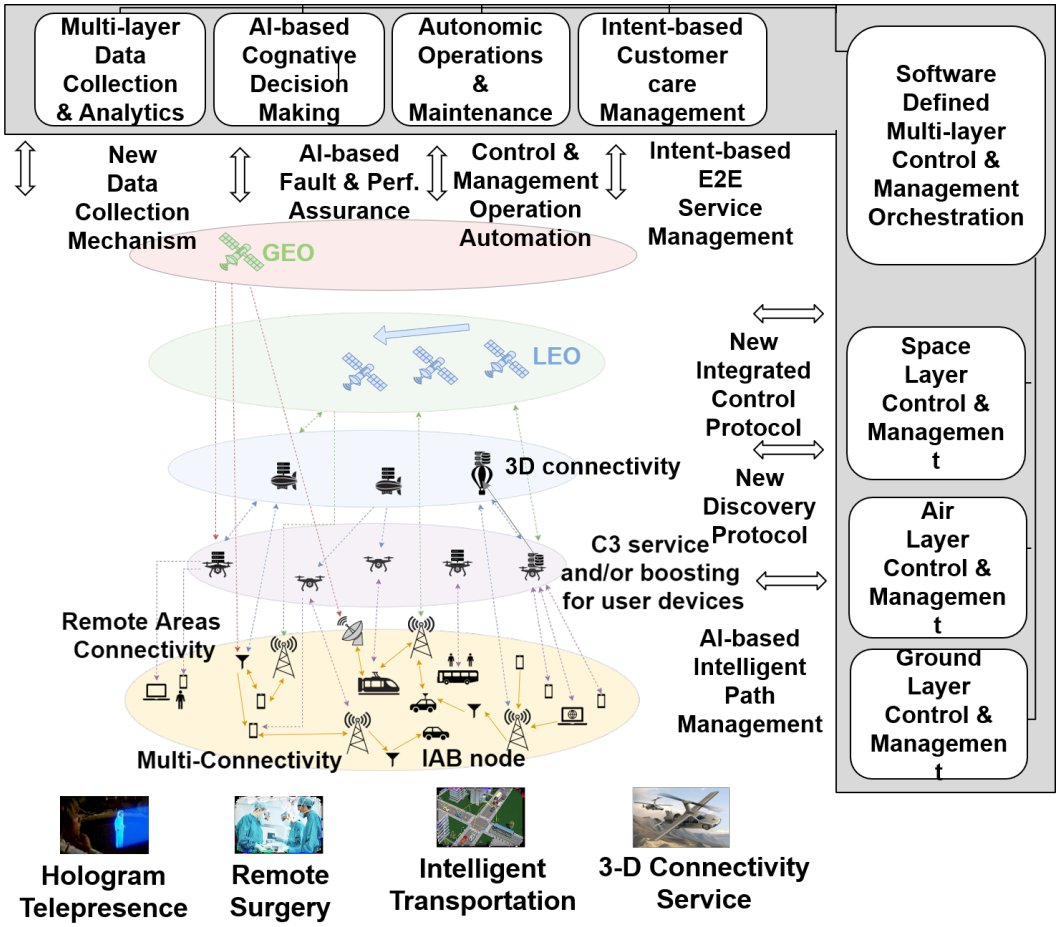


Figure 5.1. Hierarchical 3D Network System Architecture

and decision-making problems, distributed AI approaches are enablers to allow the network functionalities to learn about the network and take the best decisions accordingly.

Looking into the predictions of new technologies and services for the next decade, there is a clear need to move beyond 2D service coverage to truly 3D native services. 6G networks will enable end users moving in the 3D space to perceive a surrounding "huge artificial brain" offering virtually zero-latency services, unlimited storage, and immense cognition capabilities [119]. To match this vision, future 6G networks will seamlessly incorporate terrestrial, aerial, and satellite radio access points to teleport *on demand* cloud functionalities *where* and *when* the intelligence support is needed in the 3D space.

5.2 System Architecture

Hierarchical 3D networks with multiple and heterogeneous types of flying layers are key to provide enhanced 2D services [120] and to 3D native services including connectivity and intelligence support. Fig. 5.1 illustrates a high-level architecture of the hierarchical 3D networks unifying diverse 3D network nodes distributed over ground and flying layers. Different types of aerial nodes such as UAV, and more generally LAPS, HAPS and LEO/GEO satellites are located on different flying layers. Since aerial nodes can be equipped with on-board computation/storage capabilities, they can serve as 3D Base Stations, alone or in swarm formation, or 3D relay, which comprises an integrated access and backhaul (IAB) based hierarchical 3D networks. Although the current IAB standardization in 3GPP focuses on the ground network, in 6G, it will be extended to air and space networks as well as their integrated network.

Low and high-altitude platforms have several key potential applications in wireless communication systems due to their high mobility, flexibility, adaptive coverage capacity, and low cost. Equipped with MEC servers, these aerial vehicles can provide opportunities for ground mobile users to offload heavy computation tasks, and then after computation, the mobile users can download the computation results via reliable, cost-effective wireless communication links, as well as download each kind of needed content. The proposed integrated 3D architecture enables the boosting of Command, Control, and Communications (C3) performance in areas with existing infrastructure, and provides a network infrastructure for C3 services in areas without coverage. 3D connectivity services exploit the flexibility to accommodate a wide spectrum of applications ranging from two-way telecommunications (e.g., interactive 3D video, 3D intelligent services), to remote sensing, pollution monitoring, meteorological measurements, real-time earth monitoring, traffic monitoring, control, land management, and agriculture.

Connectivity of UEs, BS, and relays placed on different flying layers might lead to much larger connectivity handover instances, mainly due to the difference in heights and speeds of nodes belonging to different flying layers. A today's open axe of research for offering 3D service continuity and handover instance minimization is the cross-layer harmonization of selected UAV, HAPS and satellite placement and, the optimization of flying trajectories. In addition, already in 2005, NASA proposed the vision of a *Space Wide Web network*, where messages can hop between intermediate nodes to reach close planets having each orbiter, rover, space-borne telescope, and any other skyward-launched device working as a node of the 3D network [121]. At the horizon of 2030, with 6G, 3GPP standards will not go so far. Nevertheless, an *Sky Wide Web or Internet of Sky* might be already possibly interconnected with 6G

non-terrestrial 3D networks. In this hierarchical 3D network, 3D multi-connectivity will allow UEs to establish multiple different traffic links with 3D network nodes, thereby significantly improving the service performance of the UE with a dynamic load balancing scheme over the established links. This, however, requires specifically designed highly efficient, and intelligent control and management of 3D layers.

In our view, future 3D system architectures will apply network slicing not only across terrestrial nodes as it is designed for 5G networks but also across non-terrestrial nodes to facilitate different use cases and services provisioned in 3D space. The proposed architecture shall then be able to offer services that go beyond pure connectivity and at the same time offer deep customization of connectivity and intelligent mobile network services at different granularity levels, spacing from dedicated slices per data of users, to slices per individual and groups of users and to slices dedicated to 3D applications and 3D sub-networks. This will require a new adaptable *Midhaul* for an era of services that goes well beyond the services of today's 5G networks and the ones envisaged in most studies that focus on integrating UAVs into 5G networks.

AI-based approaches for network control also play a pivotal role in intelligent routing selection across 3D network layers and load balancing. For this reason, the proposed architecture shall be able to provide network intelligence capabilities at various levels and also entail device-to-device (D2D) communication, which may be enhanced by the addition of the new dimension and moving network equipment such as UAVs. In 3GPP, the first version of NR sidelink for the support of advanced V2X applications has been developed in Rel-16, and in 3GPP Rel-17, sidelink-based relaying functionality will be studied on top of the Rel-16 sidelink specification for the purpose of sidelink/network coverage extension and power efficiency improvement. In 6G, device-to-device (D2D) communications will be further extended to 3D layers, which could have great potential in facilitating a wider range of applications and services such as the next-generation intelligent transportation services.

5.2.1 Terrestrial and Non-Terrestrial Networks

Since many years, researchers have been advocating solutions for a converged integration of terrestrial and satellite communication into handheld devices and mission control centers [100], which ranged from Over-The-Top multi-RAT approaches [122] to fully unified air-interfaces [123]. Conducted field trials with e.g. adapted 4G-LTE system parameters [124] proved feasibility but only recent advances in 5G-NR standardization [67] finally bring commercial impact into graspable reach.

Continuous efforts were made by the satellite community to engage and contribute to the 3GPP process, which was focused on land mobile networks for decades.

The inclusion of NTN use cases and deployment options into the 3GPP technology feature roadmap is a best practice example of how vertical industries can actively push boundaries and get vital technologies included in an evolving standard. Initial skepticism by many critics was overcome by a gradual approach, first to study the impact from NTN use cases on 5G-NR and to provide suitable channel models [67] and simulation assumptions [125] matched to the well-established 3GPP evaluation procedures and, after successful completion, continued with nominal work in Release 16 and 17.

The 3D component is new territory for network design, in particular when aspects and KPIs like e.g. coverage, capacity, reliability, interference, and mobility are to be extended and evaluated in 3D. It is expected that providing ubiquitous connectivity in 3D will require significant changes in architecture, function placement, and network node design beyond the current approaches for terrestrial 5G base stations and satellites deployed or launched today. One example is MEC placement in a LEO satellite network to provide, e.g., a virtual private network slice for maritime or air fleet applications with low latency service requirements. MEC placement may require fundamentally new approaches to dynamic allocation of computation, caching, and communication resources on LEO nodes, including inter-node connections in space and between space and ground. Thus, the standardization impact goes beyond 3GPP and will touch standardization groups in charge of MEC, SDN, Fronthaul, and other interfaced involved to build a fully functional communication network.

Latest satellite network deployments will increasingly populate the LEO at 500-1000 km altitude. Various corporations and consortia e.g. Amazon's Project Kuiper, OneWeb, Telesat, or Elon Musk's Starlink plan to provide internet services from 2021, with current deployments ranging from a few dozen to hundreds of satellites, some targeting more than 10.000 in the future. Bend-pipe satellites keep flexibility for air-interface selection, e.g. DVB S2X or LEO adapted variants of LTE, NB-IoT or 5G-NR. On the other hand, onboard signal processing help to reduce e2e latency, in space packet routing and MEC. This will further open the existing satellite ecosystem toward interoperability and scalability in market size on-chip, module, device, and signal processing platform manufacturers, system, and service provider levels. Since satellite networks provide coverage footprints beyond the boundaries of countries or continents, infrastructure and spectrum sharing will become increasingly important for cost and spectrum-efficient deployment and operation of terrestrial and NTN including VLEO, cube sats, and HAPSs.

5.2.2 UAVs as Radio Access Network

Instead of connecting UAV with an existing RAN for control and communication from on-board equipment and/or sensors, UAVs may serve as deployed base stations or provide relaying functionality between devices and base stations of the RAN. Prominent examples of flying base stations for emergency networks or networks in remote areas are Google's Loon project [126] or unmanned airplanes supporting a larger coverage area while moving above the targeted coverage area at an altitude of 10-20km. Alternative approaches consider drones at very low altitudes of 10-50m to provide extra capacity at hotspots [127] e.g. during large public gatherings. Considering non-stationary positions and a varying number of infrastructure components e.g. UAV mounted base stations to provide an extended cellular coverage, such dynamic topology with all its flexibility comes at the cost of additional features at the RAN side to be standardized. So far, moving base stations and/or networks have been tested and deployed in relative isolation, using proprietary interfaces in particular for backhaul and interlinking between several base stations using line-of-sight links over potentially hundreds of kilometers with mmWave or laser technology. For a wider acceptance in co-existence with terrestrial RAN deployments, further studies have to be made beyond the ongoing discussions for 5G-NR.

5.3 From 5G NR 2D Enhanced Services to 6G 3D Services

In this section, we focus on the coverage extension from 2D to 3D. First, we analyze the benefits of the inclusion of aerial devices in terms of connectivity. Then, we move to the service level, highlighting the need for moving to a holistic approach that looks at communication, computation, and caching as components of a single system. We distinguish between 2D services involving devices on the ground potentially benefiting from 3D connectivity, and 3D services involving devices on the ground and in the air. We discuss these aspects also from the point of mobility management, handover and live migration of virtual machines and, control of C3 services. Finally, we focus on the importance of including artificial intelligence mechanisms to design a cost-effective system, able to incorporate proactive mechanisms and learn from online observations.

5.3.1 3D Connectivity

Including UAV-based devices in wireless communication networks provides a cost-effective solution to improve connectivity, especially if the data traffic is non-homogeneous

and non-stationary, i.e. it is expected to be highly varying across space and/or time. In such a case, a fixed infrastructure is highly ineffective, for both CAPEX and OPEX expenditures. As in many real-world situations, the opportunities offered by the UAV-based devices come along with several challenges. To highlight these challenges, it is first necessary to classify the role that UAV-based devices can play in the network. The UAV devices may act as flying bases stations (UAV-BS), as flying user equipments (UAV-UE), and as flying relays (UAV-R).

The UAV-BS brings connectivity to mobile devices on demand. The challenges come from the nature of the UAVs. HAPS have sufficient energy availability and are typically supported by solar-powered batteries so that they are able to support continuous coverage for a long time. They can typically be used to support long-term coverage purposes. On the contrary, the support of coverage in highly time-varying situations is better handled with LAPS, which can be flown on the spot of interest on demand. However, LAPS have very limited energy availability and can hover over a given area for a relatively short period of time. This means that flight and energy constraints should be taken into account in allocating the resources of the network. The limited weight payload that can be placed on a LAPS suggests the use of higher frequency bands, e.g., mmWave bands, to use smaller size antennas and to achieve better spectral efficiency. However, the use of mmWave links faces the problems of link attenuation, in case of rain, and blocking effects. To reduce link attenuation it is necessary to limit the coverage area, possibly flying at the lowest permitted altitude. However, flying at low altitudes increase the probability of blocking. Momentary blocking severely impacts also the reliability of high-capacity radio links and the MEC-assisted service continuity [128]. In 2D networks, the detrimental effects of blocking are reduced using multiple RATs or multiple interfaces of the same RAT. The adoption of 3D connectivity to enhance the performance of 2D networks brings interesting new opportunities and challenges to be solved [102] for the next decade in 5G and beyond 5G networks. The selection of the altitude plays a key role. Intuitively, the higher the altitude, the larger the coverage offered by the platform and the lower the chance of suffering shadowing effects, due to favorable Line-of-Sight propagation conditions. However, high altitudes also imply larger distances and then higher attenuation. The altitude has then to be carefully selected, also depending on the distribution of the UE's [115].

To enable the several applications of UAV-assisted services, the UAV's needs to communicate with the existing wireless network, either cellular or Wi-Fi. In such a scenario, the UAV's acts as the UEs of the wireless networks. The UAVs can also act as UEs in applications such as delivery drones, real-time surveillance, and UAV-assisted transportation networks. In this case, we have a really *3D service*

exploiting a 3D network architecture. An interesting example of 3D service is a virtual reality scenario, where the UAV flies over a location of interest carrying a 360-degree camera, which is controlled by the end-user equipment to select the view angle specifying which part of the video needs to be transmitted with sufficient quality. To handle these 3D services properly, it is necessary to handle the interference that UAV-UEs can bring to the terrestrial UEs. Typically, the antennas of current terrestrial BSs are designed to handle an essentially 2D coverage problem, so that the radiation patterns are usually attenuated at high elevation angles. As a consequence, the communication between UAV-UEs and conventional BSs typically relies on sidelobes or back lobes of the BS antenna. Clearly, a better design involves a proper redesign of 3D beamforming at the BS, able to track the UAV-UEs. In [129], 3GPP specifies new BS antenna design and cellular communication techniques for UAV coverage up to the maximum altitude of 300 m. Most likely, it will be necessary for the BS to distinguish between the aerial and terrestrial UEs, to handle them separately.

Finally, UAV devices can act as relays (UAV-R) to provide backhaul the connectivity between the terrestrial/aerial UEs and the terrestrial/aerial BSs. In such a case, a key challenge is to devise effective cooperative communication strategies that take into account the mobility of aerial devices. In principle, one could make near distance UAV devices operate as a huge virtual antenna, with also the possibility to adapt the shape of the constellation by making the UAVs move as needed, provided that the resulting synchronization problems are properly handled. In general, using the UAVs as wireless relays can boost (on demand) the link quality between the ground BSs and the terrestrial UEs, but it raises also an interference issue towards the neighboring BSs that should be handled consequently.

5.4 Dynamic Resource Management for 3D Connectivity

5.4.1 Multi-RAT Connection Admission Control

Mobile nodes acting as embarked relays or BSs in UAV can handle sporadic congestion events in the radio access network occurring in specific areas, by offloading communication and MEC traffic from the fixed terrestrial links (from the protocol stack viewpoint).

This scenario impacts on connection admission control (CAC) algorithms which now have to consider not only UE mobility, but also the mobility of the AP. To show the capacities of this new scenario, a simulation study made through an ad-hoc open-source 5G network simulator [51] is presented, in which the multi-RAT

Numerology (μ)	Subcarrier spacing (KHz)	Timeslot length (ms)
0	15	1
1	30	0.5
2	60	0.25
3	120	0.125
4	240	0.0625

Table 5.1. Subcarrier bandwidth and timeslot length for the various numerologies (table 4.3.2-1 and 4.2-1 from [3])

simulation environment is composed of one type of fixed RAT, provided by a satellite cell, and two types of mobile RATs: a 5G NR mobile relay node and a 5G NR mobile BS.

Besides the fact that the admission control must be capable of handling mobile APs in a 3D environment, the key to the success is the readiness for intervention. Based on traffic and mobility data, AI algorithms are needed to foresee when and where traffic peaks are going to occur for UAV to reach the identified area timely.

5.4.2 Resource allocation

The resource allocation process differs depending on the RAT. For 5G NR RATs, we consider the Type 1 frame structure defined by 5G NR standards, which uses Frequency Division Duplexing (FDD) for both downlink and uplink, with minimum allocation unit defined as PRB. A PRB, as mentioned before, is composed by 12 frequency subcarriers, whose bandwidth depends on the numerology μ [3], as represented in table 5.1. The NR frame structure is composed of 10 ms frames, in turn, composed by a number of time-slots depending again on the numerology μ. Each RB is made by 12 or 14 Orthogonal Frequency Division Multiplexing (OFDM) symbols (respectively with extended and normal Cyclic Prefix).

A different number of PRBs is defined for each channel bandwidth, depending on the frequency band used (either FR1 [53] or FR2 [54]), and on the subcarrier bandwidth.

Once the UE requests a bitrate to an NR RAT, the AP computes the required number of PRBs. Firstly, the AP computes the SINR for the UE. The inter-AP interference is estimated as in equation (4.12). The data rate which can be transmitted by a single PRB is computed using Best Modulation Coding Scheme (MCS) [130] with the Shannon formula: $r_{ij} = B_{\mathrm{RB}} \log_2(1 + \mathrm{SINR_{ij}})$, where B_{RB} is the bandwidth of a single PRB (i.e. $12 \cdot 15 \cdot 2^{\mu}$ kHz). Finally, the number of PRB needed to satisfy UE requirements is

$$n_{ij} = \left\lceil \frac{R_i}{r_{ij}} \right\rceil \tag{5.1}$$

If the relative position between the UE and the AP changes, the SINR changes, and the number of allocated PRB to the UE has to be updated.

As detailed in section 4.4.5, the simulated satellite RAT uses TDM for concurrent UE access. Given a time frame, a certain number n_{tot} of symbols are available to the UE transmissions. Moreover, for each time frame, part of the symbols is used for synchronization purposes (n_{sync}), each communication contains a header (of length n_{head}) and there is a guard space of n_{space} symbols between each communication to avoid intra-RAT interference. The simulated satellite is an Inmarsat implementation, with $n_{\text{tot}} = 120832$ symbols, equivalent to a time frame of 2ms, $n_{\text{sync}} = 288$ symbols, with 2 synchronization messages inside the time frame, $n_{\text{head}} = 280$ symbols for each UE communication, $n_{\text{space}} = 64$ symbols, $n_{\text{slice}} = 39104$ symbols, that are about a third of the total symbols, $n_{\text{block}} = 64$ symbols. [55] The data rate that can be obtained by a single block is obtained from the Shannon formula, and the number of blocks to be allocated to satisfy the UE request R_i is computed as in equation (5.1). The actual integer number of symbols occupied by an UE are equal to $\bar{n}_{ij} = n_{\text{head}} + n_{ij} + n_{\text{space}}$.

5.4.3 Simulation Setup

The environment is represented by a 4 Km $\times$ 4 Km grid containing 50 UE, a single satellite AP, and two mobile 5G NR APs. Each UE requires a bitrate of 10 Mbps, its starting position is randomly computed and it moves on a straight line with a random direction at a speed of 10 m/s. We also consider the service is interrupted if the bitrate falls below 5 Mbps. The satellite AP is geostationary and uses a carrier frequency of 28.4GHz with 220 MHz bandwidth [57]. Its antenna Equivalent Isotropic Radiated Power (EIRP) is 62 dBw [56]. The path loss considers both the FSPL and the atmospheric loss (0.1 dB) and the user terminal G/T (-9.7 dB/K). The mobile 5G NR APs transmit a power of 15W, have an antenna gain of 15dB, a feeder loss of 1 dB and at 800 MHz carrier frequency with a bandwidth of 100 MHz and with numerology $\mu = 2$.

