UNIVERSITY OF CALIFORNIA, SAN DIEGO

Multi-agent coordination algorithms for control of distributed energy resources in smart grids

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Engineering Sciences (Mechanical Engineering)

by

Andrés Cortés

Committee in charge:

Professor Sonia Martínez, Chair
Professor Robert Bitmead
Professor Phillip Gill
Professor Jan Kleissl
Professor Tajana Simunic-Rosing

2015

The dissertation of Andrés Cortés is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California, San Diego

2015

DEDICATION

To my family, in this world and the next.

EPIGRAPH

Never did any man yet repent of having spoken too little;
whereas many have been sorry that they spoke so much.
—One Thousand and One Nights

TABLE OF CONTENTS

Signature Page . . . iii

Dedication . . . iv

Epigraph . . . v

Table of Contents . . . vi

List of Figures . . . ix

List of Tables . . . xi

Acknowledgements . . . xii

Vita . . . xiv

Abstract of the Dissertation . . . xvi

Chapter 1 Introduction . . . 1
1.1 Notations . . . 3
1.2 Organization of the thesis . . . 4

Chapter 2 Hierarchical plug-in electric vehicle V1G charging control . . . 7
2.1 Problem formulation . . . 10
2.1.1 A closer look into usage schedules . . . 12
2.1.2 Optimization problem . . . 17
2.2 PRICE LEVELING algorithm for PEV charging . . . 25
2.3 Hierarchical PEV control under communication failures . . . 33
2.3.1 NON-ANONYMOUS PRICE LEVELING algorithm under communication failures . . . 33
2.3.2 ANONYMOUS PRICE LEVELING algorithm under communication failures . . . 35
2.4 Simulations . . . 36
2.5 Auxiliary section: invariance theory . . . 41
2.6 Details of the proof of Theorem 2.1 . . . 45
2.7 Summary . . . 49

Chapter 3 A hierarchical V2G protocol for PEVs . . . 51
3.1 Preliminary notation for this chapter . . . 52
3.2 Problem formulation . . . 53
3.2.1 Structure of the power network . . . 53
3.2.2 PEV battery model . . . 55

3.2.3 Load buses in the distribution feeders 55
3.2.4 The generation/pricing node 56
3.3 Optimal control problem 57
3.3.1 Analysis and design of the penalty method 63
3.4 A hierarchical control architecture 66
3.4.1 V2G HIERARCHICAL algorithm 67
3.5 Simulations and discussion 72
3.6 Auxiliary results for this chapter 74
3.7 Summary 75

Chapter 4 Hierarchical demand response with on/off Loads 79
4.1 Management of demand response events 81
4.1.1 Modeling a demand response event 82
4.1.2 Optimal DRE control problem formulation 84
4.2 Solution approach 85
4.2.1 Step 1: convex optimization 87
4.2.2 Step 2: threshold Computation 88
4.3 Decentralized solution 89
4.3.1 Step 1: Decentralized convex optimization 90
4.3.2 Step 2: Decentralized thresholding 92
4.4 Model predictive control implementation 93
4.5 Simulations 94
4.6 Summary 97

Chapter 5 Distributed control of user-side resources in microgrids 99
5.1 Notions of graph theory for this chapter 101
5.2 Problem statement 102
5.2.1 Mathematical model of a microgrid 103
5.2.2 Mathematical model of a battery 104
5.2.3 Communication network 105
5.2.4 The MICROGRID CONTROL PROBLEM 105
5.3 Distributed reactive power and storage control 108
5.3.1 Convergence analysis of the algorithm 112
5.3.2 Study of the rate of convergence 121
5.4 The MICROGRID OPTIMIZATION ALGORITHM 124
5.4.1 Handling the non-distributed terms 132
5.5 Voltage prediction 133
5.5.1 A multilayer control approach 135
5.5.2 Distributed approximation 135
5.6 Simulation results and discussion 138
5.7 Summary 141

Chapter 6 Self-triggered best response dynamics for continuous games . . . 145
6.1 Game theoretical notions . . . 147
6.2 Continuous-time best-response dynamics . . . 148
6.3 Self-triggered Communications in Best-Response Dynamics 153
6.4 System analysis via invariance theory . . . 157
6.5 Summary . . . 160

Chapter 7 Conclusions . . . 162

Bibliography . . . 164

LIST OF FIGURES

Figure 2.1: Time window representation, and Z_i, W_i for $i = 1, 2$. 14
Figure 2.2: Communication scheme for the Price Leveling algorithm. 26
Figure 2.3: Algorithm state and its relation with the PEV charging problem, for a fixed PEV $i \in I$ 27
Figure 2.4: Top figure: initial estimated demand profile. Bottom: optimal estimated demand profile. The non-PEV demand is shown in light grey. Other colors show each PEV's charging profile. 37
Figure 2.5: Evolution of the estimated cost function for the Non-Anonymous Price Leveling 39
Figure 2.6: Actual demand profile using the optimal charging strategy 39
Figure 2.7: Overall cost of the consumed energy (PEV and non-PEV) 40

Figure 3.1: Graph model of the power grid 54
Figure 3.2: Grid hierarchical structure 66
Figure 3.3: Topology for the simulation scenario 73
Figure 3.4: The optimal aggregate EV demand is shown in red. The aggregate EV demand from the V2G hierarchical algorithm is shown in black. 74
Figure 3.5: The non-PEV demand is shown in blue. The aggregate optimal demand is shown in red. The aggregate demand obtained through the V2G hierarchical algorithm is shown in black. 75
Figure 3.6: The non-PEV demand is shown in blue. The optimal aggregate demand is shown in red. The aggregate demand obtained through the V2G hierarchical algorithm is shown in black. The green line shows the upper bound for the power capacity. 76
Figure 3.7: The non-PEV demand is shown in blue. The optimal aggregate demand is shown in red. The aggregate demand obtained through the V2G hierarchical algorithm is shown in black. The green line determines upper bound for the power capacity. 77
Figure 3.8: Top figure: Evolution of $J(u^k, v^k) - J(u^\star, v^\star)$ vs number of iterations, for four scenarios with different No. of PEVs. Bottom figure: Optimal aggregate demand for each scenario. 78

Figure 4.1: Communication architecture for the decentralized implementation of the algorithm 89
Figure 4.2: $\log_{10}$ of maximum constraint violation vs iterations 95
Figure 4.3: In blue: discomfort (cost) for the solution to the convex relaxation of the problem $\mathcal{P}1$ (infeasible). In red: normalized discomfort for solution to $\mathcal{P}1$ provided by our proposed method (feasible). Simulation over 5 different scenarios. 96

Figure 4.4: In blue: energy dispatched according to the solution to the convex relaxation of the problem $\mathcal{P}1$, for each type of flexible load. In red: energy dispatched according to our proposed method for each type of load. Simulation over 5 different scenarios. 96
Figure 4.5: In green: normalized discomfort (cost) for the solution provided by our method. In blue: normalized discomfort for the MPC implementation of our method. 98

Figure 5.1: a) Evolution of v_4, b) evolution of q_4, as a function of the iteration number K, respectively, for each $t \in \tau$. Dashed lines show the optimal values. 121
Figure 5.2: Evolution of η_4, as a function of the iteration number K, respectively, for each $t \in \tau$. 122
Figure 5.3: Evolution of μ_4 as a function of the iteration number K, respectively, for each $t \in \tau$. 123
Figure 5.4: Control computation using voltage prediction. 138
Figure 5.5: Top: Evolution of $v_1(t)$ for all $t \in \tau$, using the MICROGRID OPTIMIZATION ALGORITHM. Bottom: Evolution of $v_1(t)$ for all $t \in \tau$ using the dual decomposition algorithm. 140
Figure 5.6: Top: Evolution of $q_1(t)$ for all $t \in \tau$, using the MICROGRID OPTIMIZATION ALGORITHM. Bottom: Evolution of $q_1(t)$ for all $t \in \tau$ using the dual decomposition algorithm. 141
Figure 5.7: Top plots: $\lambda_1(t)$, for all $t \in \tau$, for the MICROGRID OPTIMIZATION 142
Figure 5.8: Evolution of Lagrange multipliers for the dual decomposition algorithm 143
Figure 5.9: Evolution of Lagrange multipliers for the dual decomposition algorithm 144

LIST OF TABLES

Table 5.1: Parameters for each testbed 139
Table 5.2: Practical Convergence 139

ACKNOWLEDGEMENTS

I would like to thank professor Sonia Martínez for the countless hours, support and patience that she granted me. The trust that she has put on me, has allowed me to grow as a person and as a professional. Her ideas and mentorship helped me go through countless obstacles that I found during the development of my research. I have nothing but gratitude to her.

I also acknowledge my committee members, Professor Bob Bitmead, Professor Philip Gill, Professor Jan Kleissl, and Professor Tajana Simunic Rosing, for their time and feedback. I thank my advisor during my Masters, professor Nicanor Quijano, whose mentorship and support were fundamental for having me come to the University of California, San Diego.

I thank my lab friends Eduardo Ramírez, Daniele Cavaglieri, Lukas Nonnenmacher Federico Bribiesca, Pedro Franco, and Ashish Cherukuri, for the academic and non-academic discussions, barbecues, and coffee times, and for making school a more enjoyable place to be.

My parents and my brother, for their support, sacrifice and encouragement. To my extended family, those who are here, and the ones who passed during this time, for believing in me and making my life worth to be lived. There are no words to thank for everything I have received from them.

My girlfriend, who has been my family in the United States, and has supported me, pushed me and comforted during the harder times, and enjoyed with me during the better ones. She deserves my endless gratitude.

Coauthor acknowledgments: This thesis contains parts of the following publications with Professor Sonia Martínez as a coauthor:

- A. Cortés and S. Martínez, "A Projection-Based Dual Algorithm for Fast Computation of Control in Microgrids ", *Submitted to the SIAM Journal on Control and Optimization*, 2015.

- A. Cortés and S. Martínez, "A Hierarchical Demand-Response Algorithm for Vehicle-to-Grid Integration", *Submitted to the IEEE Transactions on Control Systems Technology*, 2015.

- A. Cortés and S. Martínez, "On Distributed Reactive Power and Storage Control on Microgrids", *Submitted to the International Journal of Robust and Nonlinear Control*, 2014, revised 2015.

- A. Cortés and S. Martínez, "A Decentralized Algorithm for Optimal Plug-in Electric Vehicle Charging with Usage Constraints", *Automatica*, conditionally accepted, 2015.

- A. Cortés and S. Martínez, "Self-triggered Best Response Dynamics for Continuous Games", *IEEE Transactions on Automatic Control*, vol 30 (4), 2015, pp 1115-1120.

- A. Cortés and S. Martínez,, "Hierarchical Management of Demand Response Events with On/Off Loads", *Submitted to the American Control Conference*, 2016.

- A. Cortés and S. Martínez, "A Projection-based Dual Algorithm for Fast Computation of Control in Microgrids", *Proceedings of the SIAM Conference on Control and its Applications*, 2015.

- A. Cortés and S. Martínez, "A Hierarchical Demand-Response Algorithm for Optimal Vehicle-to-Grid Coordination", *Proceedings of the European Control Conference*, 2015, pp 2425-2430.

- A. Cortés and S. Martínez, "Distributed Control of Reactive Power and Storage in Microgrids", *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems* , 2014, pp 563-570.

- A. Cortés and S. Martínez, "Optimal Plug-In Electric Vehicle Charging with Schedule Constraints", *Proceedings of Allerton*, 2013, pp 262-266.

- A. Cortés and S. Martínez, "Self-triggered Best Response Dynamics for Mobile Sensors Deployment", *Proceedings of the American Control Conference*, 2015, pp 6370-6375.

At the end of each chapter, we specify the publication from where the content was taken.

VITA

2008	B. S. in Electronics Engineering *Honors degree*, Universidad Pontificia Bolivariana, Colombia
2010	M. Sc. in Electronics Engineering *cum laude*, Universidad de los Andes, Colombia
2015	Ph. D. in Mechanical Engineering, University of California, San Diego

PUBLICATIONS

A. Cortés and S. Martínez, "A Projection-Based Dual Algorithm for Fast Computation of Control in Microgrids ", *Submitted to the SIAM Journal on Control and Optimization*, 2015.

A. Cortés and S. Martínez, "A Hierarchical Demand-Response Algorithm for Vehicle-to-Grid Integration", *Submitted to the IEEE Transactions on Control Systems Technology*, 2015.

A. Cortés and S. Martínez, "On Distributed Reactive Power and Storage Control on Microgrids", *Submitted to the International Journal of Robust and Nonlinear Control*, 2014, revised 2015.

A. Cortés and S. Martínez, "A Decentralized Algorithm for Optimal Plug-in Electric Vehicle Charging with Usage Constraints", *Automatica*, conditionally accepted, 2015.

A. Cortés and S. Martínez, "Self-triggered Best Response Dynamics for Continuous Games", *IEEE Transactions on Automatic Control*, vol 30 (4), 2015, pp 1115-1120.

A. Cortés and S. Martínez,, "Hierarchical Management of Demand Response Events with On/Off Loads", *Submitted to the American Control Conference*, 2016.

A. Cortés and S. Martínez, "A Projection-based Dual Algorithm for Fast Computation of Control in Microgrids", *Proceedings of the SIAM Conference on Control and its Applications*, 2015.

A. Cortés and S. Martínez, "A Hierarchical Demand-Response Algorithm for Optimal Vehicle-to-Grid Coordination", *Proceedings of the European Control Conference*, 2015, pp 2425-2430.

A. Cortés and S. Martínez, "Distributed Control of Reactive Power and Storage in Microgrids", *Proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems* , 2014, pp 563-570.

A. Cortés and S. Martínez, “Optimal Plug-In Electric Vehicle Charging with Schedule Constraints”, *Proceedings of Allerton*, 2013, pp 262-266.

A. Cortés and S. Martínez, “Self-triggered Best Response Dynamics for Mobile Sensors Deployment”, *Proceedings of the American Control Conference*, 2015, pp 6370-6375.

ABSTRACT OF THE DISSERTATION

Multi-agent coordination algorithms for control of distributed energy resources in smart grids

by

Andrés Cortés

Doctor of Philosophy in Engineering Sciences (Mechanical Engineering)

University of California, San Diego, 2015

Professor Sonia Martínez, Chair

Sustainable energy is a top-priority for researchers these days, since electricity and transportation are pillars of modern society. Integration of clean energy technologies such as wind, solar, and plug-in electric vehicles (PEVs), is a major engineering challenge in operation and management of power systems. This is due to the uncertain nature of renewable energy technologies and the large amount of extra load that PEVs would add to the power grid. Given the networked structure of a power system, multi-agent control and optimization strategies are natural approaches to address the various problems of interest for the safe and reliable operation of the power grid. The distributed computation in multi-agent algorithms addresses three problems at the same time: i) it allows for the handling of problems with millions of variables that a single processor

cannot compute, ii) it allows certain independence and privacy to electricity customers by not requiring any usage information, and iii) it is robust to localized failures in the communication network, being able to solve problems by simply neglecting the failing section of the system.

We propose various algorithms to coordinate storage, generation, and demand resources in a power grid using multi-agent computation and decentralized decision making. First, we introduce a hierarchical vehicle-one-grid (V1G) algorithm for coordination of PEVs under usage constraints, where energy only flows from the grid in to the batteries of PEVs. We then present a hierarchical vehicle-to-grid (V2G) algorithm for PEV coordination that takes into consideration line capacity constraints in the distribution grid, and where energy flows both ways, from the grid in to the batteries, and from the batteries to the grid. Next, we develop a greedy-like hierarchical algorithm for management of demand response events with on/off loads. Finally, we introduce distributed algorithms for the optimal control of distributed energy resources, i.e., generation and storage in a microgrid. The algorithms we present are provably correct and tested in simulation. Each algorithm is assumed to work on a particular network topology, and simulation studies are carried out in order to demonstrate their convergence properties to a desired solution.

Chapter 1

Introduction

As awareness toward climate change and greenhouse emissions increases, clean alternatives for energy generation and transportation are being actively investigated. The use of technologies that are considered clean, such as solar and wind generation, and electric vehicles, is growing at a significant rate, propelled by governmental subsidies and social environmental awareness.

However, the introduction of such technologies poses additional engineering challenges for the proper operation of power systems. The inherent uncertainty of renewable energy, and the immense additional electric load that electric vehicles represent may lead power systems to instability, if appropriate engineering measures are not taken.

Control and optimization strategies have been researched in the last years, aiming for the harmless introduction of such new elements into power systems. Moreover, new ideas have come up in order to operate the power grid with the oncoming elements. Demand response and electric vehicle coordination for supply/demand balancing, and distributed generation allow for a more efficient and robust power grid. All these ideas also present control and decision making problems. These problems involve the coordination of millions of elements, trying to fulfill grid operation objectives while guarantee quality of service for users.

Multi-agent algorithms have been extensively considered for control and optimization of power grids in the last few years. Probably the main advantage of using multi-agent algorithms is that the computational burden can be parallelized. This is convenient since the large-scale of a power system, may lead to extremely large decision

making problems that cannot be solved in a centralized way. In addition, many multi-agent algorithms have been shown to be robust to information losses, communication failures, or changes in the structure of the computation network.

There are two major multi-agent structures that are considered for control and optimization of power systems. First, a distributed architecture which is horizontal, i.e., all agents perform the same task, information only flows between neighboring agents according to a communication graph, and no operation is fully carried out in a centralized manner. Second, a hierarchical architecture, which is widely considered in demand response studies. In this architecture, there are agents, usually associated to the utility or the system operator, which aggregate information in order to provide a coordination signal to agents that control user-side elements, e.g., loads, electric vehicles, etc. Then, those agents in the user-side are the ones that decide the action to take locally. There is never information exchange between users, but only between users and the grid.

Among the multi-agent algorithms that have been used for control of resources in power grids, its worth mentioning distributed load balancing for demand response [1, 2], consensus for economic dispatch [3], Laplacian-gradient (continuous-time) dynamics, also for economic dispatch [4, 5], dual decomposition [6, 7] and alternate direction method of multipliers (ADMM) [8, 9] for the optimization of distributed energy resources (DERs). While the dual decomposition algorithm allows for distributed computation in many cases (separability), it also exhibits very slow convergence. The ADMM algorithm presents higher rate of convergence, but the elements enabling higher speed make usually not possible to implement distributed computation designs. There are several works oriented to demand response and electric vehicle coordination, that are based on a hierarchical information exchange [10, 11, 12, 13, 14].

In chapters 2, 3, and 4 we explore the so called hierarchical architecture for the proposed algorithms. In this configuration, all customers, e.g., PEVs, buildings, etc., possess computation capacity and communicate directly with agents in the power grid such as the utility, the Independent System Operator (ISO), distribution buses with computational power. Then, customers receive from them a coordination signal based on the system objectives. The larger amount of the computation is carried out in the customer side, while the grid side only computes and transmits a single coordination

signal.

In chapter 5, we present a more horizontal architecture, in which resources in a microgrid, i.e., generators and storage systems, communicate with each other in order to optimize the performance of the microgrid. No agent within this architecture requires to perform centralized aggregation or computation for the execution of the proposed algorithms.

Chapter 6 explores the continuous-time best response dynamics. This can be used on a multi-agent coordination setting in which the problem is formulated as a continuous game. Our main contribution in this chapter is turning the continuous-time best response dynamics into a self-triggering algorithm that does not need permanent communication between agents, but agents estimate when to communicate based on their own needs. The challenging part of the self-triggering algorithm design is choosing a communication triggering law that leads to a provably convergent algorithm to a desired solution of the game.

1.1 Notations

In what follows, $\mathbb{R}$ will denote the set of real numbers, $\mathbb{R}_{>0}$ the set of real-positive numbers, $\mathbb{N}$ the set of natural numbers, and $\mathbb{C}$ the set of complex numbers. For $x \in \mathbb{R}$, $|x|$ denotes the absolute value of x. The complex exponential function is denoted as e^{jx}, for $x \in \mathbb{R}$. notation $\|X\|$ denotes the Euclidean norm of the element $X \in \mathbb{R}^{m\times n}$, while $\|X\|_{\mathbb{C}} \in \mathbb{R}^{m\times n}_{\geq 0}$ is understood as an array whose entries are the magnitude of each entry of the variable $X \in \mathbb{C}^{m\times n}$. The operator $\angle X$ denotes the phase angle of all the complex entries of X. We let $\mathbf{e}_l \in \mathbb{R}^n$ be a vector whose $l+1$ entry is equal to one, for $l \in \{0, \ldots, n-1\}$, while all other entries are zero, and $\mathbf{1} \in \mathbb{R}^n$ be a vector which entries are all equal to one. Dimension of $\mathbf{e}_l$ and $\mathbf{1}$ depends on the context. The notation $\mathbf{0}_{m\times n} \in \mathbb{R}^{m\times n}$ represents a matrix whose entries are zero.

For $\mathcal{A}, \mathcal{B} \in X$, where X is some set, $\mathcal{A} \setminus \mathcal{B} \triangleq \{a \in \mathcal{A} \mid a \notin \mathcal{B}\}$. Given the finite set $\mathcal{A}$, $|\mathcal{A}|$ denotes its cardinality. For $n \in \mathbb{N}$, I_n denotes the identity matrix in $\mathbb{R}^{n\times n}$. Consider a set $\tau = \{1, \ldots, T\}$, and let $x \in \mathbb{C}^{|\mathcal{A}|T}$, with entries $x_l(t) \in \mathbb{C}$, for all $l \in \mathcal{A}$, $t \in \tau$; that is, $x = [x_1(1) \ldots, x_{|\mathcal{A}|}(1), \ldots, x_1(T), \ldots, x_{|\mathcal{A}|}(T)]^\top$. Then, $x(t) \in \mathbb{C}^{|\mathcal{A}|}$ denotes

the vector $x(t) = [x_1(t), \ldots, x_{|\mathcal{A}|}(t)]^\top$, for each $t \in \tau$. Similarly, $x_l \in \mathbb{C}^\top$ denotes the vector $x_l = [x_l(1), \ldots, x_l(T)]^\top$, for each $l \in \mathcal{A}$.

For a complex square matrix A, we denote its spectrum as $\text{spec}(A)$ and its spectral radius as $\rho(A)$. For the matrix A, A_{ij} represents its (i, j) entry. If A is block partitioned, $(A)_{ij}$ represents its (i, j) block. For $x \in \mathbb{C}^n$, $\mathbf{d}(x)$ denotes the diagonal matrix such that $\mathbf{d}(x)_{ii} = x_i$, for $i \in \{1, \ldots, n\}$, while for a complex square matrix A, $\mathbf{d}(A)$ is a diagonal matrix such that $\mathbf{d}(A)_{ii} = A_{ii}$. Let C be a complex array in $\mathbb{C}^{m \times n}$, we denote the null space of C as $\text{null}(C) \triangleq \{x \in \mathbb{C}^n \mid Cx = 0\}$. The row space of C, defined as the set of all linear combinations of the rows of C, is denoted as $\text{row}(C)$. Finally, $\hat{C}$ denotes the conjugate transpose of C.

For $r > 0$, define $B_r(z) = \{y \in \mathbb{R}^q \mid \|z - y\| \leq r\}$. Let S be a set in $\mathbb{R}^q$. Then, $S + B_r(0) = \{y \in \mathbb{R}^q \mid \exists z \in S \text{ s.t. } \|z - y\| \leq r\}$. Let $\mathcal{B} \subseteq \mathbb{R}^n$ be a convex set. Then, for $x \in \mathbb{R}^n$, $\text{dist}(x, \mathcal{B}) \triangleq \inf_{y \in \mathcal{B}} \|x - y\|$. Given functions $f : \mathbb{R}^m \to \mathbb{R}$ and $g : \mathbb{R}^m \to \mathbb{R}^n$, we denote the "small-o" notation $g(x) = o(f(x))$ if $\lim_{\|x\| \to \infty} g(x) f(x)^{-1} = 0$. For a vector $x \in \mathbb{R}^n$, $x \geq 0$ indicates that all entries of x are nonnegative. For a function $V : \mathbb{R}^q \to \mathbb{R}^s$, and the set-valued map $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^q$, define $V \circ F : \mathbb{R}^m \rightrightarrows \mathbb{R}^s$, such that $V \circ F(z) = \{y \in \mathbb{R}^s \mid \exists \xi \in F(z) \text{ s.t. } W(\xi) = y\}$. For a function $f : \mathbb{R} \to \mathbb{R}$, let us denote the derivative of f by f'.

Note: Some notations that are particular to each chapter, will be introduced at the beginning of the chapter itself.

1.2 Organization of the thesis

The contents of this thesis can be summarized as follows: Chapter 2 presents a hierarchical algorithm for the charging of Plug-in Electric Vehicles (PEVs) with usage constraints. The algorithm aims to optimize the price of energy provided to all PEV and non-PEV loads over a time horizon. The algorithm works on a one-to-all topology, where an aggregator provides a coordination signal that is the energy price at each time of the horizon, and PEVs update their charging strategies based on the coordination signal.

Chapter 3 introduces a hierarchical algorithm for charging of PEVs which in-

cludes line capacity constraints in the distribution grid, and considers V2G interaction, i.e., PEVs can also inject power into the grid. To this end, the power grid is modeled as a rooted tree, where the root node provides a coordination signal based on the energy price which depends on the total demand. Then, such signal is modified by nodes in the distribution-side of the grid according to whether power flows reaching them are satisfying line capacity constraints. Finally, all PEVs use those coordination signals to decide on their charging/discharging strategy for the upcoming time horizon.

Chapter 4 introduces a greedy-like algorithm for the management of demand response events with on/off loads. The algorithm also follows a one-to-all topology, in which two steps are executed sequentially: i) convex optimization and ii) thresholding to account for the on/off loads. Both steps can be carried out in the aforementioned topology using algorithms that are guaranteed to converge to the centralized solution of the problem.

Chapter 5 presents a formulation for the distributed control of storage and reactive power in microgrids, with two algorithms to address the formulated problem. The first algorithm is a dual decomposition approach that solves an approximation of the AC optimal power flow problem over a time horizon. Given the slow rate of convergence that the algorithm exhibits in simulation, we introduce an algorithm that accounts for local constraints using projections onto the feasible sets. This algorithm is shown to outperform the rate of convergence of the dual decomposition algorithm by two orders of magnitude. Both algorithms work over a communication graph that depends on the graph that represents the microgrid.

Chapter 6 introduces the self-triggered best response dynamics, which is the self-triggered implementation of the well known best response dynamics. The best response dynamics is represented by a differential inclusion in which players of a continuous game move their strategies towards the best response set for the current action of other players. The self-triggered version of such dynamics restricts the knowledge players have about other players' actions to a discrete sequence of time instants. The idea is finding a sequence of times at which each player must update information on other players' actions, in such a way that the dynamics still converges to the Nash equilibrium of the game. We characterize sufficient conditions on the game structure for the self-

triggered best response dynamics to reach the Nash equilibrium of the game.

Chapter 2

Hierarchical plug-in electric vehicle V1G charging control

Plug-in Electric Vehicles are being proposed as an important element in flexible load control that can both help alleviate environmental transportation costs and our dependency on petroleum energy sources.

However, a large penetration of PEVs may also negatively affect the operation of the power system, by creating new demand peaks and system overload. These phenomena incur into additional stress on generation, transmission and distribution systems, which translates into increased costs for users and electric generation companies. In order to lower the burden PEVs create on power systems, and at the same time decrease end-user costs, new algorithmic approaches on PEV charging are being designed with the goal of achieving peak-shaving solutions. This chapter contributes in this regard by proposing a novel algorithm that allocates PEV load at low-demand hours, while accounting for planned PEV scheduling constraints.

Diverse control architectures have been proposed to minimize power demand and to avoid the rise of new load peaks: centralized, distributed, and hierarchical. In a fully distributed setting, the network is solely comprised of PEVs, which exchange information with a subset of neighboring PEVs and make decisions based on that information. In a hierarchical architecture (also referred to as "decentralized" in the literature), agents engage in a similar process, but employing a special tree communication structure and minimal communication interaction. According to this distinction, we find

the following related works in the literature.

The paper [15] formulates an optimization problem which is solved in a centralized manner to come up with a valley-filling solution. In [16], a centralized PEV charging coordination strategy is proposed in order to shave demand peaks as well as minimize distribution losses. A supervisor controls the battery charging policies for all the PEVs in [17], with the aim of minimizing costs and regulating voltage. In [18], a distributed online approach is followed, in order to decide charging rates for each time. To this end, each electric vehicle uses measures of the instantaneous congestion of those nodes of the grid, by which power flows towards it. The authors of [19] introduce a pricing-based two-layer control algorithm for charging/discharging of PEVs. The algorithm is distributed, exploiting consensus-algorithm ideas. The characterization of the solutions and performance analysis are made via game theory and nonlinear analysis. Neither of the above works considers constraints based on usage schedule.