The connection procedure consists of the following steps: (i) the UE measures the receiving power of the APs within its range; (ii) the UE chooses the AP to be connected to according to the received power with a User-Centric, RAN-Controlled or RAN-Assisted approach; (iii) upon communications with the UE, the AP allocates the resources based on the SINR in a best-effort basis. Due to the dynamicity of traffic and network elements, connection updates are required following the same

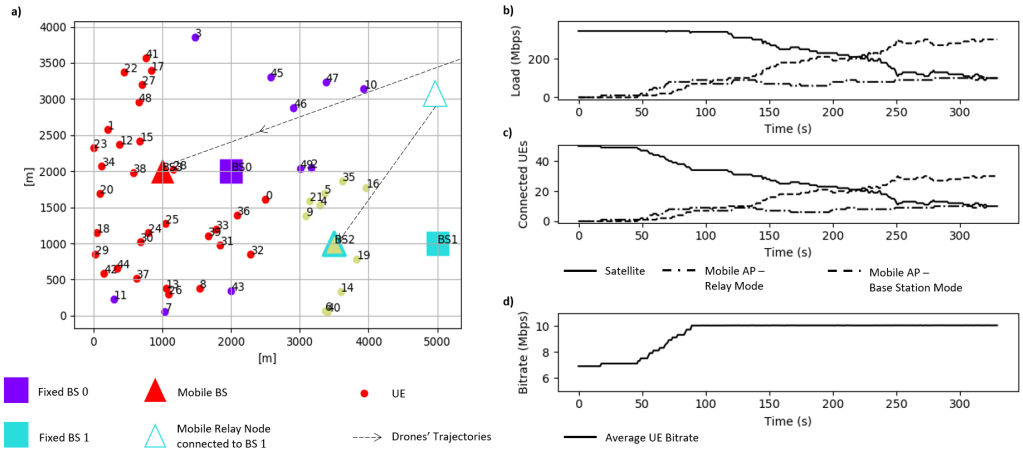


Figure 5.2. Simulation scenario 1: a) Final position of APs and UEs; b) load of the satellite AP and of the mobile APs; c) number of UE connected to the satellite AP and to the mobile APs; d) average transmission bitrate of the UEs.

procedure. The measured received power depends on the characteristics of the antenna of the generic AP j and on the path loss from the antenna to the UE i, i.e., $P_{ij} = P_j G_j L_j L_{ij}$, where P_j and G_j are the antenna power and gain, and L_j and L_{ij} represent the losses at the antenna side and the path loss between UE i and AP j, respectively. The received power depends, via the path loss L_{ij}, on the relative positions of UE i and AP j. The simulated path loss model of the satellite RAT is the free space path loss, whereas for terrestrial RATs (5G NR) we chose the COST-HATA [93] path loss model. If a UE measures a receiving power lower than a threshold $P_{\min}$ for a certain AP, then the AP is considered not visible by the UE.

5.4.4 Simulation Results

The simulation scenario shows how the use of mobile nodes can solve the congestion of fixed APs and assure *service continuity*. The initial height of the two mobile 5G NR APs is 200 m and they are far from the UE positions, which in turn can only be connected to the satellite.

In the first scenario, represented in Fig. 5.2, initially, the satellite AP is congested, i.e., all of its resources are assigned to the nearby UEs but the requests cannot be all satisfactorily fulfilled. As the simulation goes on, the two mobile 5G NR APs move towards the UEs and reduce their height to 150m. As shown in Fig. 5.2d, the average bitrate assigned to the UEs is initially equal to about 7 Mbps, below the requested one, due to the congestion of the satellite AP. The mobile APs approach the UEs and reduce their height. As a result, some UEs within range start connecting to the mobile APs, performing a handover from the satellite AP to the

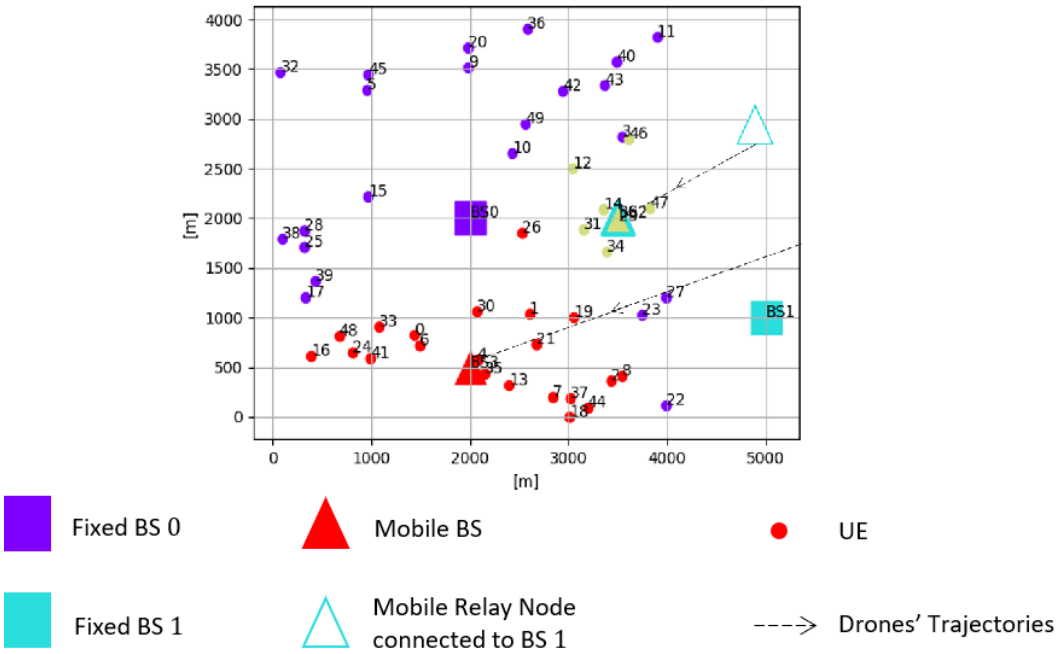


Figure 5.3. Intervention of mobile BSs in the area covered by a BS and a satellite spot. The final position of APs and UEs is shown.

mobile APs. The figure shows that, from time 90 on, the congestion of the satellite AP is solved, and all the UE requests are fulfilled.

The UEs start communicating at a random time causing the satellite AP load to increase with time. With no mobile nodes available, the satellite AP eventually becomes congested, and new UE service requests are rejected, as shown in Fig. 5.4. Moreover, some of the UE bitrates fall below 5 Mbps, causing connections' drops and service interruption, as in Fig. 5.4b. On the contrary, if UAV APs are available, as in Fig. 5.5, the UEs start connecting to the mobile APs. In this case, no UE has to interrupt the service, and service continuity is granted, maintaining the connections at 10 Mbps for the whole simulation time as well as maintaining all the UEs connected to some AP.

5.5 Conclusions

The nature of new applications in the next decade and the desire for ubiquitous availability will most likely require technologies supporting truly 3D on-demand services, rather than today's 2D service coverage.

In this view, while the integration of terrestrial with NTN for 2D service enhancement will come as a natural evolution of 5G, providing on-demand connectivity and edge intelligence to support truly 3D services will not come before 6G.

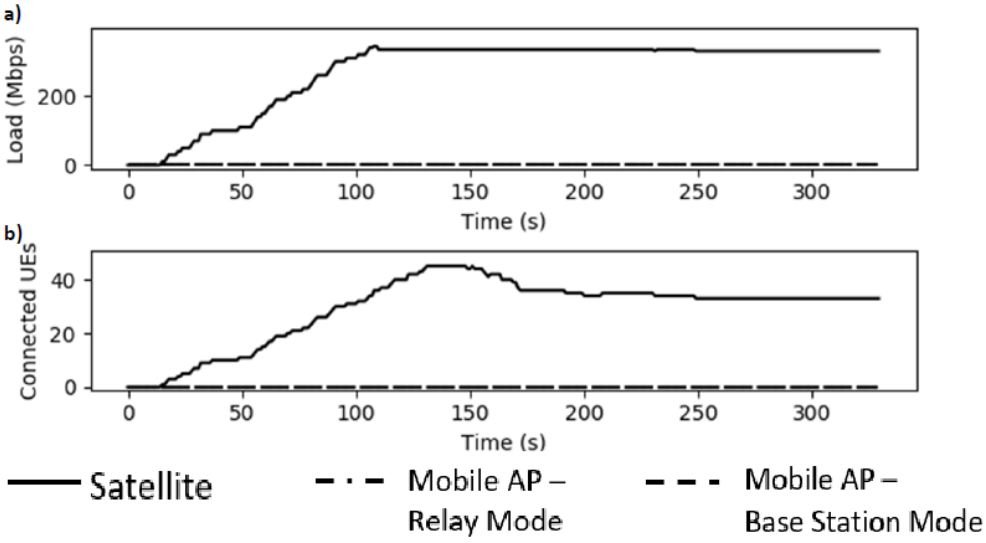


Figure 5.4. a) Load of the satellite without the mobile AP. b) Number of UEs connected to the satellite.

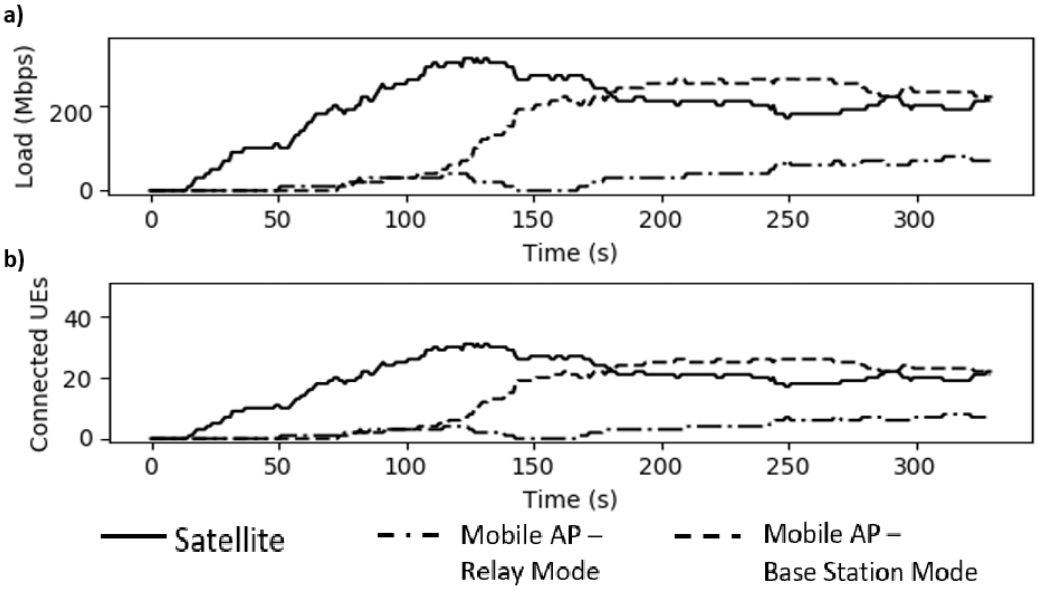


Figure 5.5. a) Load of the satellite and of the mobile APs. b) Number of UEs connected to the satellite and to the mobile APs.

In this work, an in-depth overview of a future hierarchical 3D network architecture where heterogeneous flying devices, providing different levels of mobility, coverage, and service level has been provided, enabling revolutionary new on-demand connectivity and intelligence support.

Today, NTN use cases are already being considered for new features and technology extensions in the 3GPP standard releases 16 and 17. On the roadmap for 5G NR, the integration of terrestrial and non-terrestrial networks will enable global 5G service enhancements and new functionalities. Beyond release 18 up to 6G, further extensions of 3GPP and other standardization bodies will enable advanced dynamic and meshed interconnection and relaying between NTN-nodes and MEC placement in the 3D space.

Some fundamental challenges remain open for future research. Promising innovation directions have been highlighted, like on-demand distributed C3 support, 3D-interference management, 3D-multi-link load-balancing and admission control, live intelligence handover and migration mechanisms, and AI-based joint orchestration of C4 distributed resources in the 3D-space.

Moreover, it has been shown how additional 3D nodes can effectively be exploited to dynamically handle network congestion, e.g., by using drones as on-demand mobile relay nodes or mobile BSs, to offload traffic from the fixed terrestrial links and/or to provide an extended opportunistic cellular coverage. New admission control procedures are thus needed to cope with the extended 3D network topology and, specifically, with the increased network handover occurrences implied by the dynamic 3D network.

Part II

Applicative scenario of next-generation cellular networks

Chapter 6

Fundamental concepts of Frequency Regulation in Electricity Grids

Among the many applicative scenarios of 5G and 6G cellular networks, this thesis deepened the application of next-generation cellular networks to the control and operation of electricity networks, and in particular smart grids. Indeed, next-generation cellular networks may enable completely new kinds of services on the smart grids, that were unfeasible from a technical or an economical point of view up to now with 4G LTE networks.

Among the many scenarios involving smart grids and 5G/6G, this thesis focuses on the provisioning of Frequency Regulation Services to the electricity network by exploiting Ultra-Reliable Low-Latency Communication (URLLC) slices of the next-generation cellular networks. Indeed, these kinds of services require very strict time constraints for their activation (up to 300 ms according to [2]), which are usually much faster than traditional time constants of the electricity network (5-15 minutes).

In this chapter an introduction to the problem of Frequency Regulation in Electricity Grids is detailed. In particular, the mathematical relation between the mechanical power provided to a synchronous generator and the electrical power requested to it is detailed, showing the effectiveness of aggregation of flexible loads in the participation to frequency regulation services.

6.1 Introduction

Power networks are nowadays composed of different entities that participate in the generation, transportation, and dispatching of energy to customers. In particular,

in traditional power networks, such entities had separate and precise tasks, i.e., generation companies are in charge of producing energy and providing regulation services to the network; TSOs are in charge of transporting the energy (with high-voltage lines) from the generation points to the primary substations, ensuring the safe operation of the electricity grid by requesting regulation services to the generators; Distribution System Operators (DSOs) (or very big customers) are in charge of distributing the electricity from the primary substations to the final customers (or to their plant in case of very big customers).

Nowadays, the electricity grid (and in particular, for the European context, the European Association for the Cooperation of Transmission System Operators for Electricity (ENTSO-E)) opened up the possibility for final customers to participate with a significant role in the electricity grid operation, starting from distributed generation (e.g., with small/medium-size renewable energy sources) to the contribution in the provisioning of regulation services for the electricity network.

Among the many regulation services requested by the TSO for the safe operation of the network, one of the most crucial ones is Frequency Regulation. Indeed, as detailed in Section 6.2, the frequency of the electricity network should be kept at the constant value of 50 Hz (60 Hz for the U.S.A. and a few other countries) in order to avoid damage to the synchronous generators present on the network, on the transmission and distribution lines and on the final customers' appliances.

Traditionally this kind of service is provided only by generation companies by some control loops on their synchronous generators. Moreover, with the increasing penetration of non-synchronous renewable energy sources (e.g., photovoltaic) which rely on inverters to produce energy instead of synchronous generators, the inertia of the electricity grid decreases over time, thus increasing the effects of electricity network frequency disturbances and so increasing the control effort needed to provide frequency regulation services.

In this regard, as will be detailed in the next sections, it is in principle possible to contribute to the frequency regulation services also at the customers' side, reducing the control effort requested to the synchronous generators. Indeed, the relevant regulatory bodies (ENTSO-E for the European Electricity Network, but also ARERA for the Italian Electricity Network) started to deploy pilot projects [2] for opening the ancillary service market (that includes also frequency regulation services) to other entities rather than generation companies (i.e., final customers or aggregators with a certain minimum contractual power).

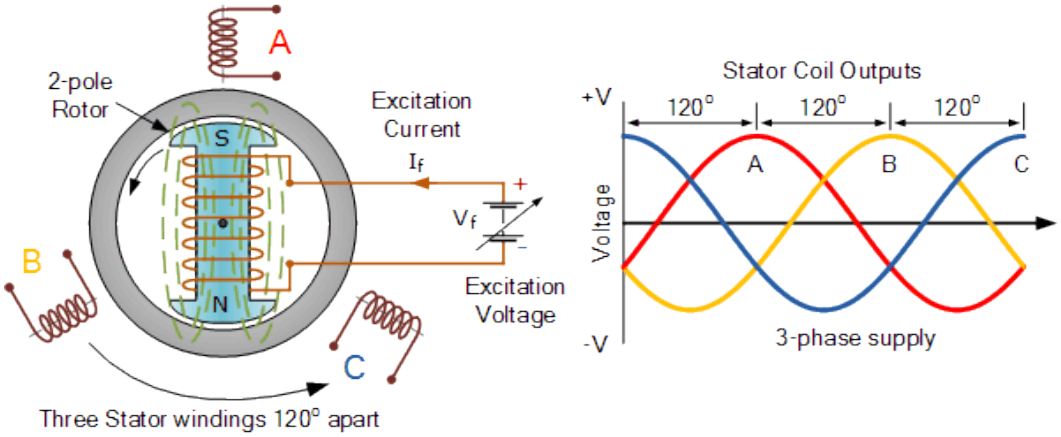


Figure 6.1. Simplified scheme of a Synchronous Generator

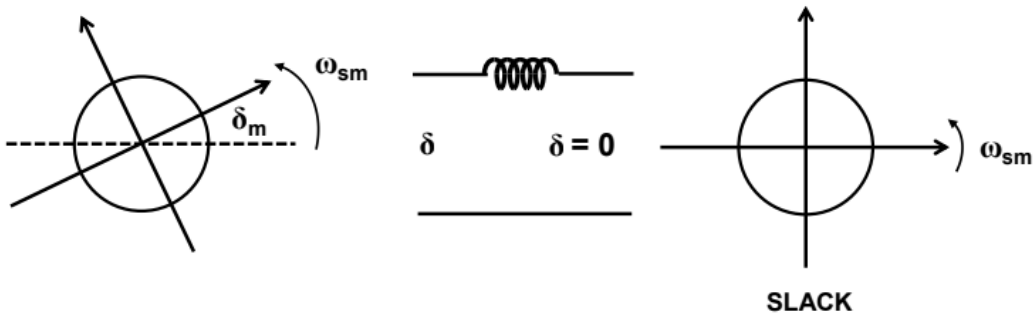


Figure 6.2. Rotor angle of a synchronous generator with respect to the slack node

6.2 Synchronous Generators and Frequency Regulation

A Synchronous Generator (depicted in Fig. 6.1) is an electrical machine that is made for generating alternate current with a frequency that is proportional to the rotation speed of the rotor. The rotor, namely an electromagnet, rotates with respect to the stator (that is fixed and contains the coils where the alternate current for the three phases is generated) thanks to the mechanical torque provided to the stator by some mechanical energy source (e.g., steam, water, wind).

Each synchronous generator has a certain rotor angle δ_m with respect to one synchronous generator that is considered as a reference (namely, the *slack*), as depicted in Fig. 6.2, and both of the generators rotate at nominal angular speed ω_{sm}.

For each generator the following equation holds:

$$J\frac{d\omega_m}{dt} + D_d\omega_m = \tau_t - \tau_e, \tag{6.1}$$

with J the moment of inertia of the turbine, ω_m the rotor shaft mechanical angular speed, D_d the mechanical damping coefficient, τ_t the mechanical torque applied to the rotor shaft, and τ_e the electromagnetic torque.

At steady state the mechanical angular speed is equal to the synchronous angular speed ω_{sm}, thus

$$\tau_m = \tau_t - D_d\omega_{sm} = \tau_e. \tag{6.2}$$

Given such a state, if an imbalance between τ_m and τ_e appears, then the rotor shaft will accelerate/decelerate in its rotational speed, with a certain ω_m defined as

$$\omega_m = \omega_{sm} + \frac{d\delta_m}{dt}. \tag{6.3}$$

If δ_m is constant, then $\omega_m = \omega_{sm}$ (i.e., no imbalance between τ_m and τ_e). Moreover, since the inertia of the power network (due to the inertia of all the generators' rotating masses connected to the electricity grid) and the frequency control loops on board, which try to keep δ_m always constant, the second term of (6.3) is almost equal to 0

$$\frac{d\delta_m}{dt} \approx 0. \tag{6.4}$$

Substituting (6.3) in (6.1), the following is obtained:

$$J\frac{d^2\delta_m}{dt^2} + D_d\left(\omega_{sm} + \frac{d\delta_m}{dt}\right) = \tau_t - \tau_e, \tag{6.5}$$

and then by adding 6.2

$$J\frac{d^2\delta_m}{dt^2} + D_d\frac{d\delta_m}{dt} = \tau_m - \tau_e. \tag{6.6}$$

By multiplying all members by ω_{sm} the following is obtained:

$$J\omega_{sm}\frac{d^2\delta_m}{dt^2} + D_d\omega_{sm}\frac{d\delta_m}{dt} = \frac{\omega_{sm}}{\omega_m}P_m - \frac{\omega_{sm}}{\omega_m}P_e. \tag{6.7}$$

Given the hypotesis of equation (6.4), i.e., $\frac{\omega_{sm}}{\omega_m} \approx 1$, (6.7) becomes

$$M_m\frac{d^2\delta_m}{dt^2} + D_m\frac{d\delta_m}{dt} = P_m - P_e. \tag{6.8}$$

From equation 6.8, which describes the behavior of the rotor shaft angle δ_m with respect to the mechanical power provided to the rotor shaft and the electrical power requested at the stator coils, it is possible to regulate the electricity network frequency (i.e., the rotor angle speed $d^2\delta_m/dt^2$ of each synchronous generator) by

modulating the mechanical power P_m at the rotor shaft (i.e., what it is traditionally done by control loops installed on-board of the generators themselves, usually by the means of a PID controller), or by modulating the active power P_e requested to the generator (i.e., the customers' electricity power request).

This means that, by modulating the electric load of the customers (i.e., by the means of an aggregate of flexible loads), it is possible to contribute to the frequency regulation services of the network. Of course, the amount of electrical power required to have an effect on the power network is quite high (due to the network inertia), so only appropriate aggregates of flexible loads could effectively participate in this type of service (according to [2] the minimum power to be guaranteed to the TSO for frequency regulation services is 1 MW).

Chapter 7

Frequency Regulation in Electricity Grids using Plug-in Electric Vehicles and 5G networks

In this chapter, an architecture for using an aggregate of PEVs as flexible loads for the provisioning of Frequency Regulation Services, by exploiting URLLC capabilities of 5G NR and the presence of MEC at the edge of the 5G radio access network, is detailed.

Indeed, as seen in the previous chapter, aggregation of (high-power) flexible loads may help in the regulation of the electricity network frequency. Charging PEVs are very appropriate for such task since their charging power goes from a few kilowatts to hundreds of kilowatts and it is possible to regulate their charging power quite easily (according to standards such as IEC 61851 for AC charging sessions and ISO 15118 for DC charging sessions). Aggregating many charging PEV it is possible to reach the minimum power to participate in frequency regulation services, but the measurement of electricity network frequency according to the requirements of [2] could make this approach unfeasible, due to the high cost of accurate frequency meters. The capabilities of 5G networks in terms of latency and the possibility of having some intelligence at the edge of the 5G radio access network (in the 5G MEC) enables the possibility to install a limited number of frequency meters (even just one, if no redundancy is needed), and then spread the frequency measurement to each PEV/charging station (CS) to provide frequency regulation services (i.e., by changing their charging power setpoints according to the measured electricity network frequency)

This work considers the requirements from [2] and is going to be implemented in the context of the European H2020 project 5G-Solutions [131], and in particular for the Smart Energy use case.

A detailed system scenario and objectives are discussed in Section 7.2 and some possible algorithms for the computation of the power-frequency (droop) curves of increasing complexity are defined in Section 7.4. Moreover, a detailed delay budget analysis has been performed to ensure the technical feasibility with respect to the requirements of [2].

7.1 Introduction

The safe and efficient operation of a power system strictly depends on two physical quantities: the frequency and the voltage level of the network [132]. The deviation of frequency and voltage from their nominal values is the effect of disequilibrium in terms of active and reactive power in the network. The so-called ancillary services are then designed for the injection/withdrawal of active and reactive power to balance the power mismatch. The evolution of the electricity network system and the spread of active components [133], make the involvement of new actors and technologies in the provisioning of ancillary services possible. The growth of electromobility in the last decade has pushed the scientific community and the power system stakeholders to develop new control strategies and concepts in order to improve and implement new paradigms to ancillary services. The high penetration of Renewable Energy Sources (RESs), with the associated transition of the power systems from synchronous-machine-based systems to inverter-dominated systems, pushed the development of the virtual inertia concept, in which also the inclusion of PEVs is expected to play an important role [134]. From this trend, several works that explore the potential and the issues of PEVs usage for frequency regulation have been studied. In [135], a review of the strategies used to include the participation of electric vehicles in frequency regulation is provided; the review considers both technical and economic aspects, showing different points of view that have to be faced to move towards the development of ancillary services that include PEVs.

Recent work in line with the activities of this work is [136], which proposes the use of PEVs for the provisioning of frequency regulation services. The work proposes some control strategies for the optimal control of PEVs' contributions, considering also the impact that such a service has on battery degradation.

This work elaborates on the design of a frequency regulation service based on the use of PEVs. The coexistence of the frequency regulation service with the smart charging one is tackled, and the advantages and challenges of the proposed scheme,

considering also relevant information and communication technologies (ICT) integration aspects, are analyzed. This work reports preliminary concepts and results established in the context of the European research project “5G Solutions" [131] where innovative use cases enabled by 5G communication technology in the field of smart energy grids are under design.

7.2 System scenario and objectives

The scenario of this work considers the evolution of the European Energy Market and the related separation of the Balance Responsible Party (BRP) and Balance Service Provider (BSP) roles [137]. The diffusion of distributed generation plants, favoured also by the European decarbonization objectives [138], and the diffusion of small-sized storage systems, together with the spread of electric mobility, bring to the need of carrying out an important revision of the role played by distribution companies. The DSOs are considering the possibility of assuming two additional roles compared to those that are traditionally under their responsibility: (i) the role of neutral facilitator for the provisioning of ancillary services made available by the BSP, that are needed for the safe operation of the overall system, and (ii) the role of the purchaser of these services.

Moreover, this evolution changes the role of the Charging Point Operator (CPO): the separation between BRP and BSP, that breaks up the physical positioning and market correlation of generation units and load plants. This separation allows the owner of energy sources to provide only ancillary services without having to care about balancing constraints. This results in the opportunity to participate in the Energy Market in an aggregated way. The separation of physical contribution and market position enables the aggregation of energy sources and, together with the possibility to sell services both at the level of distribution and transmission network, opens to the participation of new actors in the dispatching market, putting the CPO in an interesting market position.

In this context, the CPO can exploit the PEVs flexibility to create the necessary conditions for the participation in the dispatching market.

The presence of a smart charging system responsive to external signals introduces an additional factor in the context of smart charging in a load area: the possibility to have *power margins*; indeed, in [139] it is shown how it is possible to drive the aggregated charging sessions power to track a target load curve; even in case of several charging sessions running at the same time, the smart charging system proposed is able to avoid the power saturation of the load area, while ensuring the drivers' requirements [140]. The existence and the proper manage-

ment of power margins are necessary for resource qualification to the provisioning of ancillary services.

Ancillary services are historically entrusted to synchronous machines hosted by the generation units. The reasons why these services are provided by the generators are multiple; focusing on the aspects of interest for this work, there are mainly two reasons: the capability of easily controlling the power generation, thus ensuring the presence of power margins, and the unidirectionality of the distribution network power flow. The change of paradigms presented above for DSOs and Energy Market also affects the way the distribution networks are modeled and, consequently, their role in the power system. Distribution networks become hosts of active loads, storage systems, and generation units; their presence changes the distribution networks into a set of active nodes with a bidirectional flow that can potentially supply ancillary services. The vision of disseminating the provisioning of ancillary service on different portions of the network is nowadays supported and enforced by official entities and stakeholders [141].

The work presented in this chapter focuses on the Frequency Restoration Reserves with Automatic activation (aFRR). aFRR is currently entrusted to generation plants relying on synchronous machines. The extension of the aFRR service to the participation of flexible loads is subject of studies and experimentation in Europe. For example, in the Pilot Project *Fast Reserve* [2] the inclusion of Mixed Enabled Virtual Unit (UVAM) - that can be composed by PEVs [137] - in the aFRR service and their impact are investigated. This pilot project imposes specific performances on the units that are involved in the service provisioning like a specific degree of sensitivity to the frequency variations and precise reaction time requirements.

In light of the considerations made above, the reference scenario in this work is as follows: a smart charging load area is considered, and the smart charging capabilities are used to introduce power margins at a single charging session level and, consequently, in the aggregated form at load area level. The presence of charging session margins is exploited by applying a real-time frequency-based modification of the smart charging setpoints with the aim of providing the aFRR service. In this work, the Pilot Project *Fast Reserve* requirements are used to drive and validate the results.

In this context, 5G communication technologies, able to guarantee low end-to-end latency and high reliability compared to legacy technologies, together with modular virtualized network functions offered by the 5G Core Network architecture, represent enabling factors. These technologies are expected to allow the communication of frequency measurements to the charging stations within the strict

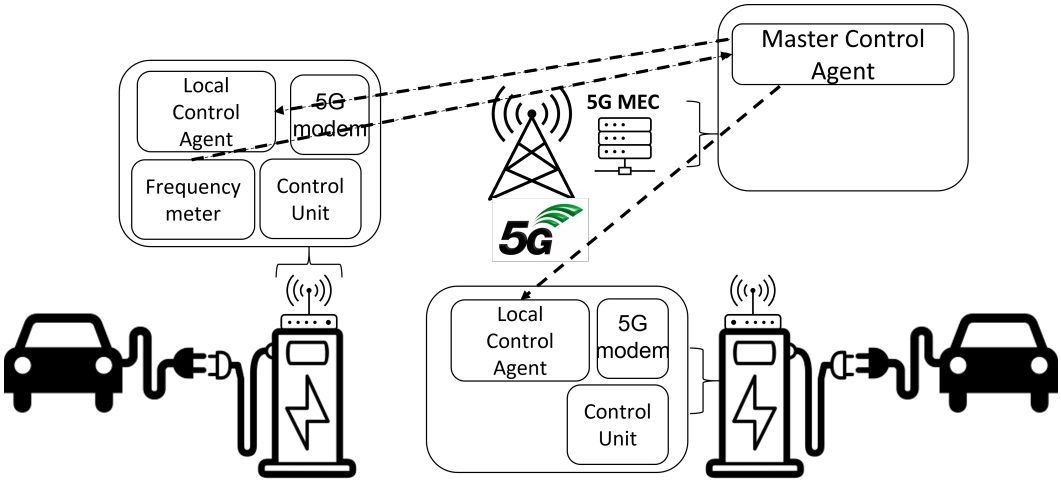


Figure 7.1. System Architecture

time constraints imposed by the frequency regulation service, instead of relying on frequency meters installed in each unit of the UVAM. Considering the size, the number, and the dispersed nature of the flexible loads needed in the aggregate, 5G technologies enable the UVAM to provide the frequency regulation service in a more cost-effective way.

The objectives of this work are: (i) to present an efficient control architecture that enables the exploitation of PEVs for the provisioning of aFRR, (ii) to provide the rationale for the integration between smart charging and the aFRR service provisioning and (iii) to discuss the impact of the telecommunication technologies on the service provisioning requirements and on the quality of the service.

7.3 System architecture

In order to match the objectives described before, a new system architecture has been designed. This system architecture represented in Fig. 7.1 shows the main components of the frequency regulation control system, together with their logical interfaces. The proposed system makes use of a single frequency meter for each Load Area (reducing a lot of the cost for the deployment of such an architecture), that can be installed inside a single Charging Station or in its neighborhood. Moreover, the proposed system exploits the novel 5G network architecture, where telecommunication operators make some computing power very near to the Radio Access Network (RAN) available for their customers. These computing resources are named MEC and enable very-low latency applications to work efficiently, since the data packets can be processed (or pre-processed) in the neighborhood of the customer requiring the service, with a substantial reduction of the end-to-end la-

tency, that is reduced almost only to the radio access link latency (that is further reduced by 5G New-Radio standards, compared to the 4G LTE one).

The proposed architecture makes use of Local Control Agents, installed inside the Charging Stations, in order to compute the control signals in response to frequency disturbances. Each of these control signals can be superimposed over its corresponding slower smart charging scheduling signal computed by a separate system, in order to enable the Load Area both to smart charging functionalities and to frequency regulation functionalities.

In the proposed architecture, the MEC hosts a Master Control Agent module that is in charge of spreading frequency measurements coming from the single frequency meter of the Load Area (that may be installed inside a Charging Station in order to make use of its 5G Modem) to all the Local Control Agents. The possibility to put the Master Control Agent inside the MEC enables a low-latency broadcast of the frequency measurements, avoiding placing a frequency meter inside each Charging Station, while still having the measurements spread with high reliability offered by 5G communication services and with a delay in line with the time requirements of the frequency regulation services, which will be better investigated in Section 7.5.

7.4 The smart charging problem and the power-frequency curve assignment

The integration between smart charging and aFRR service presented before must consider aspects related to the quality of the charging sessions, while have to guarantee the presence of power margins capable to realize a power-frequency curve that satisfies precise properties. The power-frequency curve properties can differ depending on the country. In this work, the Pilot Project Fast Reserve [2] directed by the Italian TSO *Terna* is considered as a reference for the forthcoming discussion. The requirements of the above pilot project are many, and the present work doesn't aim to address all of them. The attention is focused on the power-frequency curve shape and on the reaction time, in particular, the Pilot Project Fast Reserve requires:

- the power-frequency curve has to be symmetric (w.r.t. the frequency variations), continuous and the actuation has to be self-regulating;
- the power-frequency curve has to consider the possibility to implement a dead band;
- the fast reserve unit has to react to the frequency variation in a time window less than 300 ms and it should reach the steady state in a maximum of 1 second.

Fig. 7.2 shows an example of the expected response. In this work, the transient specifications, such as the overshoot and the steady-state error, which are strictly related to the power electronics components of the Electric Vehicle Supply Equipment (EVSE), are not considered.

The extension of smart charging with a frequency response service is explained by the example shown in Fig. 7.3. The example of Fig. 7.3 considers a PEV subject to a V1G smart charging session. The first plot represents the network frequency, with a dashed line for network frequency nominal value f_n (in Europe 50Hz), and a continuous line for measured network frequency $f(t)$. In the example, the time trajectory of the frequency network is characterized by different distortions, with corresponding frequency deviation $|\Delta f| = |f(t) - f_n| \geq \overline{\Delta f}_{min}$ where $\overline{\Delta f}_{min}$ represents the frequency deviation threshold implemented by the dead-band. The second plot shows the superposition of the frequency response service on the smart charging session: the dashed line represents the nominal charging setpoint $\tilde{p}(t)$ assigned by the smart charging system at different time instants; the continuous line represents the actual setpoint commanded by the system. The charging session presented in the example well shows the concept behind the introduction of power margins: if the power setpoint for the charging session is between maximum and minimum charging power, the power gaps can be used to change the charging power of a term $\Delta p(\Delta f)$ in response to frequency deviations. The pair of the two plots explains the concept at the basis of the service integration. During nominal frequency operation, the EVSE follows the charging reference generated by the smart charging scheduler. In presence of a frequency deviation out of the dead band, the EVSE superimposes an additional contribution $\Delta p(\Delta f)$ to the smart charging setpoint, to aim at steering the network frequency back to the reference value.

A crucial point consists in the construction of the power-frequency (p-f) curve, which defines the variation of the charging power as a function of the frequency deviation. This work aims at illustrating and discussing different possible strategies, highlighting the differences in terms of performance and compliance with the requirements discussed before.

7.4.1 Case I - linear interpolation

A first attempt for the p-f curve assignment is a strict separation between the smart charging service and the frequency regulation service. In this case, the smart charging system does not provide an active contribution for aFRR service, but it only introduces power margins (since typically the PEVs will not be charged at maximum power, and thus some power margins will be available for the provisioning of aFRR services). Each EVSE equipped with a Local Control Agent, receiving the

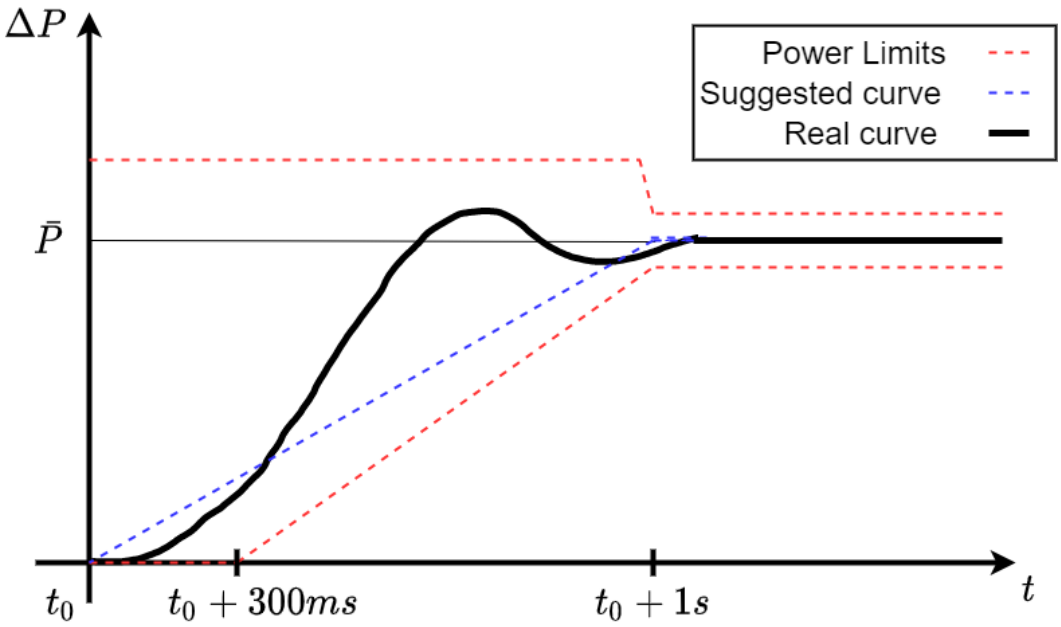


Figure 7.2. Expected response of a fast reserve unit [2]

smart charging power setpoint $\tilde{p}$ and knowing the power limits of the whole charging system, linearly interpolates in an independent way the points $(\overline{\Delta f}_{min}, \tilde{p})$, $(\overline{\Delta f}_{max}, p^{max})$ and $(-\overline{\Delta f}_{min}, \tilde{p})$, $(-\overline{\Delta f}_{max}, p^{min})$ where $\overline{\Delta f}_{max}$ characterizes the frequency deviation over which the EVSEs have to provide the full power margins. Fig. 7.4 shows a representative example of this approach in a V2G scenario with two EVSEs (blue and red curves): the main advantages of this strategy are the decoupling between smart charging service and aFRR service, the possibility to compute the p-f curves at the level of EVSE and the exploitation of all the available margins at the level of EVSE. The drawback of this strategy is in the resulting shape of the load area curve (that is composed of the curves of each active session in the Load Area). Indeed, with this approach, the symmetry requirement for the aggregated power-frequency curve is in general not ensured (see Fig. 7.4b).

7.4.2 Case II - linear interpolation with load area control

The presence of smart charging service is exploited not only to create margins at the level of the EVSE but can be also used to manage the power margins at the load area level. As in [139] [140] where an external signal is used to drive the aggregated load area power in order to satisfy a Demand Side Managment (DSM) service, the same methodology is used to impose a specific aggregated power withdrawal, i.e., half of the nominal power of the active charging sessions present in the load area at the

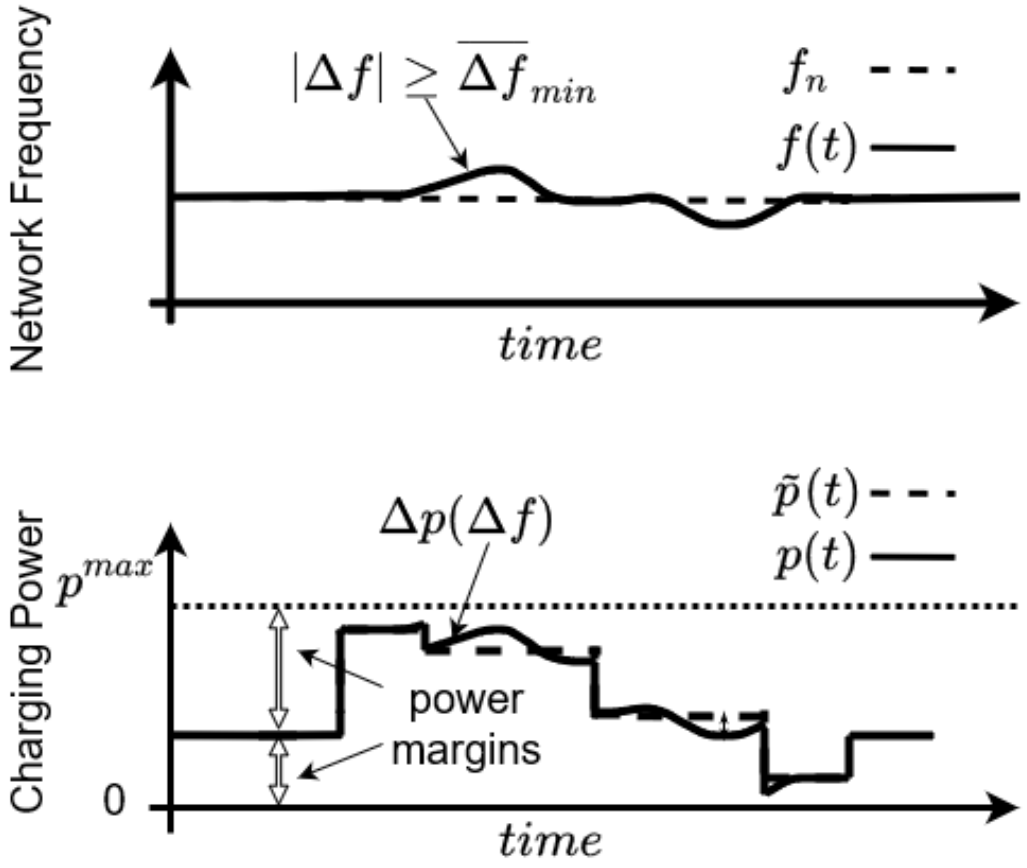


Figure 7.3. Example of the superposition of smart charging and frequency regulation services: top - network frequency time evolution, bottom - associated charging session

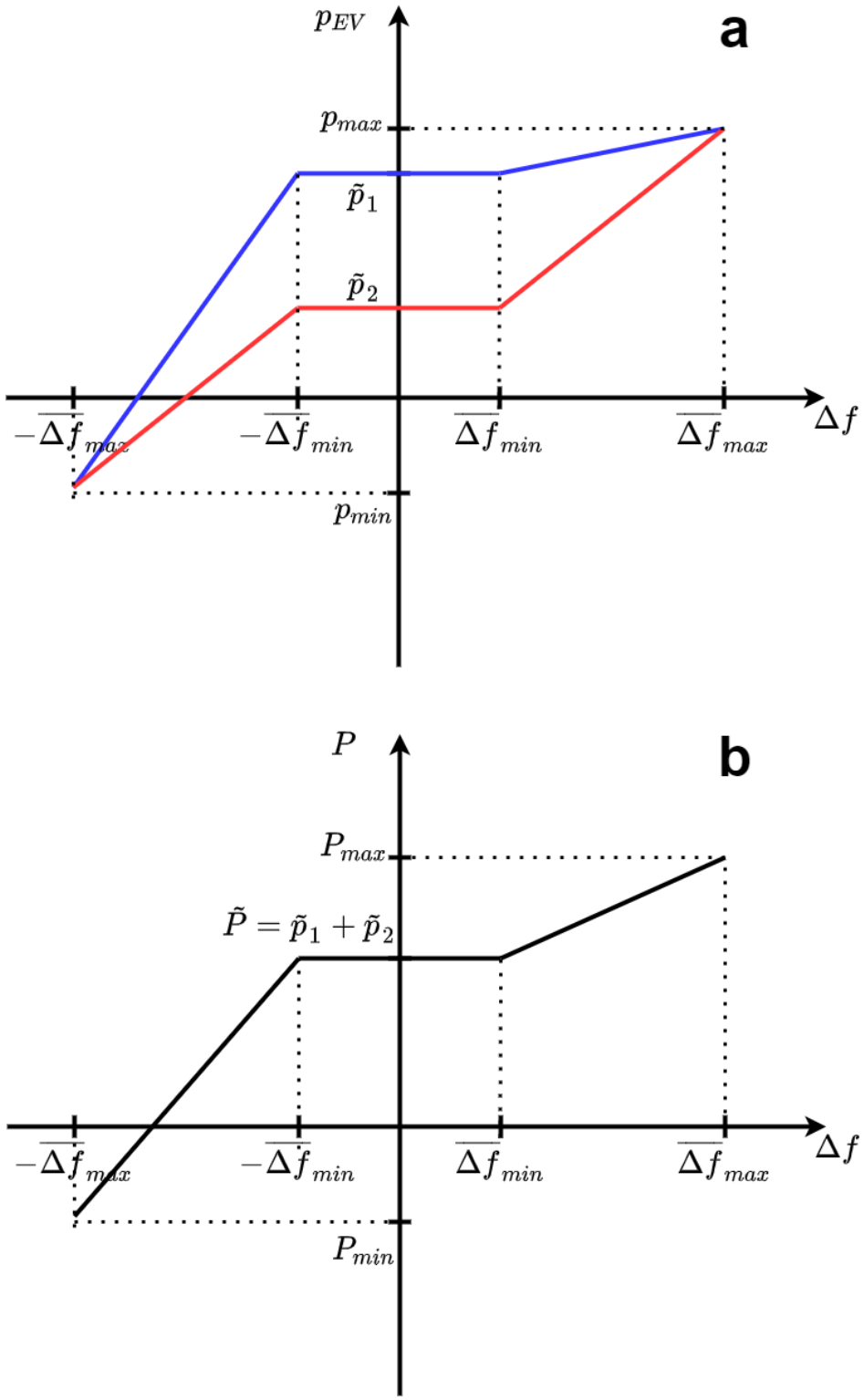


Figure 7.4. Case I - linear interpolation. a) frequency regulation p-f curves of two EVs ($\tilde{p}_1$ and $\tilde{p}_2$), b) cumulative p-f curve $\tilde{P}$