In [20], optimal charging trajectories are computed using linear programming. The authors propose two hierarchical algorithms to solve the problem. The first requires information about a centralized solution, particularly about the cost function gradient, while the second one assumes that each PEV computes a valley-filling solution based on the average charge requirements from all PEVs. No guarantee of optimality is provided. The work [10] introduces an algorithm that computes optimal charging strategies for a large population of PEVs. A bargain is performed between an energy coordinator and the PEVs, which leads to a valley-filling solution that minimizes the overall energy price. In this work, all PEVs are considered to have the same charging schedule. The paper [11], generalizes the setting of the previous work. The bargaining idea is similar to the one in [10], but it is assumed that PEVs have constraints on the maximal amount of energy that can be charged into their batteries at each time, as well as deadlines for complete charge. The result in [11] is also extended to an asynchronous iteration, under mild connectivity assumptions and for non-anonymous interactions between the utility and PEVs. The works [11, 10] present algorithms that are based on the solution of local convex optimization problems in a repeated way. Although convex optimization problems can be efficiently solved, each iteration involves several computationally expensive steps (e.g., solution of linear equations systems) which must be carried out sequentially.

These algorithms have been proven to exhibit asymptotic convergence to the optimal solution of the problem. More recently, the work [21] presents various algorithms for hierarchical V1G over a grid with line capacity constraints. The proposed algorithms require the execution of inner loops, which significantly increase computational and communication cost.

The contributions of this chapter are twofold. We present a novel hierarchical approach, the PRICE LEVELING algorithm, based on local interaction rules that meet usage schedule constraints. In this way, our algorithm is represented by a nonlinear difference equation, which only involves sums, and products. This improves on the required computational effort as compared to [11, 10], in which at each iteration a convex optimization problem must be solved by each PEV. This presents two main advantages. First, algorithms with lower computational requirements reduce errors in online implementations. Secondly, they allow for the use of cheaper computational devices in offline implementations, which is of concern to both grid operators and users.

The usage constraints we consider are described by energy requirements that must be achieved by each PEV before certain times of the day, in order to meet user needs. In addition, this algorithm also respects the bounds on the charging rate for each PEV battery. We further present a NON-ANONYMOUS PRICE LEVELING algorithm, a version of our algorithm in a non-anonymous interaction setting under communication failures. In order to analyze our algorithms, we present an invariance result for discrete-time set-valued systems, which is more general than the LaSalle invariance principle for difference inclusions. This result is instrumental for our proof of convergence to an optimal generalized valley-filling solution. Simulations demonstrate the validity of the theoretical results, and illustrate how the algorithm would perform under anonymous time-varying interactions.

This chapter is organized as follows: in Section 2.1, we formulate the PEV charging problem under scheduling constraints, as an optimization problem, and we also present some results to characterize the optimal solution of this problem. In Section 2.2, we introduce the PRICE LEVELING algorithm, and present some characterization of its behavior, as well as the convergence analysis towards the set of optimal solutions of the PEV charging problem. In Section 2.3, we present the NON-ANONYMOUS PRICE

LEVELING algorithm to work in a scenario with communication failures. In Section 2.4, we show simulation results for a specific scenario with communication failures.

2.1 Problem formulation

In this section, we describe the PEV charging problem of our interest. In simple words, a large group of PEVs must determine how to charge their batteries during a certain period of time, e.g., a day by interacting with an energy provider, in a way that alleviates the additional burden from the power grid. This energy must suffice to satisfy the user's needs at different times of the day, while keeping some amount of energy (possibly zero) in the battery at the end of the day. To this end, we first introduce a battery model and a usage schedule model. Then, we present PEV charging as a convex optimization problem. Further, we present some results that characterize any solution of the optimization problem.

Let us consider a set $I = \{1, \dots, N\}$ of PEVs, where each PEV $i \in I$ is supplied by a battery with maximum capacity $\beta_i > 0$. Based on the usage information, which is assumed to be known either from a forecasting process, or provided by the user, each PEV must plan the charging strategy for the upcoming day. The energy is obtained from an energy provider or utility through the power grid. For the remaining of this manuscript we assume that the power grid is able to provide enough power to supply any population of PEVs, along with the demand that does not come from PEVs. In order to characterize the PEV usage, the day is divided into $T \in \mathbb{N}$ time slots with equal duration. Define the set $\tau = \{1, \dots, T\}$ as the set of time slots in a day. Each slot $t \in \tau$ has a non-PEV demand D_t associated to it. This non-PEV demand is assumed to be known by the utility via demand forecast. Let $W_i \subseteq \tau$ for each $i \in I$, be the set of time slots at which the vehicle will be in use, this is, the time slots at which the vehicle is not connected to a power supply. The estimated energy usage of each PEV at each $t \in W_i$ will be characterized as a percentage of β_i. Formally, for each $t \in W_i$ and each $i \in I$, we define $w_{i,t} \in [0, 1]$ as the amount of energy (in percentage of the battery capacity) the vehicle aims to spend during the t^{th} time slot. Let us denote the set $\tau \setminus W_i$ as Z_i for all $i \in I$. This is the set of all time slots at which a PEV is connected to any power source

that is fed by the power grid.

Remark 2.1. *Note that during the time horizon, a PEV can be taking energy from different locations at different times, e.g., at the user's home, or at a parking lot. All such times are included in the set Z_i. The set Z_i is independent of the charging point location, and is solely defined by whether or not there is an available power source for the PEV at each time.*

Following, we define $w_{i,t}$ for $t \in Z_i$ as $w_{i,t} = 0$, for all $i \in I$, i.e., there is no energy usage during the charging time. Then, a linear model for the battery state at the end of each time slot is given by:

$$\vartheta_{i,t} = \vartheta_{i,t-1} + \frac{\alpha_i}{\beta_i} u_{i,t} - w_{i,t}, \quad i \in I,$$

for all $t \in \tau$, where $u_{i,t}$ is the amount of energy collected by the i^{th} PEV during the t^{th} time slot, and $\vartheta_{i,t}$ is the state-of-charge (SOC) of the i^{th} PEV battery at the end of the t^{th} time slot. The parameter $\alpha_i \in (0, 1)$ models the efficiency of the battery/charger system, which presents losses due to the internal resistance of the battery, parasitic currents in the charger, etc. This model is a simple representation of energy balances at each time slot, and has been widely used in the literature (see [12, 11, 10]). The explicit solution for $\vartheta_{i,t}$, $t \in \tau$, $i \in I$ is given by:

$$\vartheta_{i,t} = \vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{q=1}^{t} u_{i,q} - \sum_{q=1}^{t} w_{i,q}. \tag{2.1}$$

Since the i^{th} PEV has no available power supply at time slots $t \in W_i$, we define $u_{i,t} = 0$ for all $t \in W_i$. Certainly, there exist constraints on the amount of energy the i^{th} PEV battery is able to get during a time slot $t \in \tau$, which cannot be larger than some value $u_{i,\max} > 0$, given by the battery/charger specifications, for $i \in I$. Based on this battery model, which includes the aforementioned physical constraints, we define a charging strategy for the i^{th} PEV as a vector $u_i \in \mathbb{R}^T_{\geq 0}$, where $u_i = [u_{i,1}, \ldots, u_{i,T}]^\top$, $u_{i,t} = 0$ for all $t \in W_i$, with $u_{i,t} \in [0, u_{i,\max}]$ for $t \in \tau$, such that $w_{i,q} \leq \vartheta_{i,q-1}$ is guaranteed for all $q \in W_i$. This inequality means that the energy stored in the PEV battery at the beginning of the q^{th} time slot must be enough to provide the usage requirement $w_{i,q}$. The conditions

$w_{i,q} \leq \vartheta_{i,q-1}$ for each $q \in W_i$, hereinafter called *usage satisfaction constraints* can be rewritten from (2.1) as:

$$\frac{\beta_i}{\alpha_i}\left(\sum_{q=1}^{t} w_{i,q} - \vartheta_{i,0}\right) \leq \sum_{q=1}^{t-1} u_{i,q}, \tag{2.2}$$

for all $t \in \tau$, and $i \in I$. In order to include the need for a minimum amount of energy at the end of the day, we can define a virtual slot $T+1$, and $w_{i,T+1}$ as the percentage of the battery charge that must remain in the battery by the end of the T^{th} time slot. Thus, we say that (2.2) must hold for all $t \in \tau \cup \{T+1\}$. Notice that this condition also guarantees that $\vartheta_{i,t} \geq 0$ for all $t \in \tau \cup \{T+1\}$, and for all $i \in I$. The last feasibility conditions, from now on called *state-of-charge capacity constraints (SOC constraints)* regarding the amount of energy in the battery, are written as:

$$\vartheta_{i,t} \triangleq \vartheta_{i,0} + \frac{\alpha_i}{\beta_i}\sum_{q=1}^{t} u_{i,q} - \sum_{q=1}^{t} w_{i,q} \leq 1, \tag{2.3}$$

for all $t \in \tau$, $i \in I$, given by the fact that $\vartheta_{i,t} \in [0,1]$ for all $i \in I$, $t \in \tau$.

Remark 2.2. *In general it is not desired that a battery reaches states of charge of* 0 *or* 1. *Reaching such values can seriously damage the battery. Our framework allows to easily replace the operation interval from* $[0,1]$ *to* $[\vartheta_{i,\min}, \vartheta_{i,\max}]$, *for* $0 < \vartheta_{i,\min} < \vartheta_{i,\max} < 1$. *However, for simplicity in the notation we keep considering an operation interval of* $[0,1]$.

For convenience, let us introduce the notation u for the charging profile of the entire set of PEVs, namely, $u \triangleq \{u_i\}_{i \in I}$.

2.1.1 A closer look into usage schedules

In order to better understand the way in which we model usage schedules, let us take a closer look into the set W_i, for each $i \in I$.

The first important thing to realize is that for some $i \in I$ it could be that $1 \in W_i$. It means that the i^{th} PEV may have a planned continuous usage for time slots $\{1, \ldots, t_0\} \subset W_i$, for some $t_0 \geq 1$, such that $t_0+1 \in Z_i$. Since there is no time before that usage at which

the battery can be charged, the feasibility of the problem will be subject to whether the initial battery state $\vartheta_{i,0}$ is enough to satisfy the usage during time slots $\{1,\dots,t_0\}$. In order to allow for feasibility of the problem, we constrain our interest to the case in which if $1 \in W_i$ for some $i \in I$, then $\sum_{q=1}^{t_0} w_{i,q} \leq \vartheta_{i,0}$. Next, for each $i \in I$, let us define the set W_i^0 as:

$$W_i^0 = \begin{cases} \emptyset, & \text{if } 1 \in Z_i \\ \{1,\dots,t_0\}, & \text{if } 1 \in W_i. \end{cases}$$

Since we take for granted that the energy usage for W_i^0 will be satisfied, the solution to the optimization problem described below focuses on charging the battery to fulfill the energy requirement for the time determined by the set $(W_i \cup \{T+1\}) \setminus W_i^0$, for each $i \in I$.

Next, we introduce partitions on the sets $(W_i \cup \{T+1\}) \setminus W_i^0$ and Z_i, for all $i \in I$. Each element of the partition of $(W_i \cup \{T+1\}) \setminus W_i^0$ represents an uninterrupted time period of usage for the i^{th} PEV. Likewise, each element in the partition of Z_i represents an uninterrupted time period during which the battery can be recharged. Thus, the W and the Z are complementary in the timeline. This new notation allows us to reduce the amount of usage constraints in the model, and is also used to introduce the Price Leveling algorithm in Section 2.2.

First, let us partition the set $(W_i \cup \{T+1\}) \setminus W_i^0$ in connected components as $(W_i \cup \{T+1\}) \setminus W_i^0 = \bigcup_{\ell=1}^{m_i} W_i^\ell$, for each $i \in I$. That is, W_i^n is such that if $q, t \in W_i^n$, $q \leq t$, then $\{q,\dots,t\} \subseteq W_i^n$. In this way, the collection of sets $\{W_i^\ell\}_{\ell=0}^{m_i}$ is a partition of the set W_i. Notice that the elements of the collection are indexed following the order of the time slots in τ, i.e., $n < \ell$ if and only if $q < t$ for all $q \in W_i^n$ and $t \in W_i^\ell$. Further, let us define a partition $\{Z_i^\ell\}_{\ell=1}^{m_i}$ of Z_i by its connected components, similarly to the partition of $W_i \cup \{T+1\}$. That is, each set in the partition is clustering consecutive time slots; see in Figure 2.1 the sets associated with two PEVs, such that $m_1 = 3$ and $m_2 = 2$.