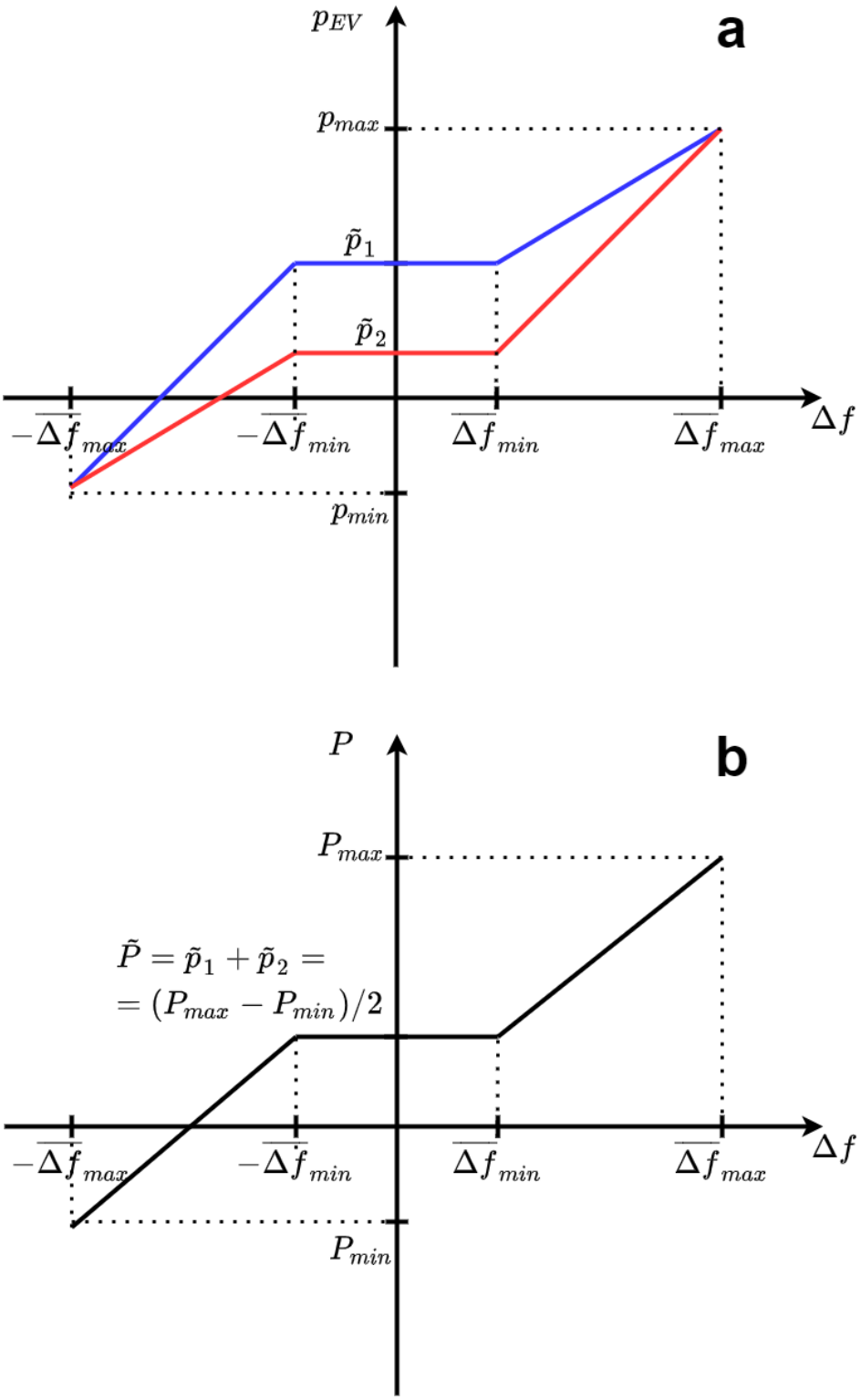


Figure 7.5. Case II - linear interpolation with load area control. a) frequency regulation p-f curves of two EVs, b) cumulative p-f curve

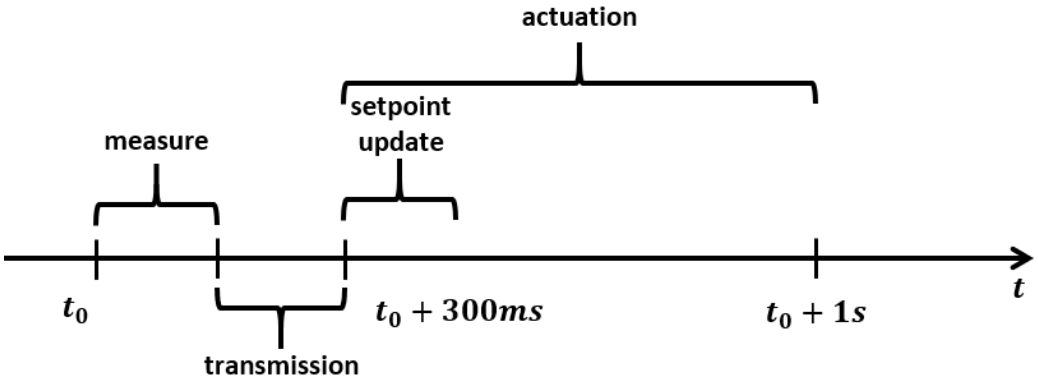


Figure 7.6. Delay Budget

given time. In this condition, by applying the linear interpolation strategy presented before, even if at the level of the single EVSE the curves are not symmetric, the aggregate p-f curve satisfies the symmetry requirement (Fig. 7.5). Nonetheless, this strategy is characterized by different issues and limitations: first of all, the load area should be operated at half of the nominal power capacity. In addition, the symmetry is strictly related to the ability of the smart charging system to follow the load area power setpoint.

7.4.3 Case III - resources allocation

A more sophisticated strategy considers the implementation of a resource allocation algorithm, which assigns to each EVSE a specific p-f curve (even a nonlinear one), such that the aggregated load area curve satisfies all the requirements discussed before. The integration between the smart charging system and the resources allocation algorithm results in an integrated system that, managing the two services, allocates the resources considering drivers' requirements, balancing and distributing the service provisioning based on user profiling. The strategies presented before are less complex and easier to implement, while this last one requires the development of a resources allocation algorithm, that implies communications between EVSE and/or deep integration with the smart charging system. In next chapter, a mathematical formulation for computing the (p-f) curves will be proposed.

7.5 The delay budget problem

In order to enable PEVs to the frequency regulation functions, some very strict constraints on the delay between the occurrence of the frequency disturbance and the actuation of the control signal by the EVSEs power unit on the PEVs have to be

considered. In particular, the Pilot Project Fast Reserve requires an initial response from the PEVs within 300ms from the occurrence of the frequency disturbance event, and a full response of the system (the end of the transient and the establishment of a steady-state at the frequency-dependent power setpoint) within 1s.

As detailed in Fig. 7.6, and having in mind the proposed architecture in Fig. 7.1, the system in the first 300ms from the frequency disturbance event (happening at time t_0) have to:

- measure, through the frequency meter, the frequency disturbance;
- transmit it to the Master Control Agent hosted in the 5G MEC;
- the Master Control Agent has to propagate this frequency measure between all the Local Control Agents;
- the Local Control Agents have to compute the updated setpoints;
- the Local Control Agents have to communicate the new setpoints to the Control Units of the EVSEs.

Then, from $t_0 + 300$ms to $t_0 + 1$s, the PEVs must completely actuate the new power setpoints.

Considering the frequency meters available on the market, it is possible to estimate a measurement delay in the order of a few hundred milliseconds (usually 100ms-200ms), while the computation of the new setpoints, as explained in section 7.4, can be very simple, and so executed in less than a millisecond, or more advanced, and so executed in few dozen of milliseconds.

The transmission delay is a critical factor in the total delay budget since the system components must receive the frequency measure well in advance of $t_0 + 300$ms, so to have time to process it and to update the power setpoints. Then, a maximum communication delay must be guaranteed by telecommunications operators, as the TSO requesting the services may apply penalties to the CPO for not providing the service within the time constraints.

Using wired telecommunication technologies, this delay can be in the order of a few milliseconds (for optic fiber), or in the order of dozen milliseconds (for copper wires), which is in line with frequency regulation service requirements. However, considering the dispersed nature of the EVSEs, this kind of solution may be very expensive for the CPO, implying a cost that may be considerably high compared to the investment needed to install a frequency meter inside each EVSE.

Moreover, most of the EVSEs are nowadays connected through cellular networks to their back-end, so, in principle, it is possible to exploit the cellular connectivity already present on mean/guaranteed latency to perform such tasks. Indeed, 3G and

4G technologies have end-to-end communication latency respectively in the order of hundreds of milliseconds and about 50-100 milliseconds. 3G communication delay occupies almost all the available time, while 4G latency may not guarantee that there is enough time for measurement and for computation. Moreover, both technologies do not support slicing, so the overall performance and the delay (at least on the radio access part of the network) cannot be guaranteed easily. On the contrary, 5G communication technologies reduced the radio access delay to a few milliseconds, as well as the end-to-end delay with a novel fronthaul/backhaul/core network architecture made of Virtual Network Functions. Moreover, 5G introduces network slicing, so the telco operator may provide an exclusive slice of its resources to the CPO's frequency regulation system. Moreover, 5G architecture introduced, as explained in section 7.3, the MEC at the edge of the radio access network. Indeed, deploying Master Control Agent module inside MEC, it is possible to communicate with the frequency meter and to the Local Control Agent with negligible delays (in the order of a few milliseconds), making feasible the provisioning of frequency regulation services with PEVs. Finally, since the EVSEs already have a modem installed, the only cost for the CPO is to replace it with a 5G-compatible one. This cost for the CPO is justified also by the obsolescence of older communication technologies, that may be discontinued in the next few years, forcing the CPO to change in any case the EVSEs modems.

7.6 Conclusion

This work has presented how plug-in electric vehicles can participate in the frequency regulation ancillary service provisioning. In the work, the reference scenario, the systems, and the actors involved are presented and discussed showing how the roles of CPOs and DSO will be expected to evolve in the next few years.

The work provided a control architecture that exploits the novel 5G network architecture and its very low latency. The work focused on the superposition between smart charging services and frequency regulation services, providing three different strategies and discussing their properties, pros, and cons.

The work also provided a deep discussion on the integration between the frequency regulation service, the ICT infrastructure, and charging, measuring, and communication components, providing quantitative considerations on the system feasibility analyzing the effects of different ICT technologies on the service quality.

In future works, the proposed control strategy and its variants will be extended and tested using hybrid real/virtual test scenarios. The real components, such as 5G network nodes, charging infrastructures, and measurement devices, will be used to

validate the feasibility of the approach, while high-fidelity real-time power network simulators will be used to analyze the impact of the service provisioning strategy.

Chapter 8

Optimal assignment of Droop Curves for Frequency Regulation services composition

The work presented in this chapter is strictly related to the one presented in the previous chapter. Indeed, the previous work presented a reference system architecture to provide Frequency Regulation Services using PEVs and 5G network, together with the delay budget required to make this approach feasible according to the pilot project in [2]. Moreover, in the previous work, three different approaches (at increasing complexity levels) are presented to provide the proposed Frequency Regulation Service. Here in this work, a linear optimization problem formulation for the third approach is presented, which generates linear and non-symmetric local droop curves for the PEVs, in such a way the aggregate global droop curve is equal to the one contracted with the TSO for the provisioning of the service, which is symmetric (as required in [2]) and characterized by a certain amount of power to absorb/release in case of frequency disturbance.

Moreover, the local droop curves computed by the proposed algorithm consider also the actual/desired state-of-charge of the PEVs and the remaining dwelling time, as to make them possible to participate with more power to the PEVs that have an actual state-of-charge near to the desired one and sufficient remaining dwelling time, while reducing the contribution of PEVs that are far from their reference state-of-charge or that have too small remaining dwelling time.

Since the proposed approach does not guarantee that, in case of the activation of the service, the desired state-of-charge of a generic PEV is reached before the remaining dwelling time. To this aim, the proposed approach has been designed to run over a smart charging controller (such as [139, 140], where [140] has been

actually implemented by the Candidate inside a 5G MEC in the context of the European H2020 project 5G Solution [131]) that is able to bring back the actual state-of-charge of the PEV towards the desired one within the remaining dwelling time by properly modulating its charging power.

8.1 Introduction

In recent years, the European energy market opened the possibility for third parties to provide regulation services (e.g., ancillary services). Among these control services, a very important role in the safe operation of the electricity network is given to the frequency control services, which aim at maintaining the network frequency in a safe range. Frequency regulation services are usually provided by traditional generators, through speed droop controllers that control the mechanical power applied to the rotor shafts based on their rotation speed. However, it is possible to control the network frequency also with other kinds of generators, like wind farms [142,143] and photovoltaic plants [144]. Moreover, many papers intercepted the possibility to modulate the energy demand of heavy flexible electric loads, like heating, ventilation, and air conditioning systems [145], and thermostatically controlled loads [146], in order to control network frequency. Among the flexible loads, PEVs emerged as good candidates for the provision of frequency regulation services, as their power consumption can be changed remotely by the charging stations, and the reaction time is very low. Anyway, the power absorbed by a single PEV is not enough to participate effectively in the ancillary service market, so it is fundamental to aggregate a set of PEVs and coordinate their action to provide the requested services. The possibility to aggregate many charging PEVs is becoming easier, due to the mass spreading of PEVs in the market, and thanks to new high-speed/low-latency cellular communication means [147], which reduce a lot the cost to update the current CPO infrastructure to provide frequency regulation services. Indeed, this work proposed a novel architecture [148] that enables CPO charging stations to provide frequency regulation services over a 5G communication network, exploiting the 5G multi-access edge computing [149] as edge computing node, without violating the time constraints to provide this kind of services. The proposed architecture [148], together with the results coming from this work, will be implemented in the context of the European H2020 Project 5G-Solutions [131].

The problem of controlling PEVs for frequency regulation services has been studied in the literature by Yao, Wong, and Schober with a focus on a robust control method to estimate the hourly regulation capacity (i.e., aggregated power margins) of the PEVs [150], or with a focus on the market bids for the frequency

regulation services provided by PEVs [151]. Moreover, Xia et al. studied in [152] the problem of computing power setpoints for fleets of PEVs in presence of fluctuating wind generators, using consensus.

In addition, the contribution of PEVs to fast frequency regulation has been studied in [153], where the frequency droop curves for each PEV cluster are computed using an improved harmony search algorithm, but without balancing the effort of each PEV based on its smart charging requirements.

Kuang et al. have studied in [154] the problem of providing frequency regulation by using dispersed PEVs while maintaining the target SOC level requested by the PEV owners, but without changing dynamically over time the control droop curves to balance the effort among the different PEVs based on the drivers' requirements.

In [155], the authors provided a hierarchical control scheme to allocate frequency control effort among the charging PEVs that is able to ensure the final SOC level for the PEVs by real-time corrections of the allocated frequency control action. Anyway, this approach does not optimize the effort among the different PEVs, allocating it based on the frequency regulation capacity of each PEV.

Jia et al. have studied in [156] the problem of load frequency control using PEVs with inertia uncertainties and time-varying delays, and in particular, focusing on the coordination among PEVs and power plants to provide such frequency control functionalities, but without considering the integration of smart charging system with load frequency control, which is one of the focuses of the presented work.

In [157], the authors have discussed a demonstration of the provisioning of frequency-controlled normal operation reserve (one of the primary frequency regulation services in the Nord Pool energy market) using three PEVs from different manufacturers. Differently from the present work, the charging setpoints to implement the frequency regulation service are computed by a central controller, which results in higher response times compared to the decentralized solution.

In [158], a model predictive control scheme to schedule the bidding in the frequency regulation market is proposed, with the aim of maximizing the payment to PEV aggregators, based on a prediction of the frequency regulation market price built using a seasonal-autoregressive integral-moving-average model. Though taking into account the current PEVs SOC level, the PEV power schedule is limited to three possible states and, again, the effort is not balanced among PEVs; moreover, the algorithm does not provide droop curves to be applied by the PEVs in response to frequency deviations.

Finally, in [159], Islam et al. propose a Markov decision process formulation of the problem of aggregating flexibility from the PEVs for providing frequency regulation services. The goal is to maximize the revenue of the aggregator. A

possible limitation of this approach is in its scalability, and in the possibility of having a fine-grained control for each PEV (due to the use of the state aggregation technique).

The distinctive features of the present work are: i) an approach in which frequency control actions are taken at each charging station based on local droop curves, which allows for very fast intervention, as required for providing fast frequency regulation. The parameters of the local droop curves are periodically updated by a central controller, which allows for equally balance frequency control effort among the PEVs while having a highly scalable scheme, since the central optimization problem is a linear one and is executed within larger time constraints with respect to the frequency regulation control loop ones; ii) the droop curves are computed as to balance the control effort among the various PEVs, taking into account the smart charging requirements of the different users, on available dwelling time, current SOC, and final desired SOC; iii) the presented algorithm exploits the availability of edge computing nodes, such as 5G multi-access edge computing, to reduce as much as possible communication latencies with the charging stations, and with the frequency meter.

In summary, the proposed approach is scalable, it allows for very fast frequency control action, and it ensures a correct balancing of the control action among the participating PEVs, by taking into account their individual conditions. Therefore, the algorithm summarizes the strengths found separately in the different works in the literature.

8.2 Reference Scenario and Problem Description

In the last few years the European energy market has moved towards a separation between the roles of the balance responsible party (i.e., the entities which are responsible for electricity network imbalances), and the balance service provider (i.e., the entities which provide balancing services for the electricity network). In this context, the DSOs can rely on the balance service providers for their balancing services [160]. The CPO can, thus, take advantage of this market separation, becoming itself a balance service provider and providing ancillary services to the power network, by leveraging the possibility of changing the power setpoint of the charging station in real-time. Anyway, the CPO cannot participate in the ancillary service market with each separate charging station, since the power provided by each charging station is not in line with the power requirements to access this market; instead, it can participate after aggregating many (even physically dispersed) charging stations, and so providing balancing services both at DSO and transmis-

sion system operator level. In this context, the authors of this work proposed a smart charging solution that can enable fleets of charging PEVs to demand-side management services [140].

Within the same scenario, it is possible, in principle, to provide also frequency regulation services, by leveraging the power margins provided by the smart charging system. Indeed, most of the time the charging PEVs charge with a power setpoint that is between zero and the maximum charging power, and so there are power margins both to increase and to reduce the power absorbed by the PEVs. These margins can be aggregated and used to participate in the frequency regulation services.

Fig. 8.1 shows the reference scenario considered in this study, which is based on the same novel architecture proposed in [148]. A set of PEVs are connected to charging stations that are remotely monitored and controlled by the CPO backend. The charging power setpoints actuated by the charging stations are periodically (e.g., every minute in [140]) computed by smart charging algorithms hosted in the "smart charging module". The goal of the smart charging module is to control the PEV recharging in the load area, in such a way that the PEVs are recharged in compliance with the user preferences (e.g., time available for recharging and amount of energy to recharge), and in compliance with the technical constraints of the grid. Hence, the role of the smart charging module is to ensure safe (for the vehicle and the grid), efficient (e.g., taking into account battery degradation), and economical smart charging service to the users. For an example of a possible implementation of the smart charging module, and the associated smart charging algorithms, the reader is referred, e.g., to the previous work [139].

The role of the "local droop curves computing module" instead is related to the provisioning of frequency control services. The module hosts a control algorithm that builds frequency control services by leveraging the flexibility offered by the PEVs. The module periodically computes and sends to the charging stations one power-frequency droop curve for each PEV performing smart charging, and agreeing to participate in frequency regulation. The droop curves are frequency/power curves that, given the current value of the network frequency, specify the power deviation with respect to the smart charging setpoint, i.e., how much the charging power setpoint computed by the smart charging module for the specific vehicle should be increased or decreased at any time, depending on the current grid frequency value. In absence of the frequency regulation service, the charging station simply implements the power setpoint it receives periodically from the smart charging module. When the PEV also participates in the frequency regulation, the charging station adds or subtracts to the smart charging setpoint a delta of power that is

given by the value of the local droop curve at the current value of the grid frequency (frequency measurement is communicated to the charging station via 5G). To be able to modulate in real-time the power setpoint, as required in this scheme, the charging station needs to be equipped with a power converter (as in direct current charging stations). As discussed in the following, the different droop curves are computed in such a way that their total, cumulative effect follows a desired power-frequency droop curve, as if the PEV fleet was a unique entity providing frequency regulation services.

The relation between the two modules introduced above will be fully characterized in future works. It is important to remark here that each local droop curve is computed on the basis of the specific parameters of the related charging session. The logic is that the contribution of each PEV to the frequency regulation should be tuned depending on the current progress of the charging session and on the technical characteristics of the PEV. On the other hand, the implementation of the droop curves by the single PEVs means that a different setpoint is actuated by the PEV, compared to the one that was computed by the smart charging module. The effect of this perturbation is compensated by the fact that the smart charging module periodically recomputes the charging schedules (see, e.g., [139]), so that the new charging schedule is computed taking into account, and to compensate, the effect of any possible deviation from the previous schedule caused by the participation of the PEV to the frequency regulation service (and by any other disturbance affecting the charging process). The effect of the perturbation on the smart charging session caused by the participation of the PEV in the frequency regulation service is measured by the smart charging module by measuring the current state of charge of the PEV, before each re-calculation of the smart charging setpoint. This scheme ensures that the two services, i.e., the smart charging service, and the frequency regulation one, can coexist, in a way that is transparent for the user.

The present work focuses entirely on the discussion of the control algorithm hosted by the local droop curves computing module. The interaction of this module with the smart charging module will be fully analyzed in a future publication.

In order to enable the charging stations to provide frequency regulation services, the system should be able to take a frequency measure and actuate the proper power setpoint to the PEV (increasing or reducing it based on the frequency deviation) within a very limited time (in the order of 300 ms, according to the *Fast Reserve* pilot project of TERNA [2]), as analyzed in [148]. It is important to notice that this very strict time constraint has to be guaranteed just for frequency regulation functions, while other smart-charging functionalities (e.g., the computation of nominal power setpoints) have much larger time constraints, that are typically in the

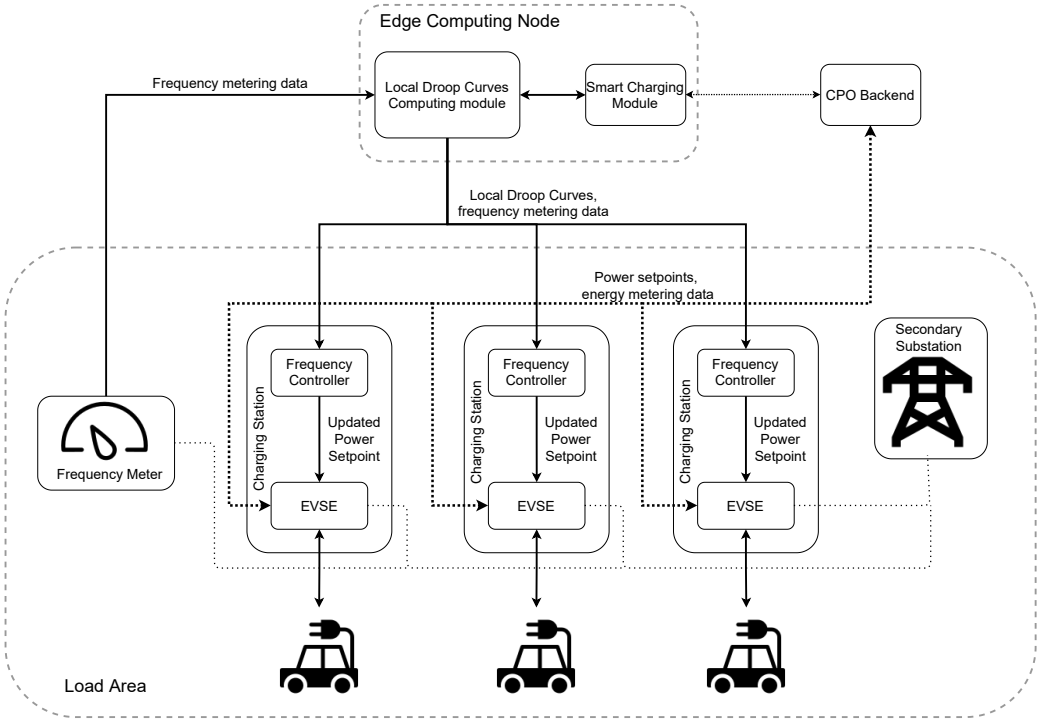


Figure 8.1. Reference System Architecture

order of minutes. Installing a frequency meter capable of the above-mentioned time requirements inside each charging station would imply very high costs. However, by leveraging the new cellular communication technologies (e.g., 5G), it is possible to install just one (or few, for redundancy) frequency meter in each load area, and then spread the information to the other charging stations with negligible delay, thus matching the strict time requirements for the service.

8.3 Nomenclature and Problem Formulation

In this section, the proposed control algorithm hosted in the "local droop curves computing module", which periodically computes the local droop curves (i.e., one for each charging station), based on the current status of the ongoing charging sessions is detailed. In particular, the local droop curves are computed with the same sampling time of the smart charging signals (e.g., nominal power setpoints for the charging sessions), so in the order of minutes. First of all, the used nomenclature is introduced.

Let N be the number of PEVs connected at the generic time k in the load area. The generic n-th PEV at time k is characterized by:

- The current charging power level $P_{n,k}$;

- The maximum and the minimum possible charging power levels, respectively $P_n^{max} > 0$ and P_n^{min} (if $P_n^{min} < 0$, the PEV is enabled to discharging);
- The current SOC level $x_{n,k}$;
- The time left until the end of the charging session, $d_{n,k} > 0$;
- The error, $e_{n,k}$, between the desired SOC, x_n^{ref}, and the current one, $x_{n,k}$, i.e., $e_{n,k} := x_n^{ref} - x_{n,k}$;
- The power deviation, $\Delta P_{n,k}$, at time k, for the n-th PEV, due to the participation in the frequency regulation service. This value is computed from a droop curve.

The proposed algorithm is aimed at optimally coordinating the connected PEVs in the participation in the provisioning of ancillary services. In more detail, the problem tackled in this work is that of optimally defining p-f droop curves at single PEV level (called *local droop curves*), in such a way that, once combined, they match a given, desired droop curve (the *global droop curve*). The global droop curve defines how, *collectively*, the connected PEVs should react to frequency mismatches, as if they formed a unique entity participating in the provisioning of the ancillary service.

Fig. 8.2 displays a general prototype of a global p-f curve. Focusing on the right-half plane, the parameter Δf_{min} defines the deadband: the power variation is zero if the deviation of the frequency with respect to the reference frequency value is in the interval $[0, \Delta f_{min}]$. Δf_{max} defines the frequency deviation limit after which the power variation saturates. m^{global} defines the droop, i.e., the ratio between the variation of the power and the frequency deviation.

The algorithm presented next recovers the global droop curve as the sum of N local droop curves. A notable aspect is that the design of each local droop curve takes into account the current status of the PEV's recharging process, as explained next.

For the design of the local droop curves, the following assumption is made.

Assumption 8.3.1 (Shape of local droop curves). Local droop curves have a positive slope, and they are linear between Δf_{min} and Δf_{max}.

Assumption 8.3.1 is included because it is in line with the common design principle of standard p-f droop curves (the algorithm proposed in this work could work also with nonlinear droop curves). As implied by Assumption 8.3.1, the focus of this work is on linear droop curves, i.e., on curves described by the equation:

$$\Delta P_{n,k}(\Delta f) = m_{n,k}\Delta f + q_{n,k}. \tag{8.1}$$

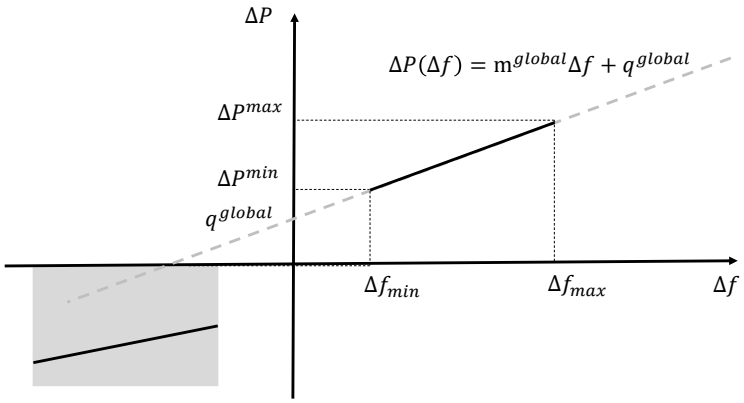


Figure 8.2. Global droop curve, with associated relevant parameters.

The focus of this work is on linear local droop curves because standard droop curves are linear, and because working with such types of functions results in a linear optimization problem, which has low computational complexity.

The algorithm proposed in this work optimally designs each local droop curve by selecting the parameters $m_{n,k}$ and $q_{n,k}$, which are therefore the optimization variables of the problem. The algorithm is presented in the next section. To keep the description of the algorithm concise, and without loss of generality, the work focuses on the design of the portion of the droop curves lying on the right half of the $\Delta f/\Delta P$ plane (see Figs. 8.2, 8.3). The part of the curve on the left half of the plane is designed in a similar way.

8.4 Proposed Local Droop Curves Design Algorithm

The design of the generic local droop curve (8.1) must respect constraints related to the shape of the local curve, and others related to the shape of the global droop curve. It must also respect certain limitations imposed by the current status of the charging sessions, as explained next.

8.4.1 Local Droop Curve Design Constraints

Fig. 8.3 displays in gray the "design space" in which the generic n-th local droop curve can be drawn. The following constraints for each local curve are included.

$$0 \leq \Delta P_{n,k}^{min} := m_{n,k}\Delta f_{min} + q_{n,k} \leq P_n^{max} - P_{n,k}, \tag{8.2}$$

$$0 \leq \Delta P_{n,k}^{max} := m_{n,k}\Delta f_{max} + q_{n,k} \leq P_n^{max} - P_{n,k}. \tag{8.3}$$

The above two constraints state that, respectively, at Δf_{min} and Δf_{max}, the power increase for the single PEV must be between zero (i.e., corresponding to a null

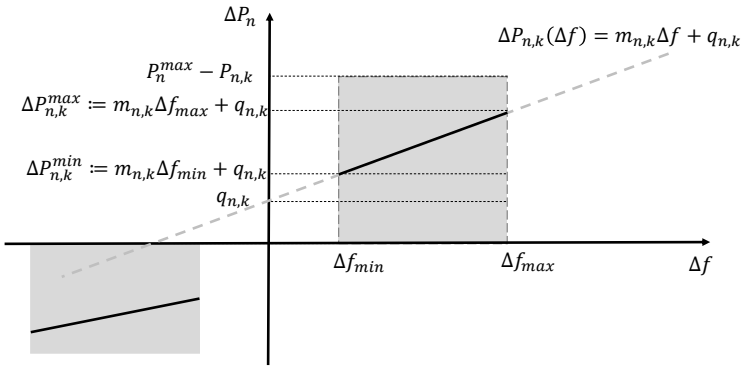


Figure 8.3. Local droop curve, with associated relevant parameters.

contribution to the frequency regulation services, and therefore having no impact on the current charging power level), and the maximum possible increase in charging power, i.e., $P_n^{max} - P_{n,k}$, which takes into account the current charging level, $P_{n,k}$, and the maximum possible one, P_n^{max}. Notice however that (8.3) alone is not sufficient, since the maximum power increase might be affected also by the current SOC. For example, the maximum power increase for a vehicle that is fully charged is zero. For this reason, the following constraint is added:

$$P_{n,k} + \Delta P_{n,k}^{max} \leq \frac{x_n^{max} - x_{n,k}}{T}. \tag{8.4}$$

Equation (8.4) states that the maximum power increment is limited by the maximum energy that can be charged into the battery in the unit of time. T is the sampling time of the algorithms, i.e., every T seconds the local droop curves are re-computed.

Given the focus on linear droop curves, (8.2), (8.3), and (8.4) ensure that the contribution of the single PEV is always feasible, also considering the current status of the charging session.

Next, in line with Assumption 8.3.1, local droop curves must have a positive slope, i.e.:

$$m_{n,k} \geq 0. \tag{8.5}$$

This of course implies also that $\Delta P_{n,k}^{min} \leq \Delta P_{n,k}^{max}$.

8.4.2 Global Droop Curve Design Constraints

The p-f droop curve arising from the superimposition of the local droop curves can be written as

$$\Delta P(\Delta f) = m_k \Delta f + q_k, \tag{8.6}$$

where

$$m_k = \sum_{n=1}^{N} m_{n,k}, \quad q_k = \sum_{n=1}^{N} q_{n,k}. \tag{8.7}$$

The following additional constraints are included to ensure that the superimposition of the local droop curves matches the desired global droop curve:

$$m_k = m^{global}, \tag{8.8}$$

$$q_k = q^{global}. \tag{8.9}$$

8.4.3 Target Function

In the above subsections, the constraints that must be respected in the selection of the parameters $m_{n,k}$ and $q_{n,k}$ were presented, in order for the deriving local curves and global curve to be feasible d-f droop curves. In the following, the formulation of the proposed target function, to ensure that the parameters $m_{n,k}$ and $q_{n,k}$ are selected in an optimal way, according to the current status of the charging session of the PEVs participating in the provisioning of the ancillary service will be discussed. The proposed target function to be *minimized* is:

$$\begin{aligned} J_k = \sum_{n=1}^{N} & - \alpha_1 e_k(\Delta P_{n,k}^{min} + \Delta P_{n,k}^{max}) + \alpha_2 d_k(\Delta P_{n,k}^{min} + \Delta P_{n,k}^{max}) + \\ & + \alpha_3 m_k^{max} + \alpha_4 \Delta P_k^{max_r} + \alpha_5 \Delta P_k^{min_r}, \end{aligned} \tag{8.10}$$

where $\alpha_1, ..., \alpha_5 \in [0, 1]$ are such that $\sum_{i=1}^{5} \alpha_i = 1$, and m_k^{max}, $\Delta P_k^{max_r}$, $\Delta P_k^{min_r}$, are auxiliary variables such that:

$$m_{n,k} \leq m_k^{max}, \quad m_k^{max} \geq 0, \tag{8.11}$$

$$\frac{\Delta P_{n,k}^{max}}{P_n^{max} - P_{n,k}} \leq \Delta P_k^{max_r} \quad \forall n \in N, \tag{8.12}$$

$$\frac{\Delta P_{n,k}^{min}}{P_n^{max} - P_{n,k}} \leq \Delta P_k^{min_r} \quad \forall n \in N. \tag{8.13}$$

Coefficients $\alpha_1, ..., \alpha_5$ can be used to weigh the terms of the objective function. It is easy to see from (8.10) and (8.11) that, at the optimum, m_k^{max} is equal to the maximum value of $m_{n,k}$, for $i = 1, ..., N$. Hence, the inclusion of this term in (8.10) has the goal of minimizing the maximum value of $m_{n,k}$, i.e., of balancing the effort of the participation in the ancillary service provisioning among the PEVs, aiming to avoid that some vehicles are assigned steep droop curves (i.e., high values

of $m_{n,k}$). The last two terms in (8.10), similarly to $m_{n,k}$, contribute to spread the effort between the vehicles, by minimizing the maximum share of the available power margin that each vehicle contributes to frequency regulation (see (8.12) and (8.13)).

The first and the second terms in (8.10) instead are included to take into account also the current status of the charging session of each PEV, and, specifically, to give priority to the PEVs with a larger SOC error (the first term - notice the minus sign), and the ones with a smaller remaining charging time (the second term). The local droop curves associated with these PEVs, will have a more pronounced slope and/or a higher power at Δf_{min}. As a result, they will contribute more to the provisioning of the ancillary service, which will help them in reaching earlier the desired SOC level. Notice that, when the SOC error e_k is negative (i.e., the current SOC is higher than the reference), then the PEV will contribute less to the provisioning of the service, which is a consistent behavior. Finally, recall that, for the sake of brevity, the proposed work focuses on the design of the right-half part of the droop curves. The design of the left-half part of the curve, also in terms of the choice of the target function, can be carried out with similar considerations.

8.5 Numerical Tests

Simulations have been performed in Julia [161], version 1.6.0, on a standard computer (3.3 GHz, I7 processor with 16 GB RAM). The simulation scenario is as follows. There is a charging load area that participates in the frequency regulation service. It is assumed that the charging sessions active in the load area are enough to provide the required power-frequency curve, i.e., that the composition algorithm presented in Section 8.4 admits a solution.

In the following, two simulation scenarios are discussed, to validate the proposed approach:

1. Scenario 1: the algorithm is run in a balanced scenario, i.e., considering a set of charging sessions that are homogeneous in terms of power margin flexibility, SOC error, and charging time availability;

2. Scenario 2: the algorithm is run in a scenario in which the charging sessions have different power margins, different SOC errors, and time flexibility.

The two scenarios are meant to show that, from one side, the algorithm is able to come up with a fair and balanced assignment of local droop curves among the participating PEVs while, on the other hand, taking always into consideration the real-time SOC and time flexibility status of the participating PEVs, as determined by the respective charging session status, controlled by smart charging algorithms.

8.5.1 Scenario 1: Local Droop Curves Assignment in a Balanced Scenario

The first set of simulations is aimed to show how the algorithm is able to design local droop curves that result in a balanced distribution of the regulation effort, among the participating PEVs. To this end, 3 active charging sessions are considered, characterized by each one by the same dwelling time and SOC error. The distinguishing attribute between the sessions is the available margin of power, $\Delta P_{n,k}^{max}$, which reflects:

- The possible presence of different charging technologies in the load area, i.e., the fact that the charging sessions are characterized in general by different maximum power, depending on the charging technology;
- The presence of smart charging sessions, i.e., the fact that the charging sessions happen at different charging levels, which are in general different from the maximum possible charging level.

Indeed, the algorithm must be able to work in presence of smart charging sessions ongoing at different charging levels.

The first simulation in this scenario shows the case in which the charging sessions are characterized by the same power margin. Specifically, the charging sessions are characterized by a maximum charging power of 150 kW, and by a common charging setpoint of 100 kW, resulting in a power margin of 50 kW. The global droop curve, for over-frequency events, is characterized by a maximum power deviation of 105 kW (70% of the available power margin of 150 kW), and by a frequency range $[\Delta f_{min}, \Delta f_{max}] = [500, 1500]$ mHz. In correspondence with the minimum of the bandwidth, the given global droop curve is characterized by a power deviation of 10 kW. Fig. 8.4 shows the result of the algorithm. The algorithm determines the optimal distribution as an equal allocation of droop curves between the charging session. The target droop curve is identified in the plot with the red-dashed line, instead, the resultant droop curve is represented by the black-dotted line. Fig. 8.4 shows that the sum of the local curves matches exactly the target global one. The presence of the last term in the cost function determines an exact balancing of the frequency regulation service among the charging sessions, i.e., leading to identical local droop curves. Fig. 8.5 shows the percentage of maximum usage of the power margin of each vehicle, i.e., the ratio between the value of the local droop curve corresponding to the frequency deviation Δf_{max} and the available margin of power. This plot confirms what was already discussed: the algorithm distributes the effort in order to assign the same maximum relative usage to each vehicle's power margin, in this case 70%.

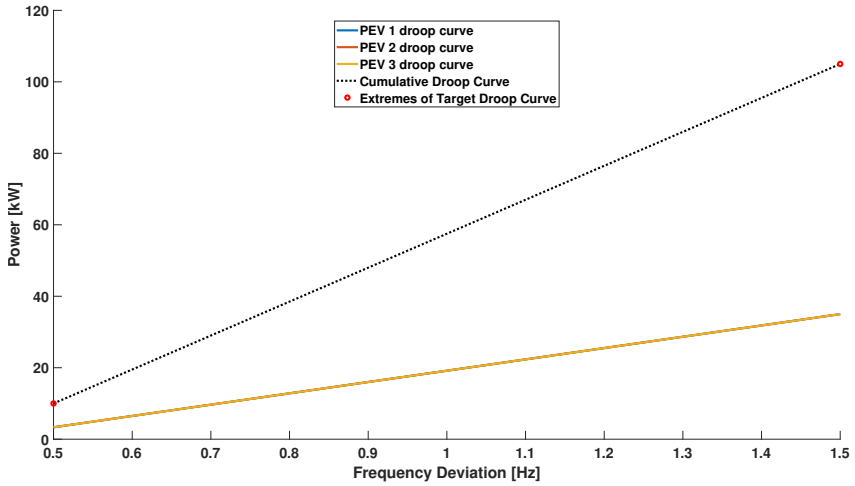


Figure 8.4. Scenario 1, balanced conditions: resulting local and global droop curves.

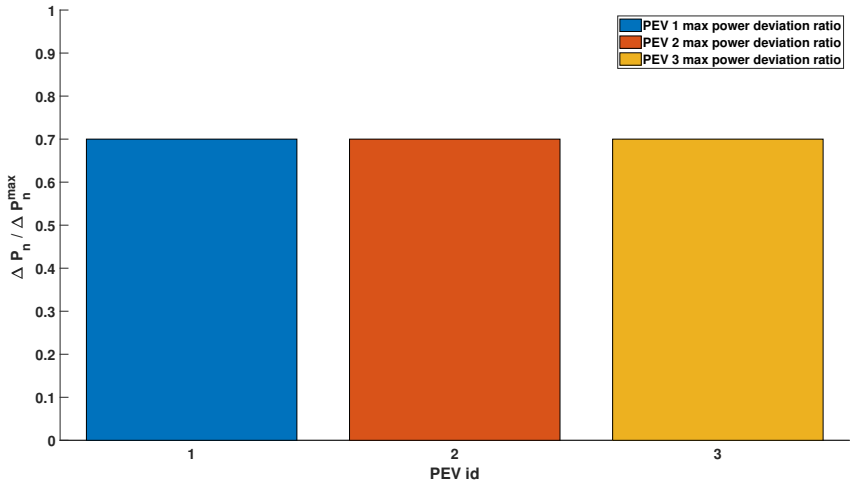


Figure 8.5. Scenario 1, balanced conditions: fraction of the maximum PEV power margin used for each PEV.

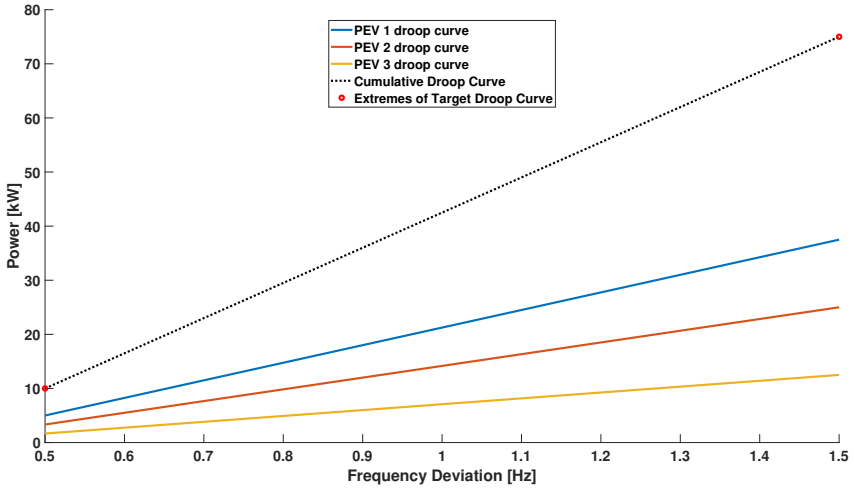


Figure 8.6. Scenario 1, different power margins: resulting local and global droop curves (request of 50% of the overall power margins).