In order to ease the problem description, we introduce further notation as follows. We let $\mathcal{Z}_i^j = \bigcup_{r=1}^{j} Z_i^r$, for $j \in \{1,\dots,m_i\}$, and $\mathcal{W}_i^j = \bigcup_{r=0}^{j} W_i^r$, for $j \in \{0,\dots,m_i\}$. The number $\mathbf{n}(i,t)$, for each $t \in Z_i$, $i \in I$, represents the set of the partition $\{Z_i^\ell\}_{\ell=1}^{m_i}$ that contains t. As an example, for $t = 11$ in Figure 2.1, we have that $\mathbf{n}(2,11) = 1$. In terms

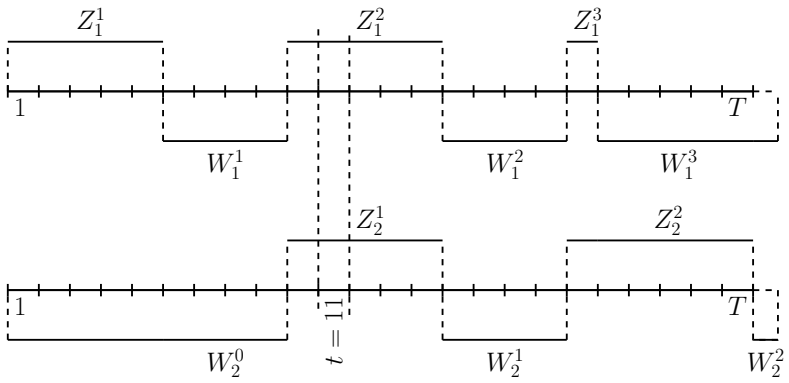

Figure 2.1: Time window representation, and Z_i, W_i for $i = 1, 2$.

of the new notation, we can recast our constraints using the partitions $\{W_i^\ell\}_{\ell=1}^{m_i}$, $\{Z_i^\ell\}_{\ell=1}^{m_i}$, for all $i \in I$. Since $u_{i,t} = 0$ for all $t \in W_i$, we have that if u satisfies the condition $\vartheta_{i,t-1} \geq w_{i,t}$ for the last $t \in W_i^r$, $r \in \{1, \ldots, n_i\}$, then, the condition is fulfilled for all $s \in W_i^r$. Then, the *user satisfaction constraints* can be recast as:

$$\frac{\beta_i}{\alpha_i}\left(\sum_{q\in\mathcal{W}_i^n} w_{i,q} - \vartheta_{i,0}\right) \leq \sum_{q\in\mathcal{Z}_i^n} u_{i,q}, \tag{2.4}$$

for $i \in I$, $n \in \{1, \ldots, m_i\}$. Notice that the summations above are performed over the sets that are denoted by calligraphic letters, i.e., $\mathcal{W}_i^n$ and $\mathcal{Z}_i^n$. Clearly, if (2.4) holds for $n = m_i$, $w_{i,T+1} \leq \vartheta_{i,T}$, the requirement of minimum energy at the end of the day is fulfilled. The *SOC constraints*, can be given in terms of the partitions $\{W_i^\ell\}_{\ell=1}^{n_i}$, $\{Z_i^\ell\}_{\ell=1}^{m_i}$ as follows:

$$\vartheta_{i,t_n} \triangleq \vartheta_{i,0} + \frac{\alpha_i}{\beta_i}\sum_{q\in\mathcal{Z}_i^n} u_{i,q} - \sum_{q\in\mathcal{W}_i^{n-1}} w_{i,q} \leq 1, \tag{2.5}$$

for all $i \in I$, $n \in \{1, \ldots, m_i\}$, where $t_n \triangleq \max \mathcal{Z}_i^n = \max Z_i^n$. Likewise, the summations in (2.5) are carried out over the sets that are denoted by calligraphic letters. As $w_{i,q} = 0$ for all $q \in Z_i^n$, we have that if (2.5) holds, the battery state is less or equal than one for all $q \in Z_i^n$. Furthermore, it follows that since $u_{i,q} = 0$ for all $q \in W_i$, then $\vartheta_{i,q} \leq 1$ for all $q \in W_i^n$.

We establish next some verifiable conditions on the parameters of the PEV

charging problem so that it is feasible.

Lemma 2.1 (Verification of feasibility)**.** *Define the recursion:*

$$A_i^n = B_i^{n-1} + \frac{\alpha_i}{\beta_i}|Z_i^n| \min\left\{u_{i,\max}, (1 - B_i^{n-1})\frac{\beta_i}{\alpha_i|Z_i^n|}\right\}, \tag{2.6}$$
$$B_i^{n-1} = A_i^{n-1} - \sum_{t \in W_i^{n-1}} w_{i,t},$$

for all $n \in \{1, \ldots, m_i\}$, $i \in I$. Define $A_i^0 = \vartheta_{i,0}$, for all $i \in I$. Then, the PEV charging problem is feasible if and only if for all $n \in \{0, \ldots, m_i\}$ and for all $i \in I$, it holds that $A_i^n \geq \sum_{t \in W_i^n} w_{i,t}$, and $A_i^n = \vartheta_{i,t_n}$, where $t_n = \max Z_i^n$, given the charging strategy:

$$u_{i,t} = \min\left\{u_{i,\max}, (1 - B_i^{n-1})\frac{\beta_i}{\alpha_i|Z_i^n|}\right\},$$

for all $n \in \{1, \ldots, m_i\}$, $i \in I$.

Proof. We will first verify the if part of the result. Assume that $A_i^n \geq \sum_{t \in W_i^n} w_{i,t}$ for all $n \in \{1, \ldots, m_i\}$. First, we will show that i) A_i^1 is the state at the end of the time slot $t_1 = \max Z_i^1$, given the charging strategy:

$$u_{i,t} = \min\left\{u_{i,\max}, (1 - B_i^{n-1})\frac{\beta_i}{\alpha_i|Z_i^n|}\right\}, \tag{2.7}$$

for $n = 1$, $t \in Z_i^1$, ii) $u_{i,t} \in [0, u_{i,\max}]$, for $t \in Z_i^1$, and iii) $A_i^1 \leq 1$. In order to show i), notice that the battery state at the end of t_1 can be written as:

$$\begin{aligned}
\vartheta_{i,t_1} &= \vartheta_{i,0} - \sum_{q \in W_i^0} w_{i,q} + \sum_{q \in Z_i^1} u_{i,q} \\
&= A_i^0 - \sum_{q \in W_i^0} w_{i,q} + \frac{\alpha_i}{\beta_i} \min\left\{u_{i,\max}, (1 - B_i^0)\frac{\beta_i}{\alpha_i|Z_i^1|}\right\}|Z_i^1| \\
&= B_i^0 + \frac{\alpha_i}{\beta_i} \min\left\{u_{i,\max}, (1 - B_i^0)\frac{\beta_i}{\alpha_i|Z_i^1|}\right\}|Z_i^1| \\
&= A_i^1.
\end{aligned}$$

Equalities above hold by definition of A_i^1 and B_i^0 in (2.6) and $u_{i,t}$ in (2.7) for $t \in Z_i^1$. Now, to show ii), we must prove that $1 - B_i^0 \geq 0$. This comes from the fact that $\sum_{q \in W_i^0} w_{i,q} \geq 0$,

then $B_i^0 = A_i^0 - \sum_{q \in W_i^0} w_{i,q} \leq A_i^0 \leq 1$. Then, since $u_{i,\max} > 0$, from (2.7) it follows that $u_{i,t} \geq 0$. The upper bound $u_{i,t} \leq u_{i,\max}$ follows directly from (2.7). Next, let us show iii). From (2.6) we have that:

$$\begin{aligned} A_i^1 &= B_i^0 + \frac{\alpha_i}{\beta_i} \min \left\{ u_{i,\max}, (1 - B_i^0) \frac{\beta_i}{\alpha_i |Z_i^1|} \right\} |Z_i^1| \\ &\leq B_i^0 + \frac{\alpha_i}{\beta_i} |Z_i^1| (1 - B_i^0) \frac{\beta_i}{\alpha_i |Z_i^1|} \leq 1. \end{aligned}$$

An identical analysis allows us to conclude that for $n = 2$, $A_i^2 = \vartheta_{i,t_2}$, for $t_2 = \max Z_i^2$, due to the charging strategy in (2.7), and $A_i^2 \leq 1$. Since $A_i^1 \leq 1$ and $\sum_{q \in W_i^1} w_{i,q} \geq 0$, then it holds that $1 - B_i^1 \geq 0$ and, following similar arguments as before, by (2.7), $u_{i,t} \in [0, u_{i,\max}]$, for all $t \in Z_i^2$. We can repeat this procedure recursively, to conclude that if A_i^{n-1} is the battery state at the end of $t_{n-1} = \max Z_i^{n-1}$, then A_i^n is the battery state at the end of $t_n = \max Z_i^n$, and $A_i^n \leq 1$, with $u_{i,t} \in [0, u_{i,\max}]$ for $t \in Z_i^n$, for $n \in \{1, \ldots, m_i\}$. Thus, $A_i^n = \vartheta_{i,t_n}$ for all $n \in \{1, \ldots, m_i\}$. Then, the *SOC constraints* and the upper bound on $u_{i,t}$ hold for all t, Z_i^n. The assumption $A_i^n \geq \sum_{t \in W_i^n} w_{i,t}$ for all $n \in \{1, \ldots, m_i\}$ accounts for the *user satisfaction constraints*, given that $A_i^n = \vartheta_{i,t_n}$. Then $u_{i,t}$ as defined in (2.7), is a feasible solution of the problem, hence the problem is feasible.

The only if part follows easily from the fact that $u_{i,t}$, for $t \in Z_i^n$, given by (2.7) charges the maximum amount of energy that does not violate $u_{i,t} \in [0, u_{i,\max}]$ or the *SOC constraints* during Z_i^n, for each $n \in \{1, \ldots, m_i\}$. Thus, if for some $n \in \{1, \ldots, m_i\}$, $A_i^n < \sum_{t \in W_i^n} w_{i,t}$, it is not possible to increase any $u_{i,t}$, $t \in \mathcal{Z}_i^j$, without violating the aforementioned constraints. Then, the problem is not feasible. □

The charging strategy in Lemma 2.1 consists on getting the largest amount of energy possible at each time. Therefore, if such strategy does not meet the *user satisfaction constraints*, there is no charging strategy that meets them, and the problem is infeasible.

2.1.2 Optimization problem

The PEV charging strategy is chosen to be optimal in the following way. Consider the cost function:

$$J(u) = \sum_{t\in\tau} p(D_t + \sum_{j\in I} u_{j,t})\left(D_t + \sum_{j\in I} u_{j,t}\right),$$

where $D \in \mathbb{R}^T_{\geq 0}$ is a known vector that represents the demand on the grid that comes from non-PEV loads, and $p : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is a continuous function that relates the instant demand with the electricity price. Then, finding the optimal strategy for each PEV will be equivalent to solving the following optimization problem:

$$\begin{aligned} &\min_{u\in\mathbb{R}^{NT}_{\geq 0}} J(u), \\ &\text{subject to } u \in \mathcal{F}, \end{aligned} \tag{2.8}$$

where $\mathcal{F} = \prod_{i\in I} \mathcal{F}_i$. Each $\mathcal{F}_i$ corresponds to the set of feasible charging strategies of the i^{th} PEV, i.e., $u_i \in \mathcal{F}_i \subset \mathbb{R}^T_{\geq 0}$ if and only if $u_{i,t} \in [0, u_{i,\max}]$, the *SOC constraints*, and the *user satisfaction constraints* hold.

Assumption 2.1 (Properties of the price function p)**.** *The function p is convex, strictly increasing for all its domain.* □

Notice that the cost function, along with assumptions on the function p capture the fact that the higher the demand, the higher the production cost is, and, as a consequence, the higher the end-user price becomes. This same model has been used in [22, 10, 11]. Given that the function p is convex, it follows that the cost function of (2.8) is convex, and since all the constraints are affine, the optimization problem is convex, with a convex optimal solution set.

Finally, we introduce some additional shorthand notation. For each $t \in \tau$, we let $x_t = \sum_{i\in I} u_{i,t}$, and given $u^\star$, an optimal solution to the problem in (2.8), we define $x^\star_t = \sum_{i\in I} u^\star_{i,t}$.

Remark 2.3. *Notice that, trivially, no solution is optimal if $\frac{\beta_i}{\alpha_i}\left(\sum_{q\in W_i\cup\{T+1\}} w_{i,q} - \vartheta_{i,0}\right) \leq \sum_{q\in Z_i} u_{i,q}$ does not hold with a strict equality for all $i \in I$. If the equality does not hold*

for some $i \in I$, for some optimal solution $u^\star$, there must exist some $t \in \tau$ such that the amount of energy i obtains at t, i.e., $u^\star_{i,t}$, can be decreased while all other components of $u^\star$ remain equal. The new vector would satisfy all constraints, and, by monotonicity of p, it produces a lower J, contradicting the fact that $u^\star$ is optimal. Therefore, we slightly change the user satisfaction constraints by replacing the inequality in (2.2)*, for $t = T+1$ (equivalently the inequality in* (2.4)*, for $n = m_i$) by an equality, for all $i \in I$, without modifying the solution set of the problem.* □

Next, we characterize the solution to this problem.

Lemma 2.2 (Optimal solutions)**.** *Let $u^\star$ be an optimal solution for the optimization problem in* (2.8)*. Then, for each $i \in I$:*

O1: For $s, q \in Z_i^\ell$, such that $u^\star_{i,q}, u^\star_{i,s} \in (0, u_{i,\max})$, it holds that $D_q + x^\star_q = D_s + x^\star_s$, for all $\ell \in \{1, \dots, m_i\}$.

O2: For $q \in Z_i^\ell$ such that $u^\star_{i,q} = 0$, it must be that $D_q + x^\star_q \geq D_s + x^\star_s$, for all $s \in Z_i^\ell$ such that $u^\star_{i,s} > 0$, for $\ell \in \{1, \dots, m_i\}$.

O3: For $q \in Z_i^\ell$ such that $u^\star_{i,q} = u_{i,\max}$, it must be that $D_q + x^\star_q \leq D_s + x^\star_s$, for all $s \in Z_i^\ell$ such that $u^\star_{i,s} \in (0, u_{i,\max})$, for $\ell \in \{1, \dots, m_i\}$.

O4: For all $s \in Z_i^\ell$, $q \in Z_i^n$, with $\ell, n \in \{1, \dots, m_i\}$, $n < \ell$, such that $u^\star_{i,s} \in (0, u_{i,\max})$, $u^\star_{i,q} \in [0, u_{i,\max})$, and the SOC constraints are not active, i.e., $\vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{t \in \mathcal{Z}_i^r} u^\star_{i,t} - \sum_{t \in \mathcal{W}_i^{r-1}} w_{i,t} < 1$ for all $r \in \{n, \dots, \ell - 1\}$, it must hold that $D_s + x^\star_s \leq D_q + x^\star_q$.

O5: For all $s \in Z_i^\ell$, $q \in Z_i^n$, with $\ell, n \in \{1, \dots, m_i\}$, $n > \ell$, such that $u^\star_{i,s} \in (0, u_{i,\max})$, $u^\star_{i,q} \in [0, u_{i,\max})$, and the user satisfaction constraints are not active, i.e., $\frac{\beta_i}{\alpha_i}\left(\sum_{t \in \mathcal{W}_i^{r-1}} w_{i,t} - \vartheta_{i,0}\right) < \sum_{t \in \mathcal{Z}_i^{r-1}} u^\star_{i,t}$, for all $r \in \{\ell, \dots, n-1\}$, it must hold that $D_s + x^\star_s \leq D_q + x^\star_q$.