Typically, in smart charging, different charging power set-points are assigned to each charging session. As explained, this leads to different power margins, i.e., different levels of maximum contribution that could be provided by each PEV. This unbalance is considered by the proposed algorithm during the composition of the local droop curves. This is shown in the next simulation, which assumes three charging sessions characterized by different power margins, as summarized in Table 8.1.

Table 8.1. Charging sessions of the first scenario.

PEV ID	$P_{n,k}$ [kW]	P_n^{max} [kW]	e_k [%]	d_k [%]
1	75	150	10	10
2	50	100	10	10
3	25	50	10	10

The global droop curve that has to be composed is characterized by a maximum droop value of 75 kW, i.e., 50% of the overall available margins. Note that all the charging sessions are performed at the 50% of maximum power, so they are characterized by different absolute power margins, $\Delta P_{n,k}$, but by the same percentage margin.

Fig. 8.6 displays the droop curves assigned to each vehicle. The algorithm allocates the curves in order to distribute the relative effort equally between the sessions. Fig. 8.7 highlights the equal distribution.

It is interesting to see the result of the same simulation, when the overall maximum power deviation increases. The same simulation is now performed with a maximum global droop curve value of 105 kW, i.e. 70% of available margins. The

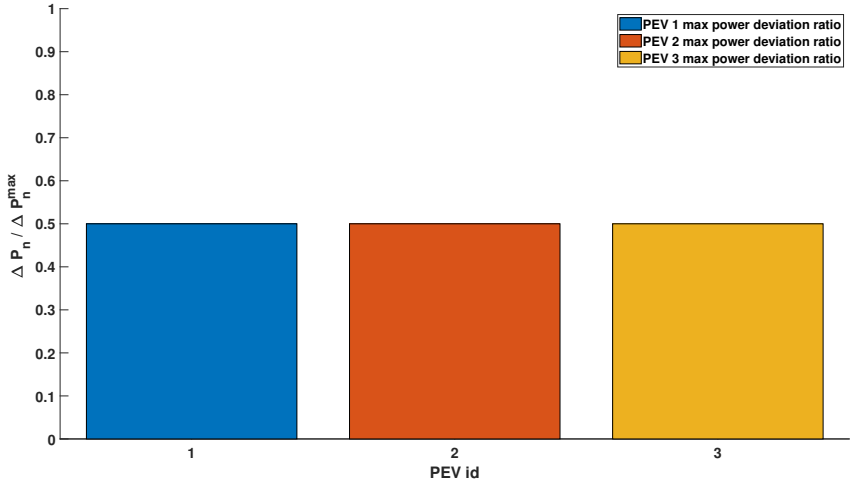


Figure 8.7. Scenario 1, different power margins: fraction of the maximum PEV power margin used for each PEV (request of 50% of the overall power margins).

increased request for power could be accomplished using the same strategy as before, so splitting the effort equally between the sessions. This strategy does not consider the fact that an equal distribution of power margins will impose a steeper droop curve on the sessions with larger margins, with a negative impact on the battery system. A steeper curve means a more aggressive response of the battery system to the frequency variations, with the consequent stress on power electronics and its effects on the battery temperature and cells health. The strategy presented in this work instead takes into account also the curve slope $m_{n,k}$ and the intercept $q_{n,k}$. As a result, the system distributes the additional effort needed to reach the new target droop curve less equally (Fig. 8.9) but fair distributes the power electronics stress (Fig. 8.8).

8.5.2 Scenario 2: Local Droop Curves Assignment in an Unbalanced Scenario

The interconnection between the smart charging and the frequency regulation service is done by considering the charging preferences expressed by the driver, i.e., the desired final state of charge and the charging time. Summarizing what already discussed about the target function (8.10), the algorithm gives priority in the assignment of droop curves to the PEVs with a greater state of charge error and/or a shorter dwelling time (i.e., to those PEVs that will benefit from an increase in the charging rate).

The simulations presented below will illustrate the capability of the designed algorithm to link the droop curve assignment with the charging status and user preferences.

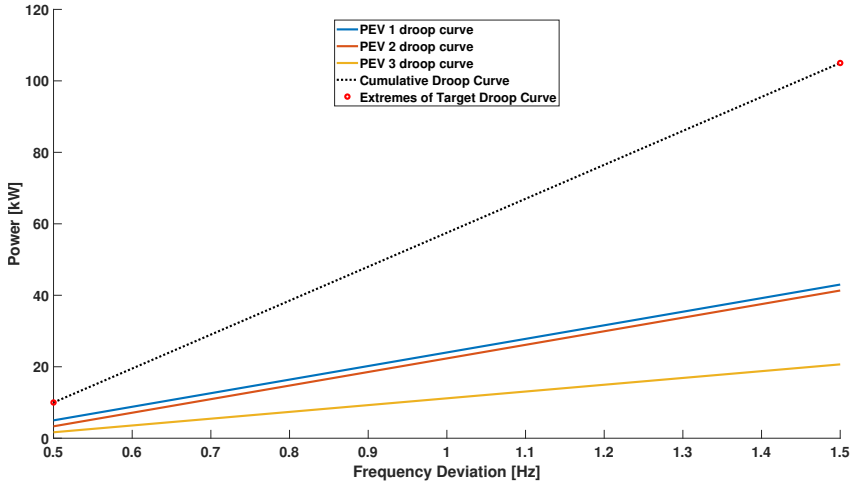


Figure 8.8. Scenario 1, different power margins: resulting local and global droop curves (request of 70% of the overall power margins).

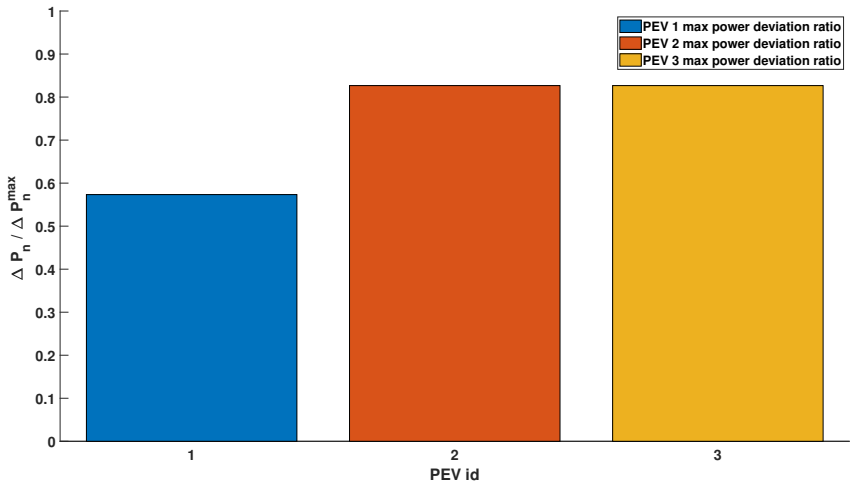


Figure 8.9. Scenario 1, different power margins: fraction of the maximum PEV power margin used for each PEV (request of 70% of the overall power margins).

Table 8.2. Charging sessions of the second scenario.

PEV ID	$P_{n,k}$ [kW]	P_n^{max} [kW]	e_k [%]	d_k [%]
1	150	100	80	10
2	150	100	40	10
3	150	100	30	10

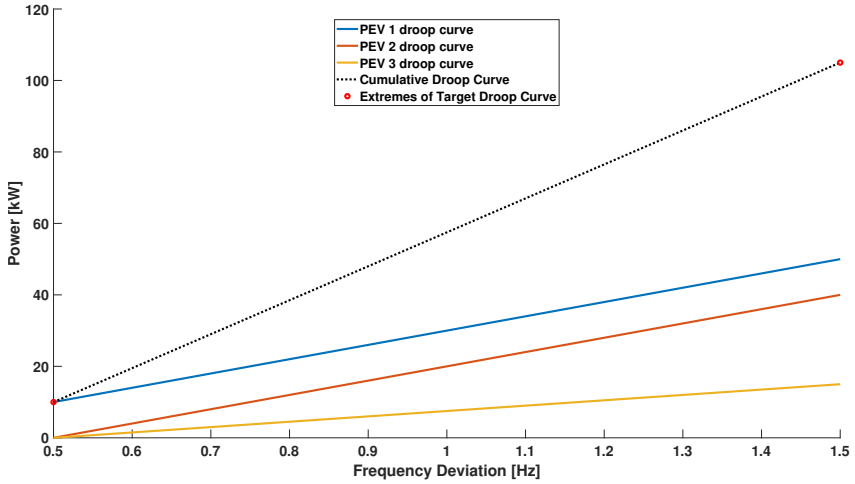


Figure 8.10. Scenario 2, balanced margins and unbalanced SOC errors: resulting local and global droop curves (request of 70% of the overall power margins).

The balanced scenario is now re-proposed, but this time with a different state of charge errors and dwelling times for the three PEVs. Table 8.2 summarizes the charging session status at the time when the proposed algorithm for allocating local droop curves is run.

PEV 1 is characterized by the grater SOC error, followed by PEV 2. The algorithm recognizes this condition and prioritizes the assignment of power to these two vehicles. The prioritization is in turn weighted by the respective SOC error.

The SOC error is one of the two charging preferences taken into account in the droop curve assignment. In the above simulation, the dwelling time was the same for the three PEVs. Fig.s 8.13 and 8.14 show the case in which the PEVs 1 charging session, characterized by a SOC error of 80%, is also characterized by a remaining dwelling time of 50%. In this case, the SOC error prioritization is compensated by the available dwelling time, so the algorithm identifies an optimal strategy to equally distribute the effort between the vehicles.

8.5.3 Notes on the Computational Complexity of the Algorithm

The optimization problem associated with the proposed algorithm is linear, with continuous variables. This is the category of optimization problems that is the

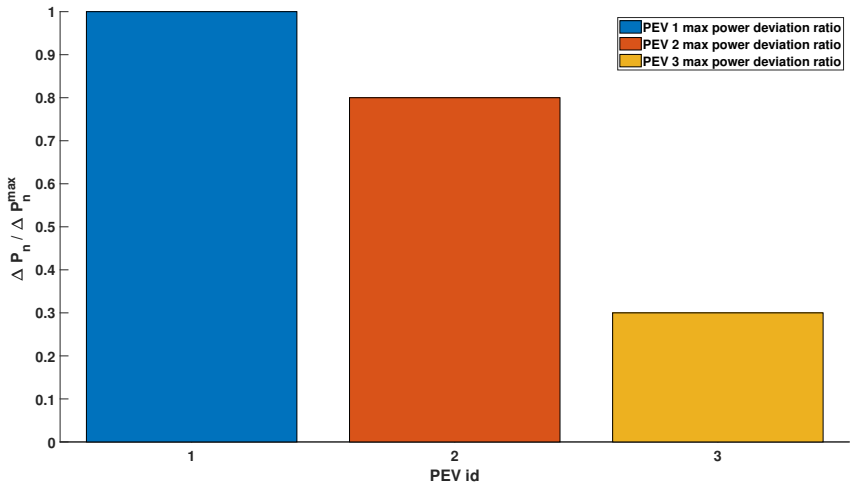


Figure 8.11. Scenario 2, balanced margins and unbalanced SOC errors: fraction of the maximum PEV power margin used for each PEV (request of 70% of the overall power margins).

most consolidated in the literature, and for which efficient and well-established optimization algorithms exist. The optimization problem was built in Julia and solved with Gurobi 9.1 [162].

Quadratic cost functions could be also used for this kind of problem, as they are an efficient choice in terms of effort distribution (e.g., the third term in the objective function could be replaced by the term $\alpha_3 m_{n,k}^2$, and similarly the last two terms). However, quadratic problems are more complex and require more computational resources than linear ones.

Considering that this algorithm is designed to work in conjunction with a smart charging system that updates the charging set points with a rate of minutes, and considering the amount of power needed for the participation to the frequency regulation services (e.g. in Italy, the Pilot Project *fast reserve* requires at least 5 MW of aggregated power [2]), the issue of finding the solution of a large scale optimization problem in a short amount of time has to be faced. For this reason, a linear formulation has been chosen. To test the scalability of the proposed algorithm, several simulations have been performed, considering scenarios of various dimensions. Fig. 8.12 reports the result of local droop curves computation in a scenario with 1000 PEVs participating in the frequency regulation. This simulation does not consider a specific charging technology, the maximum charging (and discharging) power rate for a generic charging session is randomly chosen between ±150 kW. Also, the user preferences are assigned randomly. The simulation shows how the proposed algorithm assigns different droop curves considering the current load area requirements (70% of the maximum available aggregated power margin), and the charging sessions status (the result is a uniform distribution of droop curves with no particular

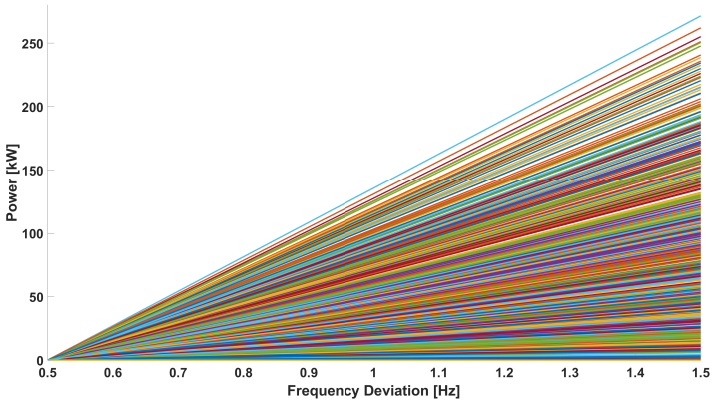


Figure 8.12. Simulation with 1000 PEVs.

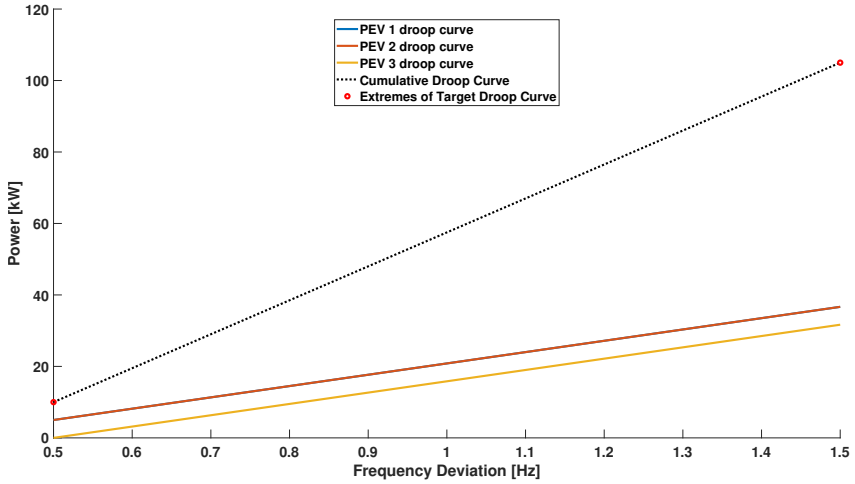


Figure 8.13. Scenario 2, balanced margins and unbalanced SOC errors and dwelling times: resulting local and global droop curves (request of 70% of the overall power margins).

clustering). As already mentioned, being the problem linear with continuous variables, very low computational time was observed in various experiments, always between 1-1.3 seconds.

Finally, the algorithm was tested with 100000 contemporary charging sessions. The solution time was always between 30 and 50 seconds, which is acceptable, in view of integrating the proposed algorithm with the smart charging algorithm.

8.6 Conclusions

This work has presented a novel control algorithm for enabling the participation of smart charging PEVs in the provisioning of frequency regulation services. This has a positive value for the grid, providing an additional source of flexibility to ensure

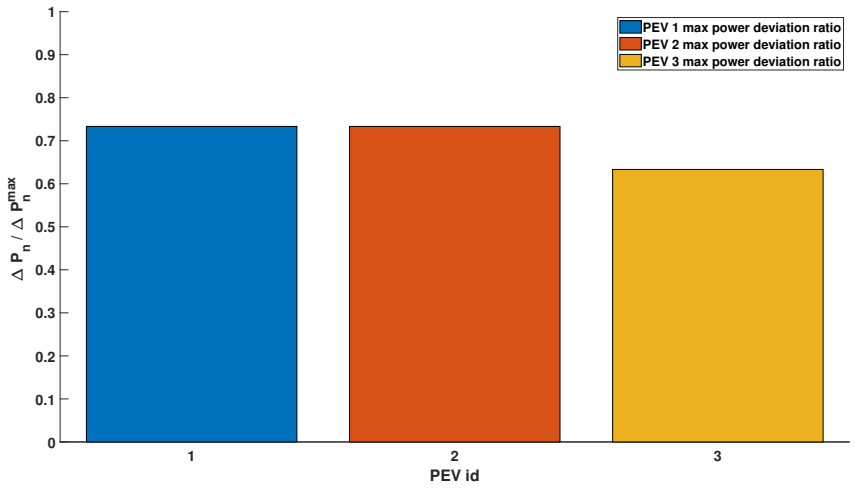


Figure 8.14. Scenario 2, balanced margins and unbalanced SOC errors and dwelling times: fraction of the maximum PEV power margin used for each PEV (request of 70% of the overall power margins).

grid stability, and for the PEV drivers as well, providing them an additional stream of revenues, which will lower the cost of ownership of the PEVs.

The proposed algorithm computes local frequency-power droop curves, one for each active charging session. These curves specify how the charging power setpoint should be changed in real-time, in response to frequency deviations from the nominal value. The local curves computation takes into account the real-time status of the charging sessions (in terms of time left until the end of the charging session, current state-of-charge, and energy left to charge), which is fundamental to ensure that the frequency regulation service is interdependent and harmonized with the smart charging service, and thus transparent to the PEV user. The superimposition of the local droop curves has to match a desired load area droop curve, which specifies how the PEVs in the load area should collectively react to a frequency deviation. In this way, the aggregate of PEVs can provide frequency regulation services to the market.

Numerical simulations have shown that the proposed algorithm is effective in assigning local droop curves in a fair way, which takes into account the different statuses of the charging sessions. Also, the algorithm scales well and is able to cope with aggregates of tens of thousands of PEVs.

Chapter 9

General Conclusion and Perspectives

This thesis collected two of the main research activities the Candidate has been focused in during the three years of his PhD. The Candidate had the opportunity to deepen several control methodologies to be applied in 5G NR radio access network and in multi-RAT systems, both for Connection Admission Control, Network Selection, and Traffic Splitting & Steering, thanks to his participation in the H2020 joint EU-Korea project 5G-ALLSTAR. At the same time, the Candidate had the opportunity to focus on applicative scenarios of 5G networks, in particular in the smart grids one thanks to his participation in the EU-funded H2020 project 5G-Solutions.

During the activities he carried out in the context of the 5G-ALLSTAR project, the Candidate studied several control methodologies to enable multi-connectivity in 5G heterogeneous networks. In particular, the problem of traffic splitting and steering has been deepened in the first work presented in this thesis. This work was based on the concept of Beckmann User equilibria and aims at equalizing some metrics on the network (i.e., the latency functions) among the different APs. In this particular case, the physical radio resources (i.e., PRBs) of the APs have been equalized by splitting the data flows from/to the UEs among different APs and steering the traffic dynamically between these APs to attain load balancing.

In the first work, no proper Connection Admission Control or Network Selection process is made. This means that, if the number of requests is too high, many UEs may suffer from network congestion. To this aim, the second study addressed the problem of selecting the APs to handle data flow requests from/to the UEs. In this first approach, an AP is selected for each UE connection request by the means of a Multi-Agent Reinforcement Learning methodology named Friend-or-Foe Q-Learning, which is based on the concept of Nash equilibria in Game Theory.

The third study tackles the same problem, but using Deep Reinforcement Learn-

ing methodology to select the best APs to maximize the users' perceived QoE. Some QoE models have been provided, to model typical user behaviors during Internet connection, and the DQN algorithm has been trained by using a reward function that depends on the estimated QoE of the current connection request by a UE.

All three works have been tested against a 5G heterogeneous radio access network simulator developed by the Candidate, to validate the proposed algorithms against realistic scenarios. The simulator has been developed starting from the relevant standards on 5G NR, 4G LTE, and Satellite radio access networks and it has been released as open-source software to the research community.

The fourth study presents a vision on 6G radio access networks and its possible new functionalities, such as mobile APs (e.g., UAVs, HAPSs, LAPSs, GEO/LEO satellites). In this work, the initial feasibility of the concept and the improvement of radio access network performances by using mobile APs as relay and as base stations has been demonstrated, also by using the over-mentioned radio access network simulator, that has been adapted to support these new functionalities.

Among the activities the Candidate carried out in the context of the 5G-Solutions project, an applicative scenario of 5G networks for the provisioning of Frequency Regulation services on smart grids, detailed in the second part of this thesis, has been proposed. An introduction to the problem of Frequency Regulation in electricity grids has been provided and the main idea of using charging PEVs to provide such kind of services to the electricity network has been detailed in the fifth study. In this work, a control architecture that fits the system scenario has been proposed, together with three proposed control approaches of increasing complexity for the computation of the power-frequency (droop) curves for each PEV in a smart charging scenario, and an assessment on the delay budget required to make the approach feasible according to the current Italian regulatory framework.

In the last study, a control algorithm based on linear optimization problem formulation for the computation of droop curves for each PEV (namely, local droop curves) has been proposed. The control algorithm detailed in this work computes a set of droop curves (one for each charging PEV) so to differentiate the contribution of each PEV based on its actual and desired state-of-charge and on its remaining dwelling time, while composing a global droop curve (i.e., the sum of all the local droop curves) that is equal to the one agreed with the TSO. The optimal solution is re-computed at each time step, due to the evolution of the state-of-charge of the PEVs (which may be subject to disturbances due to conversion losses, but also due to previous activations of the Frequency Regulation service) and their remaining dwelling time.

The Candidate is currently continuing the study on 5G multi-RAT radio access

network control, in particular for guaranteeing connectivity in emergencies (like wildfires and landslides), by dynamically splitting the traffic generated by the UEs among the available APs; moreover, he is deepening on the problem of physical radio resources allocation in 5G networks to attain low-latency and increased bitrate at the user level, while guaranteeing defined QoS levels. Moreover, the Candidate is going through the implementation of the proposed approach for Frequency Regulation service using PEVs in the context of the 5G-Solutions project and in collaboration with Enel X Way, to demonstrate the proposed control architecture and algorithms by using a real 5G network in the city of Turin (provided by TIM) and real DC charging stations connected to real PEVs. Finally, the Candidate is also working on optimal control approaches to enable fast charging in service areas equipped with renewable energy sources and energy storage systems and on algorithms for day-ahead allocation and real-time control of an aggregate of PEVs to participate in the ancillary service market for the provisioning of ancillary services to the electricity grid.

Bibliography

[1] A. Gavras, O. Bulakci, M. Gramaglia, M. Iordache, M. Ghoraishi, A. Garcia, T. Cogalan, J. Gutiérrez, A. Tzanakaki, D. Warren, X. Li, G. Landi, J. Mangues, K. Tsagkaris, V. Frascolla, and H. Lee, "5g ppp architecture working group - view on 5g architecture, version 4.0," 2021. [Online]. Available: https://zenodo.org/record/5155657

[2] TERNA. Pilot project fast reserve. [Online]. Available: https://www.terna.it/en/electric-system/pilot-projects-pursuant-arera-resolution-300-2017-reel/fast-reserve-pilot-project

[3] 3GPP, "NR; Physical channels and modulation," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.211, 09 2019, version 15.7.1. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3213

[4] M. Sauter, *From GSM to LTE-Advanced; An Introduction to Mobile Networks and Mobile Broadband*. Chichester, UK: John Wiley & Sons, Ltd, Jun. 2014.

[5] F. Capozzi, G. Piro, L. Grieco, G. Boggia, and P. Camarda, "Downlink packet scheduling in LTE cellular networks: Key design issues and a survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013. [Online]. Available: https://doi.org/10.1109/surv.2012.060912.00100

[6] H. Chen, R. Abbas, P. Cheng, M. Shirvanimoghaddam, W. Hardjawana, W. Bao, Y. Li, and B. Vucetic, "Ultra-reliable low latency cellular networks: Use cases, challenges and approaches," *IEEE Communications Magazine*, vol. 56, no. 12, pp. 119–125, Dec. 2018. [Online]. Available: https://doi.org/10.1109/mcom.2018.1701178

[7] K.-C. Chen, T. Zhang, R. D. Gitlin, and G. Fettweis, "Ultra-low latency mobile networking," *IEEE Network*, vol. 33, no. 2, pp. 181–187, Mar. 2019. [Online]. Available: https://doi.org/10.1109/mnet.2018.1800011

[8] M. Fei and P. Fan, "Position-assisted fast handover schemes for LTE-advanced network under high mobility scenarios," *Journal of Modern Transportation*, vol. 20, no. 4, pp. 268–273, Dec. 2012. [Online]. Available: https://doi.org/10.1007/bf03325809

[9] P. Fan, J. Zhao, and C.-L. I, "5g high mobility wireless communications: Challenges and solutions," *China Communications*, vol. 13, no. 2, pp. 1–13, 2016. [Online]. Available: https://doi.org/10.1109/cc.2016.7405718

[10] R. D. Gaudenzi, P. Angeletti, D. Petrolati, and E. Re, "Future technologies for very high throughput satellite systems," *International Journal of Satellite Communications and Networking*, vol. 38, no. 2, pp. 141–161, Oct. 2019. [Online]. Available: https://doi.org/10.1002/sat.1327

[11] A. Pérez-Neria, M. A. Lagunas, and M. A. Vázquez, "High throughput satellites in 5g and MIMO interference limited communications," *MATEC Web of Conferences*, vol. 76, p. 03008, 2016. [Online]. Available: https://doi.org/10.1051/matecconf/20167603008

[12] H. Kameda, J. Li, C. Kim, and Y. Zhang, *Optimal Load Balancing in Distributed Computer Systems*. Springer London, 1997. [Online]. Available: https://doi.org/10.1007/978-1-4471-0969-3

[13] J. R. Correa, A. S. Schulz, and N. E. Stier-Moses, "Selfish routing in capacitated networks," *Mathematics of Operations Research*, vol. 29, no. 4, pp. 961–976, Nov. 2004. [Online]. Available: https://doi.org/10.1287/moor.1040.0098

[14] M. E. Helou, M. Ibrahim, S. Lahoud, K. Khawam, D. Mezher, and B. Cousin, "A network-assisted approach for RAT selection in heterogeneous cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 6, pp. 1055–1067, Jun. 2015. [Online]. Available: https://doi.org/10.1109/jsac.2015.2416987

[15] A. Morgado, K. M. S. Huq, S. Mumtaz, and J. Rodriguez, "A survey of 5g technologies: regulatory, standardization and industrial perspectives," *Digital Communications and Networks*, vol. 4, no. 2, pp. 87–97, Apr. 2018. [Online]. Available: https://doi.org/10.1016/j.dcan.2017.09.010

[16] I. D. Silva, G. Mildh, J. Rune, P. Wallentin, J. Vikberg, P. Schliwa-Bertling, and R. Fan, "Tight integration of new 5g air interface and LTE to fulfill 5g requirements," in *2015 IEEE 81st Vehicular*

Technology Conference (VTC Spring). IEEE, May 2015. [Online]. Available: https://doi.org/10.1109/vtcspring.2015.7146134

[17] L. Wang and G.-S. G. Kuo, "Mathematical modeling for network selection in heterogeneous wireless networks — a tutorial," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 271–292, 2013. [Online]. Available: https://doi.org/10.1109/surv.2012.010912.00044

[18] J.-O. Kim, "Feedback-based traffic splitting for wireless terminals with multi-radio devices," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 476–482, May 2010. [Online]. Available: https://doi.org/10.1109/tce.2010.5505958

[19] A. A. Sabbagh, R. Braun, and M. Abolhasan, "A comprehensive survey on rat selection algorithms for heterogeneous networks," 2011. [Online]. Available: https://zenodo.org/record/1080882

[20] N. Zhang, S. Zhang, S. Wu, J. Ren, J. W. Mark, and X. Shen, "Beyond coexistence: Traffic steering in LTE networks with unlicensed bands," *IEEE Wireless Communications*, vol. 23, no. 6, pp. 40–46, Dec. 2016. [Online]. Available: https://doi.org/10.1109/mwc.2016.1600059wc

[21] M. Dryjanski and M. Szydelko, "A unified traffic steering framework for LTE radio access network coordination," *IEEE Communications Magazine*, vol. 54, no. 7, pp. 84–92, Jul. 2016. [Online]. Available: https://doi.org/10.1109/mcom.2016.7509383

[22] Q.-T. Nguyen-Vuong, Y. Ghamri-Doudane, and N. Agoulmine, "On utility models for access network selection in wireless heterogeneous networks," in *NOMS 2008 - 2008 IEEE Network Operations and Management Symposium*. IEEE, 2008. [Online]. Available: https://doi.org/10.1109/noms.2008.4575128

[23] H. Chan, P. Fan, and Z. Cao, "A utility-based network selection scheme for multiple services in heterogeneous networks," in *2005 International Conference on Wireless Networks, Communications and Mobile Computing*. IEEE, 2005. [Online]. Available: https://doi.org/10.1109/wirles.2005.1549578

[24] M. Cesana, N. Gatti, and I. Malanchini, "Game theoretic analysis of wireless access network selection: Models, inefficiency bounds, and algorithms," in *Proceedings of the 3rd International Conference on Performance Evaluation Methodologies and Tools*. ICST, 2008. [Online]. Available: https://doi.org/10.4108/icst.valuetools2008.4341

[25] J. Antoniou and A. Pitsillides, "4g converged environment: Modeling network selection as a game," in *2007 16th IST Mobile and Wireless Communications Summit*. IEEE, Jul. 2007. [Online]. Available: https://doi.org/10.1109/istmwc.2007.4299242

[26] D. Niyato and E. Hossain, "Dynamics of network selection in heterogeneous wireless networks: An evolutionary game approach," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 2008–2017, May 2009. [Online]. Available: https://doi.org/10.1109/tvt.2008.2004588

[27] T. Roughgarden and É. Tardos, "How bad is selfish routing?" *Journal of the ACM*, vol. 49, no. 2, pp. 236–259, Mar. 2002. [Online]. Available: https://doi.org/10.1145/506147.506153

[28] J. G. Wardrop, "Some theoretical aspects of road traffic reseach," *Proceedings of the Institution of Civil Engineers*, vol. 1, no. 3, pp. 325–362, May 1952. [Online]. Available: https://doi.org/10.1680/ipeds.1952.11259

[29] T. Roughgarden, *Selfish routing and the price of anarchy*. MIT press, 2005.

[30] H. Z. Aashtiani, H. Poorzahedy, and M. Nourinejad, "Extending wardrop's first principle for capacitated networks," *Transportmetrica A: Transport Science*, pp. 1–30, Jul. 2018. [Online]. Available: https://doi.org/10.1080/23249935.2018.1504255

[31] A. S. Schulz and N. S. Moses, "On the performance of user equilibria in traffic networks," in *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, ser. SODA '03. USA: Society for Industrial and Applied Mathematics, 2003, p. 86–87.

[32] P. Marcotte, S. Nguyen, and A. Schoeb, "A strategic flow model of traffic assignment in static capacitated networks," *Operations Research*, vol. 52, no. 2, pp. 191–212, Apr. 2004. [Online]. Available: https://doi.org/10.1287/opre.1030.0091

[33] H. Kameda, J. Li, C. Kim, and Y. Zhang, *Optimal Load Balancing in Distributed Computer Systems*, 1st ed. Springer Publishing Company, Incorporated, 2011.

[34] D. Grosu and A. T. Chronopoulos, "Noncooperative load balancing in distributed systems," *Journal of Parallel and Distributed Computing*, vol. 65, no. 9, pp. 1022–1034, Sep. 2005. [Online]. Available: https://doi.org/10.1016/j.jpdc.2005.05.001

[35] E. Altman, H. Kameda, and Y. Hosokawa, "Nash equilibria in load balancing in distributed computer systems," *International Game Theory Review*, vol. 04, no. 02, pp. 91–100, Jun. 2002. [Online]. Available: https://doi.org/10.1142/s0219198902000574

[36] R. J. Smeed, M. Beckmann, C. B. McGuire, C. B. Winsten, and T. C. Koopmans, "Studies in the economics of transportation." *The Economic Journal*, vol. 67, no. 265, p. 116, Mar. 1957. [Online]. Available: https://doi.org/10.2307/2227641

[37] G. Como, K. Savla, D. Acemoglu, M. A. Dahleh, and E. Frazzoli, "Robust distributed routing in dynamical networks—part i: Locally responsive policies and weak resilience," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 317–332, Feb. 2013. [Online]. Available: https://doi.org/10.1109/tac.2012.2209951

[38] ——, "Robust distributed routing in dynamical networks–part II: Strong resilience, equilibrium selection and cascaded failures," *IEEE Transactions on Automatic Control*, vol. 58, no. 2, pp. 333–348, Feb. 2013. [Online]. Available: https://doi.org/10.1109/tac.2012.2209975

[39] V. Borkar and P. Kumar, "Dynamic cesaro-wardrop equilibration in networks," *IEEE Transactions on Automatic Control*, vol. 48, no. 3, pp. 382–396, Mar. 2003. [Online]. Available: https://doi.org/10.1109/tac.2003.809145

[40] D. Barth, O. Bournez, O. Boussaton, and J. Cohen, "Distributed learning of wardrop equilibria," in *Unconventional Computing*. Springer Berlin Heidelberg, 2008, pp. 19–32. [Online]. Available: https://doi.org/10.1007/978-3-540-85194-3_5

[41] S. Fischer, L. Olbrich, and B. Vöcking, "Approximating wardrop equilibria with finitely many agents," *Distributed Computing*, vol. 21, no. 2, pp. 129–139, Mar. 2008. [Online]. Available: https://doi.org/10.1007/s00446-008-0057-1

[42] S. Fischer and B. Vöcking, "Adaptive routing with stale information," *Theoretical Computer Science*, vol. 410, no. 36, pp. 3357–3371, Aug. 2009. [Online]. Available: https://doi.org/10.1016/j.tcs.2008.01.055

[43] S. Fischer, H. Räcke, and B. Vöcking, "Fast convergence to wardrop equilibria by adaptive sampling methods," *SIAM Journal on Computing*, vol. 39, no. 8, pp. 3700–3735, Jan. 2010. [Online]. Available: https://doi.org/10.1137/090746720

[44] A. Pietrabissa, L. R. Celsi, F. Cimorelli, V. Suraci, F. D. Priscoli, A. D. Giorgio, A. Giuseppi, and S. Monaco, "Lyapunov-based design of a distributed wardrop load-balancing algorithm with application to software-defined networking," *IEEE Transactions on Control Systems Technology*, vol. 27, no. 5, pp. 1924–1936, Sep. 2019. [Online]. Available: https://doi.org/10.1109/tcst.2018.2842044

[45] A. Giuseppi and A. Pietrabissa, "Wardrop equilibrium in discrete-time selfish routing with time-varying bounded delays," *IEEE Transactions on Automatic Control*, vol. 66, no. 2, pp. 526–537, Feb. 2021. [Online]. Available: https://doi.org/10.1109/tac.2020.2981906

[46] V. Sundarapandian, "An invariance principle for discrete-time nonlinear systems," *Applied Mathematics Letters*, vol. 16, no. 1, pp. 85–91, Jan. 2003. [Online]. Available: https://doi.org/10.1016/s0893-9659(02)00148-9

[47] W. Mei and F. Bullo, "Lasalle invariance principle for discrete-time dynamical systems: A concise and self-contained tutorial," 2017. [Online]. Available: https://arxiv.org/abs/1710.03710

[48] H. Zhang, N. Liu, X. Chu, K. Long, A.-H. Aghvami, and V. C. M. Leung, "Network slicing based 5g and future mobile networks: Mobility, resource management, and challenges," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 138–145, Aug. 2017. [Online]. Available: https://doi.org/10.1109/mcom.2017.1600940

[49] I. Milchtaich, "Topological conditions for uniqueness of equilibrium in networks," *Mathematics of Operations Research*, vol. 30, no. 1, pp. 225–244, Feb. 2005. [Online]. Available: https://doi.org/10.1287/moor.1040.0122

[50] 3GPP, "Study on access traffic steering, switch and splitting support in the 5G System (5GS) architecture," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 23.793, 12 2018, version 16.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3254

[51] E. De Santis, A. Giuseppi, A. Pietrabissa, and F. D. Priscoli, "trunk96/wireless-network-simulator-v2: v0.2," 2021. [Online]. Available: https://zenodo.org/record/5558396

[52] N. Tabia, A. Gondran, O. Baala, and A. Caminada, "Interference model and evaluation in LTE networks," in *2011 4th Joint IFIP Wireless and*

Mobile Networking Conference (WMNC 2011). IEEE, Oct. 2011. [Online]. Available: https://doi.org/10.1109/wmnc.2011.6097237

[53] 3GPP, "NR; User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.101-1, 10 2018, version 15.3.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3283

[54] ——, "NR; User Equipment (UE) radio transmission and reception; Part 2: Range 2 Standalone," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.101-2, 10 2018, version 15.3.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3284

[55] G. Maral, M. Bousquet, and Z. Sun, *Satellite Communications Systems*. Wiley, Apr. 2020. [Online]. Available: https://doi.org/10.1002/9781119673811

[56] H. Fenech, S. Amos, A. Tomatis, and V. Soumpholphakdy, "High throughput satellite systems: An analytical approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 1, pp. 192–202, Jan. 2015. [Online]. Available: https://doi.org/10.1109/taes.2014.130450

[57] R. Swinford and B. Grau, "High Throughput Satellites: Delivering future capacity needs," pp. 1–20, 2015.

[58] S. Andreev, M. Gerasimenko, O. Galinina, Y. Koucheryavy, N. Himayat, S.-P. Yeh, and S. Talwar, "Intelligent access network selection in converged multi-radio heterogeneous networks," *IEEE Wireless Communications*, vol. 21, no. 6, pp. 86–96, Dec. 2014. [Online]. Available: https://doi.org/10.1109/mwc.2014.7000976

[59] L. Hui, W. Ma, and S. Zhai, "A novel approach for radio resource management in multi-dimensional heterogeneous 5g networks," *Journal of Communications and Information Networks*, vol. 1, no. 2, pp. 77–83, Aug. 2016. [Online]. Available: https://doi.org/10.1007/bf03391559

[60] A. Wilson, A. Lenaghan, and R. Malyan, "Optimising wireless access network selection to maintain qos in heterogeneous wireless environments," in *wireless personal multimedia communications*. Citeseer, 2005, pp. 18–22.

[61] X. Gelabert, J. Perez-Romero, O. Sallent, and R. Agusti, "A markovian approach to radio access technology selection in heterogeneous

multiaccess/multiservice wireless networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 10, pp. 1257–1270, Oct. 2008. [Online]. Available: https://doi.org/10.1109/tmc.2008.50

[62] N. Vučević, J. Pérez-Romero, O. Sallent, and R. Agustí, "Reinforcement learning for joint radio resource management in LTE-UMTS scenarios," *Computer Networks*, vol. 55, no. 7, pp. 1487–1497, May 2011. [Online]. Available: https://doi.org/10.1016/j.comnet.2010.12.029

[63] M. L. Littman, "Friend-or-foe q-learning in general-sum games," *Morgan Kaufmann Publishers Inc.*, 2003.

[64] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, 2nd ed. The MIT Press, 2018. [Online]. Available: http://incompleteideas.net/book/the-book-2nd.html

[65] R. Ahuja, "Minimax linear programming problem," *Operations Research Letters*, vol. 4, no. 3, pp. 131–134, Oct. 1985. [Online]. Available: https://doi.org/10.1016/0167-6377(85)90017-3

[66] "5G-ALLSTAR Project." [Online]. Available: https://5g-allstar.eu/

[67] 3GPP, "Study on New Radio (NR) to support non-terrestrial networks," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 38.811, 10 2020, version 15.4.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3234

[68] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing atari with deep reinforcement learning," 2013. [Online]. Available: https://arxiv.org/abs/1312.5602

[69] K. S. S. Anupama, S. S. Gowri, and B. P. Rao, "A comparative study of outranking MADM algorithms in network selection," in *2018 Second International Conference on Computing Methodologies and Communication (ICCMC)*. IEEE, Feb. 2018. [Online]. Available: https://doi.org/10.1109/iccmc.2018.8487931

[70] Y. Zhong, H. Wang, and H. Lv, "A cognitive wireless networks access selection algorithm based on MADM," *Ad Hoc Networks*, vol. 109, p. 102286, Dec. 2020. [Online]. Available: https://doi.org/10.1016/j.adhoc.2020.102286

[71] S. Radouche, C. Leghris, and A. Adib, "MADM methods based on utility function and reputation for access network selection in a multi-access

mobile network environment," in *2017 International Conference on Wireless Networks and Mobile Communications (WINCOM)*. IEEE, Nov. 2017. [Online]. Available: https://doi.org/10.1109/wincom.2017.8238177

[72] Q. Song and A. Jamalipour, "Network selection in an integrated wireless LAN and UMTS environment using mathematical modeling and computing techniques," *IEEE Wireless Communications*, vol. 12, no. 3, pp. 42–48, Jun. 2005. [Online]. Available: https://doi.org/10.1109/mwc.2005.1452853

[73] T. Ding, L. Liang, M. Yang, and H. Wu, "Multiple attribute decision making based on cross-evaluation with uncertain decision parameters," *Mathematical Problems in Engineering*, vol. 2016, pp. 1–10, 2016. [Online]. Available: https://doi.org/10.1155/2016/4313247

[74] R. K. Goyal, S. Kaushal, and A. K. Sangaiah, "The utility based non-linear fuzzy AHP optimization model for network selection in heterogeneous wireless networks," *Applied Soft Computing*, vol. 67, pp. 800–811, Jun. 2018. [Online]. Available: https://doi.org/10.1016/j.asoc.2017.05.026

[75] X. Yan, P. Dong, T. Zheng, and H. Zhang, "Fuzzy and utility based network selection for heterogeneous networks in high-speed railway," *Wireless Communications and Mobile Computing*, vol. 2017, pp. 1–14, 2017. [Online]. Available: https://doi.org/10.1155/2017/4967438

[76] M.-M. R. Mou and M. Z. Chowdhury, "Service aware fuzzy logic based handover decision in heterogeneous wireless networks," in *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, Feb. 2017. [Online]. Available: https://doi.org/10.1109/ecace.2017.7912992

[77] R. Trestian, O. Ormond, and G.-M. Muntean, "Game theory-based network selection: Solutions and challenges," *IEEE Communications Surveys & Tutorials*, vol. 14, no. 4, pp. 1212–1231, 2012. [Online]. Available: https://doi.org/10.1109/surv.2012.010912.00081

[78] M. T. Rahman, M. Z. Chowdhury, and Y. M. Jang, "Radio access network selection mechanism based on hierarchical modelling and game theory," in *2016 International Conference on Information and Communication Technology Convergence (ICTC)*. IEEE, Oct. 2016. [Online]. Available: https://doi.org/10.1109/ictc.2016.7763451

[79] L. Rajesh, K. B. Bagan, and B. Ramesh, "User demand wireless network selection using game theory," in *Lecture Notes in Electrical*

Engineering. Springer Singapore, 2017, pp. 39–53. [Online]. Available: https://doi.org/10.1007/978-981-10-2999-8_4

[80] Meenakshi and N. P. Singh, "A comparative study of cooperative and non-cooperative game theory in network selection," in *2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*. IEEE, Mar. 2016. [Online]. Available: https://doi.org/10.1109/icctict.2016.7514652

[81] Z.-H. Zhang, X.-F. Jiang, and H.-S. Xi, "Optimal content placement and request dispatching for cloud-based video distribution services," *International Journal of Automation and Computing*, vol. 13, no. 6, pp. 529–540, Oct. 2016. [Online]. Available: https://doi.org/10.1007/s11633-016-1025-z

[82] F.-S. Lin, B.-Q. Yin, J. Huang, and X.-M. Wu, "Admission control with elastic QoS for video on demand systems," *International Journal of Automation and Computing*, vol. 9, no. 5, pp. 467–473, Oct. 2012. [Online]. Available: https://doi.org/10.1007/s11633-012-0668-7

[83] Z. Du, C. Wang, Y. Sun, and G. Wu, "Context-aware indoor VLC/RF heterogeneous network selection: Reinforcement learning with knowledge transfer," *IEEE Access*, vol. 6, pp. 33 275–33 284, 2018. [Online]. Available: https://doi.org/10.1109/access.2018.2844882

[84] Y. Yang, Y. Wang, K. Liu, N. Zhang, S. Gu, and Q. Zhang, "Deep reinforcement learning based online network selection in CRNs with multiple primary networks," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7691–7699, Dec. 2020. [Online]. Available: https://doi.org/10.1109/tii.2020.2971735