Proof. The Lagrangian of the optimization problem in (2.8) is given by:

$$
\begin{aligned}
\mathcal{L}(u,\lambda,\eta,\mu,\nu) = J(u) &- \sum_{t\in\tau}\sum_{i\in I}\lambda_{i,t}u_{i,t} \\
&+ \sum_{i\in I}\sum_{n=1}^{m_i}\left(\sum_{t\in\mathcal{Z}_i^n} u_{i,t} - \frac{\beta_i}{\alpha_i}\left(\sum_{t\in\mathcal{W}_i^{n-1}} w_{i,t} - \vartheta_{i,0} + 1\right)\right)\xi_{i,n} \\
&- \sum_{i\in I}\sum_{n=1}^{m_i}\left(\sum_{t\in\mathcal{Z}_i^n} u_{i,t} - \frac{\beta_i}{\alpha_i}\left(\sum_{t\in\mathcal{W}_i^{n}} w_{i,t} - \vartheta_{i,0}\right)\right)\eta_{i,n} \\
&+ \sum_{i\in I}\sum_{t\in\tau}(u_{i,t} - u_{i,\max})\,\mu_{i,t} + \sum_{i\in I}\sum_{t\in W_i} u_{i,t}\nu_i \\
&+ \sum_{i\in I}\left(\sum_{t\in Z_i} u_{i,t} - \frac{\beta_i}{\alpha_i}\left(\sum_{t\in W_i\cup\{T+1\}} w_{i,t} - \vartheta_{i,0}\right)\right)\rho_i,
\end{aligned}
$$

where $\lambda_{i,t} \geq 0$, $\mu_{i,t} \geq 0$ for $t \in \tau$, $\eta_{i,n} \geq 0$, $\xi_{i,n} \geq 0$, for $n \in \{1,\ldots,m_i\}$, $\rho_i \in \mathbb{R}$, and $\nu_i \in \mathbb{R}$, for all $i \in I$ are the Lagrange multipliers. The multiplier ρ accounts for the equality constraint that we employ instead of the *user satisfaction constraints* in (2.2) for $t = T+1$, following the discussion of Remark 2.3. Define $\sigma_i = [\lambda_i^\top, \mu_i^\top, \eta_i^\top, \xi_i^\top, \rho_i, \nu_i]^\top$. We have that the Jacobian of the Lagrangian with respect to u is such that, for $i \in I$ and $t \in \tau$:

$$
\begin{aligned}
\frac{\partial\mathcal{L}}{\partial u_{i,t}} =& \frac{\partial J}{\partial u_{i,t}} - \lambda_{i,t} - \sum_{r=\mathbf{n}(i,t)}^{m_i}\eta_{i,r} \\
&+ \sum_{r=\mathbf{n}(i,t)}^{m_i}\xi_{i,r} + \mu_{i,t} + \nu_i\mathbb{1}_{W_i}(t) + \rho_i,
\end{aligned}
\tag{2.9}
$$

where recall that $\mathbf{n}(i,t)$ is the index of the set Z_i^ℓ such that $t \in Z_i^\ell$, and $\mathbb{1}_{W_i}(t)$ is the indicator function, such that $\mathbb{1}_{W_i}(t) = 1$ if $t \in W_i$, and zero otherwise, and:

$$
\frac{\partial J}{\partial u_{i,t}} = p'(D_t + \sum_{j\in I} u_{j,t})(D_t + \sum_{j\in I} u_{j,t}) + p(D_t + \sum_{j\in I} u_{j,t}).
$$

O1: For $s, q \in Z_i^\ell$, such that $u_{i,q}^\star, u_{i,s}^\star \in (0, u_{i,\max})$, applying the KKT conditions associated to the constraints $0 \leq u_{i,q} \leq u_{i,\max}$, and $0 \leq u_{i,s} \leq u_{i,\max}$, it holds that $\lambda_{i,q}^\star, \lambda_{i,s}^\star, \mu_{i,q}^\star, \mu_{i,s}^\star = 0$ by complementary slackness. Given that $q, s \in Z_i^\ell$, it follows that $q, s \notin W_i$, then $\nu_i^\star\mathbb{1}_{W_i}(q), \nu_i^\star\mathbb{1}_{W_i}(s) = 0$. Since $q, s \in Z_i^\ell$, we have $\mathbf{n}(i,q) = \mathbf{n}(i,s) = \ell$, then $\sum_{r=\mathbf{n}(i,q)}^{m_i}\eta_{i,r}^\star =$

$\sum_{r=\mathbf{n}(i,s)}^{m_i} \eta_{i,r}^\star$, and $\sum_{r=\mathbf{n}(i,q)}^{m_i} \xi_{i,r}^\star = \sum_{r=\mathbf{n}(i,s)}^{m_i} \xi_{i,r}^\star$. Then, from (2.9), by using the KKT conditions $\frac{\partial \mathcal{L}}{\partial u_{i,q}}\Big|_{\sigma_i^\star} = 0$, $\frac{\partial \mathcal{L}}{\partial u_{i,s}}\Big|_{\sigma_i^\star} = 0$, we obtain that $\frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star} = \frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star}$. Since p is increasing and convex in its argument, $\frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star}$ is also increasing and convex for all $t \in \tau$. Then it must hold that $D_q + x_q^\star = D_s + x_s^\star$.

O2: If $q \in Z_i^\ell$ such that $u_{i,q}^\star = 0$, and $s \in Z_i^\ell$ such that $u_{i,s}^\star > 0$, it holds that $\lambda_{i,s}^\star = 0$ and $\lambda_{i,q}^\star \geq 0$, by complementary slackness. Since $u_{i,q}^\star = 0$, $u_{i,\max} - u_{i,q}^\star > 0$, then, again by complementary slackness, we obtain $\mu_{i,q}^\star = 0$. By definition, $\mu_{i,s}^\star \geq 0$. As $q, s \in Z_i$, $\nu_i^\star = 0$. Since $q, s \in Z_i^\ell$, $\mathbf{n}(i,s) = \mathbf{n}(i,q)$, then $\sum_{r=\mathbf{n}(i,s)}^{m_i} \eta_{i,r}^\star = \sum_{r=\mathbf{n}(i,q)}^{m_i} \eta_{i,r}^\star$, and $\sum_{r=\mathbf{n}(i,s)}^{m_i} \xi_{i,r}^\star = \sum_{r=\mathbf{n}(i,q)}^{m_i} \xi_{i,r}^\star$. Therefore, from (2.9) and the KKT conditions $\frac{\partial \mathcal{L}}{\partial u_{i,q}}\Big|_{\sigma_i^\star} = 0$, $\frac{\partial \mathcal{L}}{\partial u_{i,s}}\Big|_{\sigma_i^\star} = 0$, we have that $\frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} + \mu_{i,s}^\star = \frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star} - \lambda_{i,q}^\star$. Then, $\frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} \leq \frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star}$, and the result follows similarly to the proof of *O1*.

O3: Using complementary slackness, we have that $\lambda_{i,q}^\star = \lambda_{i,s}^\star = 0$, $\mu_{i,q}^\star \geq 0$, while $\mu_{i,s}^\star = 0$. Since $\mathbf{n}(i,s) = \mathbf{n}(i,q)$, then $\sum_{r=\mathbf{n}(i,s)}^{m_i} \eta_{i,r}^\star = \sum_{r=\mathbf{n}(i,s)}^{m_i} \eta_{i,r}^\star$ and $\sum_{r=\mathbf{n}(i,s)}^{m_i} \xi_{i,r}^\star = \sum_{r=\mathbf{n}(i,q)}^{m_i} \xi_{i,r}^\star$. Since $q, s \notin W_i$, $\nu_i^\star = 0$. Then, using (2.9) and the KKT conditions $\frac{\partial \mathcal{L}}{\partial u_{i,q}}\Big|_{\sigma_i^\star} = 0$, $\frac{\partial \mathcal{L}}{\partial u_{i,s}}\Big|_{\sigma_i^\star} = 0$, we obtain $\frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} = \frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star} + \mu_{i,q}^\star$. Hence, we have that $\frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} \geq \frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star}$ and the result follows.

O4: By complementary slackness, $\lambda_{i,s}^\star = \mu_{i,s}^\star = \mu_{i,q}^\star = 0$ and $\xi_{i,r}^\star = 0$ for each $r \in \{n \ldots, \ell - 1\}$. It means that $\sum_{r=\ell}^{m_i} \xi_{i,r}^\star = \sum_{r=n}^{m_i} \xi_{i,r}^\star$. Since $n < \ell$, we have that $\sum_{r=n}^{m_i} \eta_{i,r}^\star = \sum_{r=\ell}^{m_i} \eta_{i,r}^\star + \sum_{r=n}^{\ell-1} \eta_{i,r}^\star$. From (2.9) and the KKT condition $\frac{\partial \mathcal{L}}{\partial u_{i,q}}\Big|_{\sigma_i^\star} = 0$, we obtain $\frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star} - \sum_{r=\ell}^{m_i} \eta_{i,r}^\star - \sum_{r=n}^{\ell-1} \eta_{i,r}^\star - \lambda_{i,q}^\star = 0$. From (2.9) and the KKT condition $\frac{\partial \mathcal{L}}{\partial u_{i,s}}\Big|_{\sigma_i^\star} = 0$, we have that $\frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} - \sum_{r=\ell}^{m_i} \eta_{i,r}^\star = 0$. We connect these two expressions to obtain $\frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star} - \sum_{r=n}^{\ell-1} \eta_{i,r}^\star - \lambda_{i,q}^\star = \frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star}$. Then, it must hold that $\frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} \leq \frac{\partial J}{\partial u_{i,q}}\Big|_{\sigma_i^\star}$, and the result follows.

O5: By complementary slackness, we have $\lambda_{i,s}^\star = \mu_{i,s}^\star = \mu_{i,q}^\star = 0$ and $\eta_{i,r}^\star = 0$ for each $r \in \{\ell, \ldots, n-1\}$. Therefore, $\sum_{r=\ell}^{m_i} \eta_{i,r}^\star = \sum_{r=n}^{m_i} \eta_{i,r}^\star$. Then $\sum_{r=\ell}^{m_i} \xi_{i,r}^\star = \sum_{r=n}^{m_i} \xi_{i,r}^\star + \sum_{r=\ell}^{n-1} \xi_{i,r}^\star$. We can proceed as in the proof of *O4*, by using the KKT conditions $\frac{\partial \mathcal{L}}{\partial u_{i,q}}\Big|_{\sigma_i^\star} = 0$, $\frac{\partial \mathcal{L}}{\partial u_{i,s}}\Big|_{\sigma_i^\star} = 0$, to obtain $\frac{\partial J}{\partial u_{i,q}} - \lambda_{i,q}^\star\Big|_{\sigma_i^\star} = \frac{\partial J}{\partial u_{i,s}}\Big|_{\sigma_i^\star} + \sum_{r=\ell}^{n-1} \xi_{i,r}^\star$, then it follows that $D_q + x_q^\star \geq D_s + x_s^\star$. □

Given $u^\star$ we define a partition of the set τ, denoted as $\{\Upsilon_l\}_{l=1}^{m+1}$, $m \leq T$, corresponding to the connected clusters of times where the load price is constant at the

optimal solution, and there is PEV load greater than zero. In other words, for any pair $t, q \in \Upsilon_l$, $l \in \{1, \dots, m\}$, it holds that $x_t^\star, x_q^\star > 0$ and $p(D_t + x_t^\star) = p(D_q + x_q^\star)$. The set Υ_{m+1} consists of those $t \in \tau$ such that $x_t^\star = 0$. The collection $\{\Upsilon_l\}_{l=1}^{m+1}$ is ordered according to the corresponding price values, from the cheapest to the priciest. In other words, $l_1 < l_2$, $l_1, l_2 \in \{1, \dots, m\}$, if and only if $p(D_t + x_t^\star) < p(D_q + x_q^\star)$ for all $t \in \Upsilon_{l_1}$, and $q \in \Upsilon_{l_2}$.

Lemma 2.3 (Uniqueness of $x^\star$, $\Upsilon_1, \dots, \Upsilon_{m+1}$)**.** *Let $u^\star$, and $v^\star$ be optimal solutions of problem in* (2.8) *with associated aggregated loads $x_{1,t}^\star \triangleq \sum_{i \in I} u_{i,t}^\star$ and $x_{2,t}^\star \triangleq \sum_{i \in I} v_{i,t}^\star$, respectively. Then $x_{1,t}^\star = x_{2,t}^\star$, for all $t \in \tau$, and $\{\Upsilon_l\}_{l=1}^{m+1}$ is unique.*

Proof. The proof is similar to that of Theorem 1 in [11]. We only include it here for the sake of completeness.

With a slight abuse of notation, denote $J(x) = \sum_{t \in \tau} p(D_t + x_t)(D_t + x_t)$. Consider the set $\mathcal{B} = \{x \in \mathbb{R}_{\geq 0}^T \mid \exists u \in \mathcal{F} \text{ s.t. } x = \sum_{i \in I} u_i\}$. If $u^\star$ and $v^\star$ are optimizers of J in $\mathcal{F}$, then $x_1^\star$ and $x_2^\star$ are optimizers of J in $\mathcal{B}$. Then, by the convexity of J in x, it must hold that $\nabla J(x_1^\star) \cdot (x - x_1^\star) \geq 0$, $\nabla J(x_2^\star) \cdot (x - x_2^\star) \geq 0$ for all $x \in \mathcal{B}$. Therefore, in particular it must hold that $\nabla J(x_1^\star) \cdot (x_2^\star - x_1^\star) \geq 0$, $\nabla J(x_2^\star) \cdot (x_1^\star - x_2^\star) \geq 0$. Hence, $(\nabla J(x_1^\star) - \nabla J(x_2^\star)) \cdot (x_2^\star - x_1^\star) \geq 0$. We can write this expression out as:

$$\sum_{t \in \tau} (\nabla_t J(x_1^\star) - \nabla_t J(x_2^\star))(x_{2,t}^\star - x_{1,t}^\star) \geq 0, \tag{2.10}$$

where $\nabla_t J(x) = p'(D_t + x_t)(D_t + x_t) + p(D_t + x_t)$, for all $t \in \tau$. Since p is convex, then p' is increasing in its argument, and given that p is also increasing, it follows that $\nabla_t J$ is strictly increasing in x_t. Then, we have that $(\nabla_t J(x_1^\star) - \nabla_t J(x_2^\star))(x_{1,t}^\star - x_{2,t}^\star) \geq 0$ for all $t \in \tau$; hence, from (2.10), $\sum_{t \in \tau} (\nabla_t J(x_1^\star) - \nabla_t J(x_2^\star))(x_{1,t}^\star - x_{2,t}^\star) = 0$, and the result follows. □

Lemma 2.3 allows us to establish that if a charging profile u does not satisfy that $\sum_{i \in I} u_{i,t} = x^\star$, then u is not optimal.