[85] D. D. Nguyen, H. X. Nguyen, and L. B. White, "Reinforcement learning with network-assisted feedback for heterogeneous RAT selection," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6062–6076, Sep. 2017. [Online]. Available: https://doi.org/10.1109/twc.2017.2718526

[86] F. Liberati, A. Giuseppi, A. Pietrabissa, V. Suraci, A. D. Giorgio, M. Trubian, D. Dietrich, P. Papadimitriou, and F. D. Priscoli, "Stochastic and exact methods for service mapping in virtualized network infrastructures," *International Journal of Network Management*, vol. 27, no. 6, p. e1985, Jul. 2017. [Online]. Available: https://doi.org/10.1002/nem.1985

[87] X. Wang, J. Li, L. Wang, C. Yang, and Z. Han, "Intelligent user-centric network selection: A model-driven reinforcement learning framework,"

IEEE Access, vol. 7, pp. 21 645–21 661, 2019. [Online]. Available: https://doi.org/10.1109/access.2019.2898205

[88] K.-S. Shin, G.-H. Hwang, and O. Jo, "Distributed reinforcement learning scheme for environmentally adaptive IoT network selection," *Electronics Letters*, vol. 56, no. 9, pp. 462–464, Apr. 2020. [Online]. Available: https://doi.org/10.1049/el.2019.3891

[89] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," 2015. [Online]. Available: https://arxiv.org/abs/1509.02971

[90] Y. Zhou, Z. M. Fadlullah, B. Mao, and N. Kato, "A deep-learning-based radio resource assignment technique for 5g ultra dense networks," *IEEE Network*, vol. 32, no. 6, pp. 28–34, Nov. 2018. [Online]. Available: https://doi.org/10.1109/mnet.2018.1800085

[91] B. Mao, F. Tang, Y. Kawamoto, and N. Kato, "Optimizing computation offloading in satellite-UAV-served 6g IoT: A deep learning approach," *IEEE Network*, vol. 35, no. 4, pp. 102–108, Jul. 2021. [Online]. Available: https://doi.org/10.1109/mnet.011.2100097

[92] F. Delli Priscoli, A. Giuseppi, F. Liberati, and A. Pietrabissa, "Traffic steering and network selection in 5g networks based on reinforcement learning," in *2020 European Control Conference (ECC)*. IEEE, May 2020. [Online]. Available: https://doi.org/10.23919/ecc51009.2020.9143837

[93] "Final report for COST Action 231," Tech. Rep., 1996. [Online]. Available: http://www.lx.it.pt/cost231/final_report.htm

[94] 3GPP, "Service requirements for the 5G system," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 26.261, 03 2020, version 17.2.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3107

[95] ——, "Study on using satellite access in 5G," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 22.822, 07 2018, version 16.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3372

[96] ——, "Unmanned Aerial System (UAS) support in 3GPP," 3rd Generation Partnership Project (3GPP), Technical Specification (TS) 22.822, 12 2019,

version 17.1.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3545

[97] J. Kim, G. Casati, A. Pietrabissa, A. Giuseppi, E. C. Strinati, N. Cassiau, G. Noh, H. Chung, I. Kim, M. Thary, J.-M. Houssin, F. Pigni, S. Colombero, P. D. Zotto, L. Raschkowski, and S. Jaeckel, "5g-ALLSTAR: An integrated satellite-cellular system for 5g and beyond," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, Apr. 2020. [Online]. Available: https://doi.org/10.1109/wcncw48565.2020.9124751

[98] E. C. Strinati, M. Mueck, A. Clemente, J. Kim, G. Noh, H. Chung, I. Kim, T. Choi, Y. Kim, H. K. Chung, G. Destino, A. Pärssinen, N. Chuberre, B. Vautherin, T. Deleu, M. Gineste, and A. Korvala, "5gchampion - disruptive 5g technologies for roll-out in 2018," *ETRI Journal*, vol. 40, no. 1, pp. 10–25, Feb. 2018. [Online]. Available: https://doi.org/10.4218/etrij.2017-0237

[99] S. C. Arum, D. Grace, and P. D. Mitchell, "A review of wireless communication using high-altitude platforms for extended coverage and capacity," *Computer Communications*, vol. 157, pp. 232–256, May 2020. [Online]. Available: https://doi.org/10.1016/j.comcom.2020.04.020

[100] S. Chandrasekharan, K. Gomez, A. Al-Hourani, S. Kandeepan, T. Rasheed, L. Goratti, L. Reynaud, D. Grace, I. Bucaille, T. Wirth, and S. Allsopp, "Designing and implementing future aerial communication networks," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 26–34, May 2016. [Online]. Available: https://doi.org/10.1109/mcom.2016.7470932

[101] A. Fotouhi, H. Qiang, M. Ding, M. Hassan, L. G. Giordano, A. Garcia-Rodriguez, and J. Yuan, "Survey on UAV cellular communications: Practical aspects, standardization advancements, regulation, and security challenges," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3417–3442, 2019. [Online]. Available: https://doi.org/10.1109/comst.2019.2906228

[102] Y. Zeng, R. Zhang, and T. J. Lim, "Wireless communications with unmanned aerial vehicles: opportunities and challenges," *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, May 2016. [Online]. Available: https://doi.org/10.1109/mcom.2016.7470933

[103] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Communications Letters*,

vol. 21, no. 3, pp. 604–607, Mar. 2017. [Online]. Available: https://doi.org/10.1109/lcomm.2016.2633248

[104] X. Li, H. Yao, J. Wang, X. Xu, C. Jiang, and L. Hanzo, "A near-optimal UAV-aided radio coverage strategy for dense urban areas," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 9098–9109, Sep. 2019. [Online]. Available: https://doi.org/10.1109/tvt.2019.2927425

[105] E. Kalantari, M. Z. Shakir, H. Yanikomeroglu, and A. Yongacoglu, "Backhaul-aware robust 3d drone placement in 5g+ wireless networks," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, May 2017. [Online]. Available: https://doi.org/10.1109/iccw.2017.7962642

[106] L. Wang, Y. L. Che, J. Long, L. Duan, and K. Wu, "Multiple access MmWave design for UAV-aided 5g communications," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 64–71, Feb. 2019. [Online]. Available: https://doi.org/10.1109/mwc.2018.1800216

[107] J. Plachy, Z. Becvar, P. Mach, R. Marik, and M. Vondra, "Joint positioning of flying base stations and association of users: Evolutionary-based approach," *IEEE Access*, vol. 7, pp. 11 454–11 463, 2019. [Online]. Available: https://doi.org/10.1109/access.2019.2892564

[108] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5g and beyond," *Proceedings of the IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019. [Online]. Available: https://doi.org/10.1109/jproc.2019.2952892

[109] M. Mozaffari, A. T. Z. Kasgari, W. Saad, M. Bennis, and M. Debbah, "Beyond 5g with UAVs: Foundations of a 3d wireless cellular network," *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 357–372, Jan. 2019. [Online]. Available: https://doi.org/10.1109/twc.2018.2879940

[110] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "UAV-enabled communication using NOMA," *IEEE Transactions on Communications*, vol. 67, no. 7, pp. 5126–5138, Jul. 2019. [Online]. Available: https://doi.org/10.1109/tcomm.2019.2906622

[111] M. Gapeyenko, V. Petrov, D. Moltchanov, S. Andreev, N. Himayat, and Y. Koucheryavy, "Flexible and reliable UAV-assisted backhaul operation in 5g mmWave cellular networks," *IEEE Journal on Selected Areas in*

Communications, vol. 36, no. 11, pp. 2486–2496, Nov. 2018. [Online]. Available: https://doi.org/10.1109/jsac.2018.2874145

[112] S. Jeong, O. Simeone, and J. Kang, "Mobile edge computing via a UAV-mounted cloudlet: Optimization of bit allocation and path planning," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 3, pp. 2049–2063, Mar. 2018. [Online]. Available: https://doi.org/10.1109/tvt.2017.2706308

[113] J. Xiong, H. Guo, and J. Liu, "Task offloading in UAV-aided edge computing: Bit allocation and trajectory optimization," *IEEE Communications Letters*, vol. 23, no. 3, pp. 538–541, Mar. 2019. [Online]. Available: https://doi.org/10.1109/lcomm.2019.2891662

[114] J. Zhang, L. Zhou, Q. Tang, E. C.-H. Ngai, X. Hu, H. Zhao, and J. Wei, "Stochastic computation offloading and trajectory scheduling for UAV-assisted mobile edge computing," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 3688–3699, Apr. 2019. [Online]. Available: https://doi.org/10.1109/jiot.2018.2890133

[115] F. Costanzo, P. D. Lorenzo, and S. Barbarossa, "Dynamic resource optimization and altitude selection in uav-based multi-access edge computing," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, May 2020. [Online]. Available: https://doi.org/10.1109/icassp40776.2020.9053594

[116] X. Hou, Z. Ren, W. Cheng, C. Chen, and H. Zhang, "Fog based computation offloading for swarm of drones," in *ICC 2019 - 2019 IEEE International Conference on Communications (ICC)*. IEEE, May 2019. [Online]. Available: https://doi.org/10.1109/icc.2019.8761932

[117] Z. Yang, C. Pan, K. Wang, and M. Shikh-Bahaei, "Energy efficient resource allocation in UAV-enabled mobile edge computing networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4576–4589, Sep. 2019. [Online]. Available: https://doi.org/10.1109/twc.2019.2927313

[118] N. D. Pietro and E. C. Strinati, "An optimal low-complexity policy for cache-aided computation offloading," *IEEE Access*, vol. 7, pp. 182 499–182 514, 2019. [Online]. Available: https://doi.org/10.1109/access.2019.2959986

[119] E. C. Strinati, S. Barbarossa, J. L. Gonzalez-Jimenez, D. Ktenas, N. Cassiau, L. Maret, and C. Dehos, "6g: The next frontier: From holographic messaging to artificial intelligence using subterahertz and visible light communication,"

IEEE Vehicular Technology Magazine, vol. 14, no. 3, pp. 42–50, Sep. 2019. [Online]. Available: https://doi.org/10.1109/mvt.2019.2921162

[120] H. Ahmadi, K. Katzis, and M. Z. Shakir, "A novel airborne self-organising architecture for 5g+ networks," in *2017 IEEE 86th Vehicular Technology Conference (VTC-Fall)*. IEEE, Sep. 2017. [Online]. Available: https://doi.org/10.1109/vtcfall.2017.8288095

[121] J. Jackson, "The interplanetary internet [networked space communications," *IEEE Spectrum*, vol. 42, no. 8, pp. 30–35, Aug. 2005. [Online]. Available: https://doi.org/10.1109/mspec.2005.1491224

[122] B. B. A. via Integrated Terrestrial & Satellite Systems, "D2.4 overall integration architecture definition," 2013, available from: https://cordis.europa.eu/docs/projects/cnect/3/317533/080/deliverables/001-BATSD24FHv1F.pdf.

[123] V. Jungnickel, H. Gaebler, U. Krueger, K. Manolakis, and T. Haustein, "LTE trials in the return channel over satellite," in *2012 6th Advanced Satellite Multimedia Systems Conference (ASMS) and 12th Signal Processing for Space Communications Workshop (SPSC)*. IEEE, Sep. 2012. [Online]. Available: https://doi.org/10.1109/asms-spsc.2012.6333083

[124] J. Dommel, G. Boccolini, L. Raschkowski, S. Jaeckel, L. Thiele, T. Haustein, and N. G. Prelcic, "5g in space: PHY-layer design for satellite communications using non-orthogonal multi-carrier transmission," in *2014 7th Advanced Satellite Multimedia Systems Conference and the 13th Signal Processing for Space Communications Workshop (ASMS/SPSC)*. IEEE, Sep. 2014. [Online]. Available: https://doi.org/10.1109/asms-spsc.2014.6934543

[125] 3GPP, "Solutions for NR to support Non-Terrestrial Networks (NTN)," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 38.821, 01 2020, version 16.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3525

[126] Google, "LOON - Ballon Powered Internet," 2020, available from: https://loon.com/.

[127] A. Fotouhi, M. Ding, and M. Hassan, "Flying drone base stations for macro hotspots," *IEEE Access*, vol. 6, pp. 19 530–19 539, 2018. [Online]. Available: https://doi.org/10.1109/access.2018.2817799

[128] S. Barbarossa, E. Ceci, M. Merluzzi, and E. Calvanese-Strinati, "Enabling effective mobile edge computing using millimeterwave links," in *2017 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, May 2017. [Online]. Available: https://doi.org/10.1109/iccw.2017.7962685

[129] 3GPP, "Enhanced LTE support for aerial vehicles," 3rd Generation Partnership Project (3GPP), Technical Report (TR) 36.777, 01 2018, version 15.0.0. [Online]. Available: https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3231

[130] K. M. Addali, S. Y. B. Melhem, Y. Khamayseh, Z. Zhang, and M. Kadoch, "Dynamic mobility load balancing for 5g small-cell networks based on utility functions," *IEEE Access*, vol. 7, pp. 126 998–127 011, 2019. [Online]. Available: https://doi.org/10.1109/access.2019.2939936

[131] 5G-SOLUTIONS Consortium. (2019) H2020. [Online]. Available: https://www.5gsolutionsproject.eu/

[132] J. Machowski, J. Bialek, and J. Bumby, *Power System Dynamics: Stability and Control*. Wiley, 2011.

[133] F. Liberati, A. D. Giorgio, A. Giuseppi, A. Pietrabissa, E. Habib, and L. Martirano, "Joint model predictive control of electric and heating resources in a smart building," *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 7015–7027, Nov. 2019. [Online]. Available: https://doi.org/10.1109/tia.2019.2932954

[134] U. Tamrakar, D. Shrestha, M. Maharjan, B. Bhattarai, T. Hansen, and R. Tonkoski, "Virtual inertia: Current trends and future directions," *Applied Sciences*, vol. 7, no. 7, p. 654, Jun. 2017. [Online]. Available: https://doi.org/10.3390/app7070654

[135] C. Peng, J. Zou, and L. Lian, "Dispatching strategies of electric vehicles participating in frequency regulation on power grid: A review," *Renewable and Sustainable Energy Reviews*, vol. 68, pp. 147–152, Feb. 2017. [Online]. Available: https://doi.org/10.1016/j.rser.2016.09.133

[136] E. Hossain, R. Perez, S. Padmanaban, and P. Siano, "Investigation on the development of a sliding mode controller for constant power loads in microgrids," *Energies*, vol. 10, no. 8, p. 1086, Jul. 2017. [Online]. Available: https://doi.org/10.3390/en10081086

[137] AREA, "Documento per la consultazione dell'Autorità , 322/2019/R/eel, Testo Integrato del Dispacciamento elettrico(TIDE)- Orientamenti complessivi," 23.7.2019. [Online]. Available: https://www.arera.it/allegati/docs/19/322-19.pdf

[138] European Parliament, Council of the European Union , "Directive (EU) 2018/2001 of the European Parliament and of the Council of 11 December 2018 on the promotion of the use of energy from renewable sources," *OJ*, vol. L 328, p. 82–209, 11/12/2018. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=OJ:L:2018:328:TOC

[139] A. D. Giorgio, F. Liberati, and S. Canale, "Electric vehicles charging control in a smart grid: A model predictive control approach," *Control Engineering Practice*, vol. 22, pp. 147–162, Jan. 2014. [Online]. Available: https://doi.org/10.1016/j.conengprac.2013.10.005

[140] R. Germana, F. Liberati, and A. D. Giorgio, "Decentralized model predictive control of plug-in electric vehicles charging based on the alternating direction method of multipliers," in *2020 28th Mediterranean Conference on Control and Automation (MED)*. IEEE, Sep. 2020. [Online]. Available: https://doi.org/10.1109/med48518.2020.9183137

[141] B. De Wachter, Gerard Doorman, S. De Carlo, C. Neumann, F. Paul Sapp, K. Smolira, M. Ángel Martínez, C. Payement, S. Heather, and M. Foresti, "Options for the design of European Electricity Markets in 2030 About ENTSO-E," no. March, 2021. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/Publications/Market/Committee/publications/210331_Market_design2030.pdf

[142] Z. Wang and W. Wu, "Coordinated control method for DFIG-based wind farm to provide primary frequency regulation service," *IEEE Transactions on Power Systems*, vol. 33, no. 3, pp. 2644–2659, May 2018. [Online]. Available: https://doi.org/10.1109/tpwrs.2017.2755685

[143] A. Buckspan, J. Aho, P. Fleming, Y. Jeong, and L. Pao, "Combining droop curve concepts with control systems for wind turbine active power control," in *2012 IEEE Power Electronics and Machines in Wind Applications*. IEEE, Jul. 2012. [Online]. Available: https://doi.org/10.1109/pemwa.2012.6316403

[144] Z. Jietan, Q. Linan, R. Pestana, L. Fengkui, and Y. Libin, "Dynamic frequency support by photovoltaic generation with "synthetic" inertia and frequency droop control," in *2017 IEEE Conference on Energy Internet and*

Energy System Integration (EI2). IEEE, Nov. 2017. [Online]. Available: https://doi.org/10.1109/ei2.2017.8245445

[145] Y. Lin, P. Barooah, S. Meyn, and T. Middelkoop, "Experimental evaluation of frequency regulation from commercial building HVAC systems," *IEEE Transactions on Smart Grid*, vol. 6, no. 2, pp. 776–783, Mar. 2015. [Online]. Available: https://doi.org/10.1109/tsg.2014.2381596

[146] H. Hao, B. M. Sanandaji, K. Poolla, and T. L. Vincent, "Frequency regulation from flexible loads: Potential, economics, and implementation," in *2014 American Control Conference*. IEEE, Jun. 2014. [Online]. Available: https://doi.org/10.1109/acc.2014.6858734

[147] K. Ko and D. K. Sung, "The effect of cellular network-based communication delays in an EV aggregator's domain on frequency regulation service," *IEEE Transactions on Smart Grid*, vol. 10, no. 1, pp. 65–73, Jan. 2019. [Online]. Available: https://doi.org/10.1109/tsg.2017.2731846

[148] R. Germana, E. De Santis, F. Liberati, and A. Di Giorgio, "On the participation of charging point operators to the frequency regulation service using plug-in electric vehicles and 5g communications," in *2021 IEEE International Conference on Environment and Electrical Engineering and 2021 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)*. IEEE, Sep. 2021. [Online]. Available: https://doi.org/10.1109/eeeic/icpseurope51590.2021.9584495

[149] T. Taleb, K. Samdanis, B. Mada, H. Flinck, S. Dutta, and D. Sabella, "On multi-access edge computing: A survey of the emerging 5g network edge cloud architecture and orchestration," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 3, pp. 1657–1681, 2017. [Online]. Available: https://doi.org/10.1109/comst.2017.2705720

[150] E. Yao, V. W. S. Wong, and R. Schober, "A robust design of electric vehicle frequency regulation service," in *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, Nov. 2014. [Online]. Available: https://doi.org/10.1109/smartgridcomm.2014.7007729

[151] ——, "Optimization of aggregate capacity of PEVs for frequency regulation service in day-ahead market," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3519–3529, Jul. 2018. [Online]. Available: https://doi.org/10.1109/tsg.2016.2633873

[152] S. Xia, S. Q. Bu, X. Luo, K. W. Chan, and X. Lu, "An autonomous real-time charging strategy for plug-in electric vehicles to regulate frequency of distribution system with fluctuating wind generation," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 2, pp. 511–524, Apr. 2018. [Online]. Available: https://doi.org/10.1109/tste.2017.2746097

[153] F. S. Gorostiza and F. Gonzalez-Longatt, "Optimised TSO–DSO interaction in unbalanced networks through frequency-responsive EV clusters in virtual power plants," *IET Generation, Transmission & Distribution*, vol. 14, no. 21, pp. 4908–4917, Sep. 2020. [Online]. Available: https://doi.org/10.1049/iet-gtd.2019.1947

[154] Y. Kuang, C. Li, B. Zhou, Y. Cao, H. Yang, and L. Zeng, "Asynchronous method for frequency regulation by dispersed plug-in electric vehicles," *International Journal of Emerging Electric Power Systems*, vol. 19, no. 3, Apr. 2018. [Online]. Available: https://doi.org/10.1515/ijeeps-2017-0158

[155] H. Liu, J. Qi, J. Wang, P. Li, C. Li, and H. Wei, "EV dispatch control for supplementary frequency regulation considering the expectation of EV owners," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 3763–3772, Jul. 2018. [Online]. Available: https://doi.org/10.1109/tsg.2016.2641481

[156] H. Jia, X. Li, Y. Mu, C. Xu, Y. Jiang, X. Yu, J. Wu, and C. Dong, "Coordinated control for EV aggregators and power plants in frequency regulation considering time-varying delays," *Applied Energy*, vol. 210, pp. 1363–1376, Jan. 2018. [Online]. Available: https://doi.org/10.1016/j.apenergy.2017.05.174

[157] S. Hashemi, N. B. Arias, P. B. Andersen, B. Christensen, and C. Traholt, "Frequency regulation provision using cross-brand bidirectional v2g-enabled electric vehicles," in *2018 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*. IEEE, Aug. 2018. [Online]. Available: https://doi.org/10.1109/sege.2018.8499485

[158] S. Cai and R. Matsuhashi, "Model predictive control for EV aggregators participating in system frequency regulation market," *IEEE Access*, vol. 9, pp. 80 763–80 771, 2021. [Online]. Available: https://doi.org/10.1109/access.2021.3085345

[159] M. M. Islam, X. Zhong, Z. Sun, H. Xiong, and W. Hu, "Real-time frequency regulation using aggregated electric vehicles in smart grid," *Computers &*

Industrial Engineering, vol. 134, pp. 11–26, Aug. 2019. [Online]. Available: https://doi.org/10.1016/j.cie.2019.05.025

[160] D. A. Sbordone, E. M. Carlini, B. D. Pietra, and M. Devetsikiotis, "The future interaction between virtual aggregator-TSO-DSO to increase DG penetration," in *2015 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*. IEEE, Oct. 2015. [Online]. Available: https://doi.org/10.1109/icsgce.2015.7454296

[161] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," 2014. [Online]. Available: https://arxiv.org/abs/1411.1607

[162] Gurobi Optimization, LLC, "Gurobi Optimizer Reference Manual," 2021. [Online]. Available: https://www.gurobi.com