Lemma 2.4 (Properties of optimal solutions)**.** *Consider any $t \in \Upsilon_l$, such that for some $i \in I$, $u_{i,t}^\star > 0$. Let $\ell \in \{1, \dots, m_i\}$ be such that $t \in Z_i^\ell$. Then, the following holds:*

C1: If there exists $q \in Z_i^n \cap \bigcup_{s=1}^{l-1} \Upsilon_s$*, for some* $n \in \{1, \ldots, \ell - 1\}$*, then it must be that at least one of the following conditions hold: i)* $u_{i,q}^\star = u_{i,\max}$*, or ii) the maximum battery capacity constraint is active, i.e.,* $\vartheta_{i,t_r} = \vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{s \in \mathcal{Z}_i^r} u_{i,s}^\star - \sum_{s \in \mathcal{W}_i^{r-1}} w_{i,s} = 1$*, for some* $r \in \{n, \ldots, \ell - 1\}$*.*

C2: If there exists $q \in Z_i^\ell \cap \bigcup_{s=1}^{l-1} \Upsilon_s$*, then it must be that* $u_{i,q}^\star = u_{i,\max}$*.*

C3: If there exists $q \in Z_i^n \cap \bigcup_{s=1}^{l-1} \Upsilon_s$*, for some* $n \in \{\ell + 1, \ldots, m_i\}$*, then, it must be that at least one of the following conditions hold: i)* $u_{i,q}^\star = u_{i,\max}$*, or ii) the user satisfaction constraint is active, i.e.,* $\frac{\beta_i}{\alpha_i} \left(\sum_{q \in \mathcal{W}_i^{r-1}} w_{i,q} - \vartheta_{i,0} \right) = \sum_{q \in \mathcal{Z}_i^{r-1}} u_{i,q}^\star$*, for some* $r \in \{\ell + 1, \ldots, n\}$*.*

Proof. Assume that for $t \in \Upsilon_l \cap Z_i^\ell$, for some $\ell \in \{1, \ldots, m_i\}$, for $i \in I$, it holds that $u_{i,t}^\star > 0$.

C1: We employ a contradiction argument. Assume that there exists $q \in Z_i^n \cap \bigcup_{s=1}^{l-1} \Upsilon_s$, for some $n \in \{1, \ldots, \ell - 1\}$, with $u_{i,q}^\star \in [0, u_{i,\max})$, and $\vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{s \in \mathcal{Z}_i^r} u_{i,s} - \sum_{s \in \mathcal{W}_i^{r-1}} w_{i,s} < 1$, i.e., the *maximum battery capacity constraint* is not active for all $r \in \{n, \ldots, \ell - 1\}$. By definition of $\{\Upsilon_l\}_{l=1}^{m+1}$, we have that since $q \in \Upsilon_s$, $t \in \Upsilon_l$, and $s < l$, then $D_q + x_q^\star < D_t + x_t^\star$. On the other hand, the assumptions satisfy *O4* in Lemma 2.2, then we have that $D_q + x_q^\star \geq D_t + x_t^\star$, which is a contradiction.

C2: We employ a contradiction argument again. Assume that there exists a $q \in Z_i^\ell \cap \bigcup_{s=1}^{l-1} \Upsilon_s$, such that $u_{i,q}^\star \in [0, u_{i,\max})$. We must consider two cases: i) $u_{i,t}^\star = u_{i,\max}$, and ii) $u_{i,t}^\star < u_{i,\max}$. Let us consider the first case. Since the assumptions satisfy *O3* in Lemma 2.2, we have that $D_t + x_t^\star \leq D_q + x_q^\star$. Recall that since $q \in \Upsilon_s$, $t \in \Upsilon_l$, and $s < l$, then $D_q + x_q^\star < D_t + x_t^\star$, a contradiction. In the second case, the assumptions hold for *O1* in Lemma 2.2, then $D_q + x_q^\star = D_t + x_t^\star$, which contradicts the fact that $D_q + x_q^\star < D_t + x_t^\star$, given by $q \in \Upsilon_s$, $t \in \Upsilon_l$, and $s < l$.

O3: Once more we proceed by contradiction. Assume that there exists a $q \in Z_i^r \cap \bigcup_{s=1}^{l-1} \Upsilon_s$, for some $r \in \{1, \ldots, \ell\}$, such that $u_{i,q}^\star \in (0, u_{i,\max})$, and $\frac{\beta_i}{\alpha_i} \left(\sum_{q \in \mathcal{W}_i^{r-1}} w_{i,q} - \vartheta_{i,0} \right) < \sum_{q \in \mathcal{Z}_i^{r-1}} u_{i,q}^\star$, i.e., it holds that the *user satisfaction constraint* is not active for all $r \in \{\ell + 1, \ldots, n\}$. All these assumptions allow us to employ *O5* in Lemma 2.2, to conclude that $D_q + x_q^\star \geq D_t + x_t^\star$. However, since $q \in \Upsilon_s$, $t \in \Upsilon_l$, and $s < l$, it must be that $D_q + x_q^\star < D_t + x_t^\star$, reaching a contradiction. □

The following result implies that each PEV will charge as much as possible at the times for which the prices are the lowest, provided no constraints are violated. In other words, if a PEV charges at time slots with higher prices, it is because at the lower-price slots at least one constraint is becoming active. This comes in handy in the proof of convergence of the proposed Price Leveling algorithm.

Lemma 2.5 (Properties of optimal solutions)**.** *For each Υ_l, $l \in \{1, \dots, m\}$, and any feasible u, if $x_t = x_t^\star$ for all $t \in \Upsilon_r$, for all $r < l$, then it holds that $\sum_{t\in\Upsilon_l} x_t \leq \sum_{t\in\Upsilon_l} x_t^\star$.*

Proof. Due to the extension of the proof we show a simpler form of it, assuming that for any $i \in I$, $m_i = 2$. Define $\Theta_l = \bigcup_{r=l}^{m+1} \Upsilon_r$. First, we show that if $\sum_{t\in\tau\setminus\Theta_l} x_t > \sum_{t\in\tau\setminus\Theta_l} x_t^\star$ for some feasible u, then we reach a contradiction with either the feasibility of u or with a property of the optimal solutions that has been established in Lemma 2.4. Then, we will be able to conclude our result.

Assume that for some feasible u, $\sum_{t\in\tau\setminus\Theta_l} x_t > \sum_{t\in\tau\setminus\Theta_l} x_t^\star$. Since $u_{i,t}^\star \geq 0$ for all $t \in \tau$, $i \in I$, it must be that there exists some $i \in I$ such that $\sum_{t\in\tau\setminus\Theta_l} u_{i,t} > \sum_{t\in\tau\setminus\Theta_l} u_{i,t}^\star$. Since $u_{i,t} = 0$, $u_{i,t}^\star = 0$ for all $t \notin Z_i$, we have that $\sum_{t\in Z_i\setminus\Theta_l} u_{i,t} > \sum_{t\in Z_i\setminus\Theta_l} u_{i,t}^\star$. Hence, there must be some $\ell \in \{1, 2\}$ such that $\sum_{t\in Z_i^\ell\cap(\tau\setminus\Theta_l)} u_{i,t} > \sum_{t\in Z_i^\ell\cap(\tau\setminus\Theta_l)} u_{i,t}^\star$. In this way, there exists $\underline{t} \in Z_i^\ell \cap (\tau \setminus \Theta_l)$, such that $u_{i,\max} \geq u_{i,\underline{t}} > u_{i,\underline{t}}^\star$.

On the other hand, since $\sum_{t\in Z_i} u_{i,t} = \sum_{t\in Z_i} u_{i,t}^\star$, which follows from Remark (2.3), then it must hold that $\sum_{t\in\Theta_l} u_{i,t} < \sum_{t\in\Theta_l} u_{i,t}^\star$. Furthermore, it must be that for some $n \in \{1, 2\}$, $\sum_{t\in Z_i^n\cap\Theta_l} u_{i,t} < \sum_{t\in Z_i^n\cap\Theta_l} u_{i,t}^\star$, and in particular, for some $\bar{t} \in Z_i^n \cap \Theta_l$, $0 \leq u_{i,\bar{t}} < u_{i,\bar{t}}^\star$.

First assume that $\ell = n$. By Lemma 2.4, Part *C2*, since $u_{i,\bar{t}}^\star > 0$, and $\bar{t} \in \Theta_l \cap Z_i^n$, it must be that for all $t \in Z_i^n \setminus \Theta_l$, $u_{i,t}^\star = u_{i,\max}$. However, by the assumption that $\ell = n$, $\underline{t} \in Z_i^n \setminus \Theta_l$, then, $u_{i,\underline{t}}^\star = u_{i,\max}$. However, by definition of $\underline{t}$, it must hold that $u_{i,\max} \geq u_{i,\underline{t}} > u_{i,\underline{t}}^\star$, a contradiction.

Now, let us assume that $\ell \neq n$. We have two cases: i) $\ell = 1$, $n = 2$ and ii) $n = 1$, $\ell = 2$.

Suppose that $\ell = 1$, $n = 2$. Since $u_{i,\bar{t}}^\star > 0$, $\bar{t} \in Z_i^2 \cap \Theta_l$ and $\underline{t} \in Z_i^1 \setminus \Theta_l$ we can employ Part *C1* in Lemma 2.4, and given that $u_{i,\max} \geq u_{i,\underline{t}} > u_{i,\underline{t}}^\star$, it must hold that:

$$\vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{t\in\mathcal{Z}_i^1} u_{i,t}^\star - \sum_{t\in\mathcal{W}_i^0} w_{i,t} = 1. \tag{2.11}$$

Moreover, notice that if there is $q \in Z_i^1 \cap \Theta_l$ such that $u_{q,t}^\star > 0$, then by Part *C2* in Lemma 2.4, $u_{i,\underline{t}}^\star = u_{i,\max}$, which is a contradiction. Then, it must be that for all $q \in Z_i^1 \cap \Theta_l$, $u_{q,t}^\star = 0$. Then, given that $\sum_{t \in Z_i^\ell \cap (\tau \setminus \Theta_l)} u_{i,t} > \sum_{t \in Z_i^\ell \cap (\tau \setminus \Theta_l)} u_{i,t}^\star$, we obtain:

$$\sum_{t \in Z_i^1} u_{i,t} > \sum_{t \in Z_i^1} u_{i,t}^\star.$$

Then, from (2.11), it follows that:

$$\vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{t \in \mathcal{Z}_i^1} u_{i,t} - \sum_{t \in \mathcal{W}_i^0} w_{i,t} > 1,$$

which means that u does not satisfy the *maximum battery capacity constraint* in (2.5), a contradiction with the feasibility of u.

Now, let us assume that $n = 1$, $\ell = 2$. Similarly to the previous case, we can see that it is possible to apply Part *C3* in Lemma 2.4, and given that $u_{i,\max} \geq u_{i,\underline{t}} > u_{i,\underline{t}}^\star$, we conclude that:

$$\vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{t \in \mathcal{Z}_i^1} u_{i,t}^\star = \sum_{t \in \mathcal{W}_i^1} w_{i,t}. \tag{2.12}$$

Now, by Part *C2* in Lemma 2.4, it holds that if there is $q \in Z_i^1 \setminus \Theta_l$, then, it must be that $u_{i,q}^\star = u_{i,\max}$. Then, since $\sum_{t \in Z_i^n \cap \Theta_l} u_{i,t}^\star > \sum_{t \in Z_i^n \cap \Theta_l} u_{i,t}$, it must hold that:

$$\sum_{t \in Z_i^1} u_{i,t}^\star > \sum_{t \in Z_i^1} u_{i,t}.$$

Therefore, from (2.12), we obtain that:

$$\vartheta_{i,0} + \frac{\alpha_i}{\beta_i} \sum_{t \in \mathcal{Z}_i^1} u_{i,t} - \sum_{t \in \mathcal{W}_i^1} w_{i,t} < 0,$$

which means that u does not satisfy the *user satisfaction constraint* in (2.4), a contradiction.

We have shown that $\sum_{t \in \tau \setminus \Theta_r} x_t \leq \sum_{t \in \tau \setminus \Theta_r} x_t^\star$, for all $r \in \{1, \ldots, m\}$. Now, let us assume as in the statement of the lemma that for some $l \in \{1, \ldots, m\}$, and a feasible

solution u, $x_t = x_t^\star$ for all $t \in \bigcup_{s=1}^{l-1} \Upsilon_s$. Then, since $\bigcup_{s=1}^{l-1} \Upsilon_s = \tau \setminus \Theta_l$, we can write $\tau \setminus \Theta_{l+1} = (\tau \setminus \Theta_l) \cup \Upsilon_l$. This leads to:

$$\begin{aligned}
&\sum_{t\in\tau\setminus\Theta_{l+1}} x_t \leq \sum_{t\in\tau\setminus\Theta_{l+1}} x_t^\star \\
&\sum_{t\in\tau\setminus\Theta_l} x_t + \sum_{t\in\Upsilon_l} x_t \leq \sum_{t\in\tau\setminus\Theta_l} x_t^\star + \sum_{t\in\Upsilon_l} x_t^\star \\
&\sum_{t\in\Upsilon_l} x_t \leq \sum_{t\in\Upsilon_l} x_t^\star,
\end{aligned}$$

completing the proof. □

2.2 Price Leveling algorithm for PEV charging

In this section, we introduce a dynamics that allows agents to reach a charging profile that is arbitrarily close to an optimal solution of the optimization problem in (2.8) by means of local interactions, i.e., in a hierarchical way. This dynamics is a modification of the Price Leveling algorithm that was introduced in [12], to account for usage schedule constraints and battery physical constraints. These constraints do not exist in the problem presented in [12], because in that work the charging objective is to obtain a full charge by the end of the day, without bounds on the charging rate. The algorithm follows a load-balancing policy by moving energy consumption from time slots with higher prices to slots with lower prices.

In order to achieve the optimal solution, the set of PEVs performs an iterative process which includes two steps at each iteration $k \in \mathbb{N}$: i) each PEV transmits its charging profile $u_i^k \in \mathbb{R}^T_{\geq 0}$ to the utility, and ii) the utility aggregates the charging profiles from all PEVs, and adds them to the non-PEV demand, to further use the pricing function p for determining and transmitting the price per kWh at each time $t \in \tau$, which is represented by $p^k \triangleq [p_1^k, \ldots, p_T^k]^\top \in \mathbb{R}^T_{\geq 0}$, to all PEVs, where:

$$p_t^k \triangleq p(D_t + \sum_{i\in I} u_{i,t}^k).$$

Then, for the next iteration, the PEVs use the pricing signal to update their charging

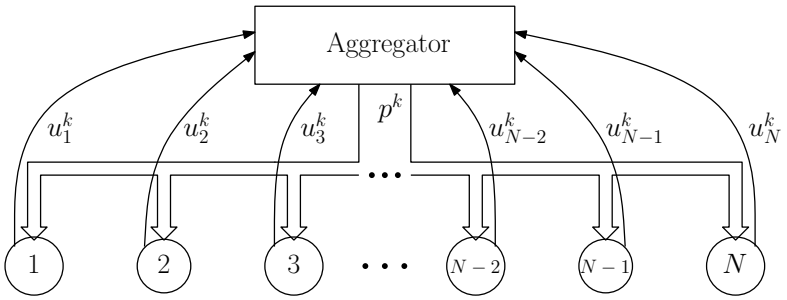

Figure 2.2: Communication scheme for the Price Leveling algorithm. The circles represent all PEVs, and the arrows represent the information that is sent by each element at iteration k.

profiles, and repeat the process. Most of the computation required by the algorithm is carried out in a parallel way by processors at the PEVs. The aggregator only has to collect the charging information, aggregate it and compute the pricing signal for the PEVs' coordination. Figure 2.2 shows how the information flows from each PEV towards the aggregator in the utility and back. By means of our scheme, the aggregator does not need to know the usage schedule of any of the PEVs, since they compute their own charging profile, and neither other PEVs need to know this from other PEVs. In this way, i) the scheme preserves user's privacy and ii) it is computationally scalable, since there is no computational procedure at any of the components of the system that grows significantly as the number of PEVs increases. This communication structure is widespread in demand-dispatch papers [13, 14], as well as in PEV coordination works [11, 10], where algorithms are termed as "decentralized." To alleviate the communication burden between the aggregator and PEVs, several several layers of aggregators transmitting aggregated signal to parent aggregators can be considered.

Next, let us introduce the computation on the PEV side for each iteration of the algorithm. For each $i \in I$, and for each $t \in \tau$, define the vector $y_{i,t} = [y_{i,t,1}, \ldots, y_{i,t,m_i}]^\top \in \mathbb{R}^{m_i}_{\geq 0}$, the vector $y_i = [y_{i,1}^\top, \ldots, y_{i,T}^\top]^\top \in \mathbb{R}^{T \cdot m_i}_{\geq 0}$, and finally $y = [y_1^\top, \ldots, y_N^\top]^\top \in \mathbb{R}^d_{\geq 0}$, where $d = NT \sum_{i \in I} m_i$. Each component of this vector is a state of the algorithm, related to $t \in \tau$, and $i \in I$, and:

$$u_{i,t}^k = \sum_{\ell=1}^{m_i} y_{i,t,\ell}^k.$$

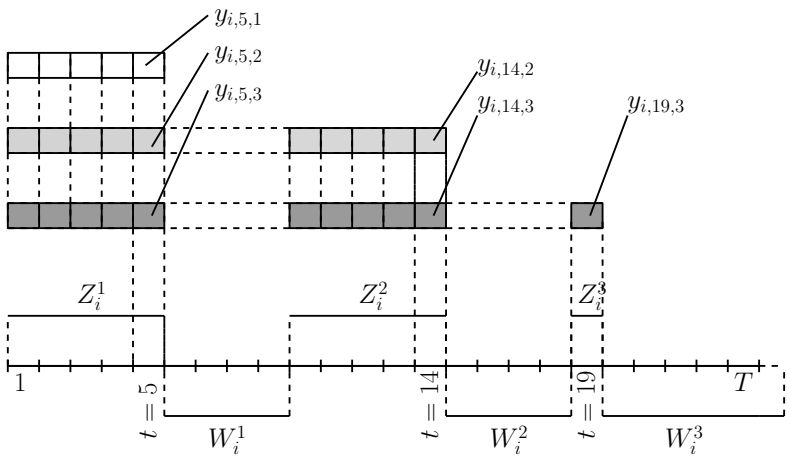

Figure 2.3: Algorithm state and its relation with the PEV charging problem, for a fixed PEV $i \in I$. Each color represents variables $y_{i,t,\ell}$ for all $t \in Z_i^\ell$, for a fixed $\ell \in \{1, \ldots, m_i\}$.

Figure 2.3 shows how different states of the algorithm are related to time slots and charging strategies of the PEV charging problem, for a fixed PEV $i \in I$.

The introduction of y is based on the following idea. Note that in Figure 2.3 there are several rows in different colors. Each level is aimed to satisfy a different usage constraint; e.g. the darker grey is for W_i^3, the lighter grey is for W_i^2, and the white color is for W_i^1. In this way, the further away W_i^ℓ is, the more opportunity there is to charge during previous times and so the length of these rows is larger. The column on top of each time t represents how the amount charged in the battery at time t is divided between the upcoming usage constraints. That is, for $t = 5$, we have an amount $y_{i,5,1}$, to satisfy W_i^1, a $y_{i,5,2}$ for W_i^2, and an amount $y_{i,5,3}$ for W_i^3.

The dynamics governing these states is aimed to update each part of the vector y to satisfy the constraints while minimizing price. Mathematically,

$$y_{i,t,\ell}^{k+1} = \begin{cases} f_{i,t,\ell}(y^k), & \text{if } t \in \mathcal{Z}_i^\ell \\ 0, & \text{otherwise,} \end{cases} \tag{2.13}$$

where $f_{i,t,\ell}(y^k)$ is given by:

$$f_{i,t,\ell}(y^k) = y^k_{i,t,\ell} + \sum_{q \in \mathcal{Z}_i^\ell} \Delta^k_{i,\ell}(q,t)\psi_i(\max\{p^k_q - p^k_t, 0\}) - \sum_{q \in \mathcal{Z}_i^\ell} \Delta^k_{i,\ell}(t,q)\psi_i(\max\{p^k_t - p^k_q, 0\}), \tag{2.14}$$

for all $i \in I$, $\ell \in \{1, \ldots, m_i\}$, and $t \in \mathcal{Z}_i^\ell$, where:

$$\Delta^k_{i,\ell}(q,t) = \frac{\min\{y^k_{i,q,\ell}, \gamma^k_1(i,t), \gamma^k_2(i,t,q)\}}{T},$$

and $\gamma^k_1(i,t), \gamma^k_2(i,t,q)$ are given by:

$$\gamma^k_1(i,t) = T\frac{u_{i,\max} - u^k_{i,t}}{\sum_{s=\mathbf{n}(i,t)}^{m_i} \left|\mathcal{Z}_i^s\right|},$$

$$\gamma^k_2(i,t,q) = \begin{cases} T \min\limits_{r \in \{\mathbf{n}(i,t),\ldots,\mathbf{n}(i,q)\}} \dfrac{\frac{\beta_i}{\alpha_i}(1 - \vartheta^k_{i,t_r})}{|\mathcal{Z}_i^r| \sum_{s=r}^{m_i} \left|\mathcal{Z}_i^s \setminus Z_i^r\right|}, \\ \qquad\qquad \text{if } \mathbf{n}(i,q) > \mathbf{n}(i,t) \\ +\infty, \\ \qquad\qquad \text{if } \mathbf{n}(i,q) \leq \mathbf{n}(i,t). \end{cases}$$

The function $\psi_i : \mathbb{R}_{\geq 0} \to [0,1]$ is continuous, increasing on its argument, with $\psi_i(0) = 0$. The value of $\gamma^k_1(i,t)$ is associated to how much the i^{th} PEV can increase $u_{i,t}$ without violating $u_{i,t} \leq u_{i,\max}$, and $\gamma^k_2(i,t,q)$ is related to how much the amount of charge that i obtains at any $t \in Z_i$, can increase without leading to $\vartheta_{i,t} > 1$, for some $q \in \tau$, $\mathbf{n}(i,q) > \mathbf{n}(i,t)$. Notice that the algorithm is trying to implement a load balancing protocol over the time slots in τ, subject to constraints and driven by the price difference between slots.

Although the PEV charging problem introduces several constraints, we present a way of choosing the algorithm initial state, so that $u^0 \in \mathcal{F}$, provided the PEV charging problem is feasible. The procedure is described in Algorithm 1. Loosely speaking, the algorithm finds the first element (denoted r) in the partition $\{Z_i^s\}_{s=1}^{m_i}$ such that W_i^r cannot be satisfied only using the initial energy stored in the battery. It implies that each PEV has to obtain energy to satisfy usage periods given by W_i^ℓ, for $\ell \in \{r, \ldots, n_i\}$.

Algorithm 1 Initial Conditions

if $\vartheta_{i,0} \geq \sum_{t\in W_i} w_{i,t}$ **then**
  $y^0_{i,t,\ell} = 0$ for all $t \in Z^\ell_i$, for all $\ell \in \{1,\dots,m_i\}$
else
  for $\ell \in \{r,\dots,m_i\}$ **do**
    for $n \in \{1,\dots,\ell\}$ **do**
      $B^n_i = A^{n-1}_i - \sum_{t\in W^{n-1}_i} w_{i,t} + \frac{\alpha_i}{\beta_i} \sum_{\ell\in\{r,\dots,\ell-1\}} \sum_{t\in Z^n_i} y^0_{i,t,\ell}$
      $y^0_{i,t,\ell} = \min(u_{i,\max}, (1 - B^n_i)\frac{\beta_i}{\alpha_i|Z^n_i|}, \frac{\beta_i}{\alpha_i}\frac{R^\ell_i}{|Z^n_i|})$, for all $t \in Z^n_i$
      $R^\ell_i = R^\ell_i - \frac{\alpha_i}{\beta_i} \sum_{t\in Z^n_i} y^0_{i,t,\ell}$

      $A^n_i = B^n_i + \frac{\alpha_i}{\beta_i} \sum_{t\in Z^n_i} y^0_{i,t,\ell}$
    end for
  end for
end if

Then it distributes the amount of energy required during W^ℓ_i, among states $y_{i,t,\ell}$ for all $t \in \mathcal{Z}^\ell_i$. This distribution is done by assigning identical values to all $y_{i,t,\ell}$, for $t \in Z^n_i$, and all $n \in \{1,\dots,\ell\}$ in ascending order, in a way that neither the constraint on charging rate nor the constraint on the battery size are violated. The assignment is also done in ascending order for $\ell \in \{r,\dots,n_i\}$. A rigorous proof of the feasibility of this solution, follows along the lines of Lemma 2.1.

By means of the next result, we state the fact that if u^0 is a feasible charging profile, then for any iteration of the algorithm, u^k is a feasible profile.

Lemma 2.6 (Invariance of the set $\mathcal{F}$)**.** *The set $\mathcal{F}$ is invariant under the dynamics defined in* (2.14)*.*

Proof. For the proof of this result, we need to show that if u^k is feasible, then, for u^{k+1}, it holds that i) the constraint in (2.4) is satisfied for $n \in \{1,\dots,m_i\}$, $i \in I$, ii) $u_{i,t} \in [0, u_{i,\max}]$ for all $t \in \tau$, $i \in$, and iii) the constraint in (2.5) holds for all $n \in \{1,\dots,m_i\}$, $i \in I$.

First, it is straightforward to show that if u^k is such that the *user satisfaction constraint* is satisfied, for all $i \in I$, $t \in \tau$, and $\ell \in \{1,\dots,m_i\}$, then, it also holds for u^{k+1}. To see this, it is enough to perform a sum on both sides of the equality in (2.14) for all $t \in \mathcal{Z}^\ell_i$, to obtain $\sum_{q\in\mathcal{Z}^\ell_i} y^{k+1}_{i,q,\ell} = \sum_{q\in\mathcal{Z}^\ell_i} y^k_{i,q,\ell}$, and the result follows.

Next, since the sum with negative sign on the right-hand side of (2.14) is less or equal than $y^k_{i,t,\ell}$, then $y^{k+1}_{i,t,\ell} \geq 0$, hence $u^{k+1}_{i,t} \geq 0$, for all $i \in I$, $t \in \tau$. Now, let us show that

if u^k satisfies that $u^k_{i,t} \leq u_{i,\max}$, then $u^{k+1}_{i,t} \leq u_{i,\max}$, for all $i \in I$, $t \in \tau$. Since the second sum in (2.14) is nonpositive, and ψ_i is upper-bounded by 1, then we have:

$$y^{k+1}_{i,t,\ell} \leq y^k_{i,t,\ell} + \sum_{q \in \mathcal{Z}^\ell_i} \frac{u_{i,\max} - u^k_{i,t}}{\sum_{s=\mathbf{n}(i,t)}^{m_i} \left|\mathcal{Z}^s_i\right|}.$$

By definition of $y^k_{i,t,\ell}$, we have that $\sum_{\ell=1}^{m_i} y^k_{i,t,\ell} = \sum_{\ell=\mathbf{n}(i,t)}^{m_i} y^k_{i,t,\ell}$. By summing both sides of this expression in $\ell \in \{\mathbf{n}(i,t), \ldots, m_i\}$, we obtain:

$$u^{k+1}_{i,t} \leq u^k_{i,t} + \frac{u_{i,\max} - u^k_{i,t}}{\sum_{s=\mathbf{n}(i,t)}^{m_i} \left|\mathcal{Z}^s_i\right|} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \left|\mathcal{Z}^\ell_i\right| = u_{i,\max}.$$

Finally, we show that if y^k is such that u^k satisfies the *maximum battery capacity constraint*, for all $r \in \{1, \ldots, m_i - 1\}$, for $i \in I$, then, the *maximum battery capacity constraint* is also satisfied by u^{k+1}. Notice that from (2.13), it follows that:

$$\begin{aligned}\sum_{t \in \mathcal{Z}^r_i} u^{k+1}_{i,t} \leq & \sum_{t \in \mathcal{Z}^r_i} u^k_{i,t} \\ & + \sum_{t \in \mathcal{Z}^r_i} \sum_{\ell=r+1}^{m_i} \sum_{q \in \mathcal{Z}^\ell_i \setminus \mathcal{Z}^r_i} \Delta^k_{i,\ell}(q,t) \psi_i(\max\{p^k_q - p^k_t, 0\}).\end{aligned}$$

Since $\ell \in \{r+1, \ldots, m_i\}$ in the sum above, then it follows that $r \in \{\mathbf{n}(i,t), \ldots, \mathbf{n}(i,q)\}$ and given that $q \in \mathcal{Z}^\ell_i \setminus \mathcal{Z}^r_i$, it follows that $\mathbf{n}(i,q) > \mathbf{n}(i,t)$. Therefore, given the definition of $\Delta^k_{i,\ell}(q,t)$, we can upper-bound it by $\gamma^k_2(i,t,q)$, to obtain:

$$\begin{aligned}\sum_{t \in \mathcal{Z}^r_i} u^{k+1}_{i,t} &\leq \sum_{t \in \mathcal{Z}^r_i} u^k_{i,t} + \sum_{t \in \mathcal{Z}^r_i} \sum_{\ell=r+1}^{m_i} \sum_{q \in \mathcal{Z}^\ell_i \setminus \mathcal{Z}^r_i} \frac{\frac{\beta_i}{\alpha_i}(1 - \vartheta^k_{i,t_r})}{|\mathcal{Z}^r_i| \sum_{s=r}^{m_i} \left|\mathcal{Z}^s_i \setminus Z^r_i\right|} \\ &= \sum_{t \in \mathcal{Z}^r_i} u^k_{i,t} + \frac{\frac{\beta_i}{\alpha_i}(1 - \vartheta^k_{i,t_r})}{|\mathcal{Z}^r_i| \sum_{s=r}^{m_i} \left|\mathcal{Z}^s_i \setminus Z^r_i\right|} \left|\mathcal{Z}^r_i\right| \sum_{\ell=r+1}^{m_i} \left|\mathcal{Z}^\ell_i \setminus \mathcal{Z}^r_i\right|.\end{aligned}$$

Then, from the definition of ϑ^k_{i,t_n}, it follows that:

$$\sum_{t \in \mathcal{Z}^r_i} u^{k+1}_{i,t} \leq \frac{\beta_i}{\alpha_i}(1 - \vartheta_{i,0}) + \frac{\beta_i}{\alpha_i} \sum_{s \in \mathcal{W}^{r-1}_i} w_{i,s},$$

which corresponds exactly to the *maximum battery capacity constraint*, for $r \in \{1, \ldots, m_i -$

1}. For $r = m_i$, the constraint is satisfied given that $\sum_{q\in\mathcal{Z}_i^{m_i}} y_{i,q,\ell}^{k+1} = \sum_{q\in\mathcal{Z}_i^{m_i}} y_{i,t,\ell}^{k}$, then we sum on both sides over $\ell \in \{1,\dots,m_i\}$, and the result follows. □

Let us define the set $\overline{\mathcal{F}} = \{y \in \mathbb{R}^d \mid u \in \mathcal{F}, u_{i,t} = \sum_{\ell=1}^{m_i} y_{i,t,\ell}\}$. This is the set of admissible states of the Price Leveling algorithm, meaning that it comprises all states that define a feasible charging profile. By Lemma 2.6, $\overline{\mathcal{F}}$ is invariant under the dynamics of the Price Leveling algorithm.

Let us define $L_t = D_t + x_t$, for all $t \in \tau$ and $L_{\min} = \min_{t\in\tau}(D_t + x_t)$. Notice that from Remark 2.3, the amount of charge the set of PEVs takes during the time window is finite. Then, there exists an upper bound $\overline{L}$ such that $\overline{L} \geq L_t^k$ for all $t \in \tau$, $k \in \mathbb{N}$. The following is a sufficient condition that allows us to prove convergence of the algorithm.

Assumption 2.2 (Properties of the functions ψ_i)**.** *Let $x_{\max}$ be such that $x_t \leq x_{\max}$, for all $t \in \tau$. The function ψ_i is Lipschitz, with Lipschitz constant r_ψ^i such that $r_\psi^i < T/((T-1)x_{\max}p'(\overline{L}))$, for all $i \in I$.* □

The above can be understood as a "coordinating property" of the PEVs' update law, and will be employed as follows. Since p is convex and increasing, we have that $p'(\overline{L}) \geq p'(L_t^k)$, and $p'(\overline{L})$ is the Lipschitz constant of p for the interval $[0, \overline{L}]$. Then, we have that:

$$p(L_1) - p(L_2) \leq p'(\overline{L})(L_1 - L_2),$$

with $L_1 \leq L_2$. Since ψ_i is increasing and Lipschitz we have:

$$\begin{aligned}\psi(p(L_1) - p(L_2)) &\leq \psi_i(p'(\overline{L})(L_1 - L_2)) \\ &\leq r_\psi^i p'(\overline{L})(L_1 - L_2).\end{aligned}$$

Using Assumption 2.2, with $L_2 = L_{\min}$, we obtain:

$$\psi_i(p(L_t^k) - p(L_{\min}^k)) < \frac{T}{(T-1)x_{\max}}(L_t^k - L_{\min}^k), \tag{2.15}$$

for all i, t, k.

The following result establishes the convergence of any trajectory of the Price Leveling algorithm towards the set of optimal solutions of the PEV charging problem.

Theorem 2.1 (Convergence)**.** *The* Price Leveling *algorithm defined by equation* (2.14) *converges to the set of optimal solutions of the PEV charging problem.*

Proof. Let us define functions V_l, for $l \in \{1, \ldots, m\}$:

$$V_l(y) = \min_{q \in \Theta_l}(D_q + x^\star_q) - \min_{t \in \Theta_l}(D_t + x_t), \tag{2.16}$$

$$\Theta_l = \cup_{r=l}^{m+1} \Upsilon_r. \tag{2.17}$$

Recall that $x_t = \sum_{i \in I} u_{i,t}$ and $u_{i,t} = \sum_{\ell=1}^{m_i} y_{i,t,\ell}$. Let us consider the sets $\mathcal{E}_l$, for $l \in \{1, \ldots, m\}$, such that $\mathcal{E}_l = \{y \in \overline{\mathcal{F}} \mid x_t = x^\star_t, \forall t \in \bigcup_{r=1}^{l} \Upsilon_r \text{ and } D_q + x_q \geq D_t + x_t, \forall q \in \Theta_l, \forall t \in \Upsilon_l\}$. Intuitively, this set contains all those states y of the algorithm that generate charging strategies u that determine optimal values for each $t \in \bigcup_{r=1}^{l} \Upsilon_r$ and such that for each time slot $q \in \Theta_l \setminus \Upsilon_l$ the price is not less than that of the slots in Υ_l. Define also $\mathcal{E}_0 = \overline{\mathcal{F}}$. Note that by this definition, $\mathcal{E}_m \subset \mathcal{E}_{m-1} \subset \ldots \subset \mathcal{E}_1 \subset \mathcal{E}_0$, and $\mathcal{E}_m$ is the set of optimizers of the PEV charging problem. We aim to employ the invariance theory introduced in Section 2.5 with height functions V_l, manifolds $\mathcal{E}_{l-1}$ and submanifolds $\mathcal{E}_l$, in a nested way, for $l \in \{1, \ldots, m\}$ to conclude our result.

This is carried out by performing the following steps: first we show that the functions V_l are strictly positive in $\mathcal{E}_{l-1} \setminus \mathcal{E}_l$, and zero in $\mathcal{E}_l$, for $l \in \{1, \ldots, m\}$. Next, we show that each function V_l is strictly decreasing along the dynamics $y^{k+1} = g(y^k)$ that represent the Price Leveling algorithm, for any $y \in \mathcal{E}_{l-1} \setminus \mathcal{E}_l$, and $V_l(g(y)) - V(y) = 0$ if $y \in \mathcal{E}_l$, for $l \in \{1, \ldots, m\}$. Then, each function V_l allows us to narrow down the set in which the omega-limit set of a solution of $y^{k+1} = g(y^k)$ into $\mathcal{E}_l$, via the invariance theory in Section 2.5. After m steps of this procedure, we obtain that the omega-limit set of an arbitrary solution of the Price Leveling algorithm converges to the set $\mathcal{E}_m$, which corresponds to the set of optimal solutions of the PEV charging problem. □

2.3 Hierarchical PEV control under communication failures

Thus far, we have only considered the Price Leveling algorithm working under the assumption that for each iteration $k \in \mathbb{N}$, every PEV $i \in I$ sends information to the utility, so that the utility can accurately compute the price p_t^k for all $t \in \tau$. However, if communication between the utility and PEVs presents failures, we must evaluate how it affects the evolution of the system state.

In order to capture the communication failures of the system we define a sequence of sets $\{I_k\}_{k=0}^{\infty}$ such that $I_k \subseteq I$ and it holds that at the k^{th} iteration only the PEVs in I_k receive price information from the utility. Likewise, only those PEVs in I_k send the charging strategy u_i^{k+1} that is computed using p_t^k back to the utility. This characterization allows for two different types of interactions between PEVs and the utility. On the one hand, we can consider an anonymous setting in which the utility does not know the identity of each PEV. On the other hand, we can assume that the utility knows for which PEVs the strategy did not arrive at each k. Then, the utility can reconstruct information on the electricity price at each time.

2.3.1 Non-Anonymous Price Leveling algorithm under communication failures

Let us assume that every PEV $i \in I$ executes the Price Leveling algorithm as has been introduced in Section 2.2. However, for any iteration k such that $i \notin I_k$, p_t^k, $t \in \tau$, is not available to i. Then, since p_t^k, $t \in \tau$, is required to compute u_i^{k+1}, for the next step agent i sets $u_i^{k+1} = u_i^k$. For each $i \in I$, $k \in \mathbb{N}$, define $k_{\text{upd}}(i,k) = \max\{s \in \{0,\dots,k\} \mid i \in I_s\}$. This leads to the Non-Anonymous Price Leveling algorithm:

$$y_{i,t,\ell}^{k+1} = \begin{cases} f_{i,t,\ell}(y^k), & \text{if } i \in I_k \text{ and } t \in \mathcal{Z}_i^\ell \\ 0, & \text{if } i \in I_k \text{ and } t \notin \mathcal{Z}_i^\ell \\ y_{i,t,\ell}^k, & \text{otherwise,} \end{cases} \tag{2.18}$$

where $f_{i,t,\ell}$ is given by equation (2.14), with a difference in p_t^k, which will be computed using the last available information to the utility, i.e., the charging strategy of i at the iteration $k_{\text{upd}}(i,k)$. Then, $p_t^k = p(D_t + \sum_{i\in I} u_{i,t}^{k_{\text{upd}}(i,k)})$, for each $t \in \tau$. Nonetheless, we can see that given the dynamics in (2.18), for all $i \in I$, $u_{i,t}^{k_{\text{upd}}(i,k)} = u_{i,t}^k$, then the actual price at the k^{th} iteration can be reconstructed by the aggregator.

Now, we study convergence of the Non-Anonymous Price Leveling algorithm under communication failures. Clearly, with this modification, we can express our system as:

$$y^{k+1} = g(y^k, I_k),$$

where $g(y^k, I_k)$ is an execution of the Non-Anonymous Price Leveling algorithm subject to communication failures $I_k \subseteq I$. In the following we provide conditions for convergence.

Theorem 2.2 (Convergence)**.** *The* Non-Anonymous Price Leveling *algorithm in Equation* (2.18) *converges to the set of optimal solutions of the PEV charging problem with communication failures $I_k \subseteq I$, $k \in \mathbb{N}$ if there is a number $n \in \mathbb{N}$ such that for all $k \in \mathbb{N}$ it holds that $\bigcup_{\ell=0}^{n} I_{k+\ell} = I$.*

Proof. We define set valued map $\mathcal{R} : \mathbb{R}^d_{\geq 0} \rightrightarrows \mathbb{R}^d_{\geq 0}$ such that $\mathcal{R}(\xi) = \{z \in \mathbb{R}^d_{\geq 0} \mid \exists \{z_\ell\}_{\ell=0}^n \subset \mathbb{R}^d_{\geq 0}, \{I_\ell\}_{\ell=0}^{n-1}, I_\ell \subseteq I \text{ s.t. } z_0 = \xi, z_\ell = z, z_{\ell+1} = g(z_\ell, I_\ell), \forall \ell \in \{1,\dots,N\}, \bigcup_{\ell=0}^{n-1} I_\ell = I\}$. Further, let us define the difference inclusion $y^{k+1} \in \mathcal{R}(y^k)$. Clearly, each solution of this system is a subsequence of a solution of $y^{k+1} = g(y^k, I_k)$, for any possible sequence of communication failures such that given a number $n \in \mathbb{N}$, all PEVs communicate with the utility at least once every n iterations. Notice that for each ξ in the domain of $\mathcal{R}$, the set-valued map returns less than 2^{Nn} points in $\mathbb{R}^d$. Since $\mathcal{R}(\xi)$ is finite, then it is compact for all ξ. The upper semicontinuity of $\mathcal{R}$ follows easily from the fact that each element of $\mathcal{R}$ comes from the composition of n continuous maps $g(\cdots g(g(\xi, I_k), I_{k+1}), \cdots, I_{k+n})$, since this composition is also continuous in ξ. Then, using the fact that $\bigcup_{\ell=1}^{n} I_{k+\ell} = I$, and similarly to Part 2 in the proof of Theorem 2.1, it follows that if $y^k \in \mathcal{E}_{l-1}$, then for each $\zeta \in [V_l \circ \mathcal{R} - V_l](y^k)$, it holds that $\zeta \leq 0$. Moreover, $0 \in [V_l \circ \mathcal{R} - V](y^k)$ only if $y^k \in \mathcal{E}_l$, for all $l \in \{1,\dots,m\}$.

Finally, we use Theorem 2.3 exactly as it has been used in Part 3 of the proof of Theorem 2.1, to conclude convergence of each solution of the Non-Anonymous Price Leveling algorithm with communication failures towards the set of optimal charging strategies. □

Remark 2.4. *Notice that the communication failures defined at the beginning of this section are symmetric, i.e., the communication link fails at iteration k in both directions. However, it is not always the case. If the link fails only in the communication from the aggregator to the* i^{th} *PEV, but the PEV can successfully transmit its charging profile at iteration k, the algorithm still converges. The analysis in Theorem 2.2 remains valid. This is because the PEV will simply not modify its charging profile. On the other hand, it may happen that the* i^{th} *PEV successfully receives the pricing information from the aggregator at iteration k, but the aggregator does not receive the new charging profile from i. In this case, a standard communication protocol with reception confirmation fixes the problem. As soon as the* i^{th} *PEV learns that the aggregator did not receive its new charging profile* u_i^{k+1}*, it simply goes back to its previous algorithm state* y_i^k*.*

2.3.2 Anonymous Price Leveling algorithm under communication failures

The Anonymous Price Leveling algorithm should follow the dynamics in (2.18), with the difference that the feedback control signal p_t^k must be computed with only the information available at iteration k. A first approach would be $p_t^k = p(D_t + \sum_{i \in I_k} u_{i,t}^k)$, for all $t \in \tau$. Notice that the aggregator does not know the set I_k, then it cannot identify which PEV's information was obtained at iteration k.

Simulations show that executions of the Anonymous Price Leveling algorithm with communication failures, lead to a neighborhood of the optimal cost. The size of this neighborhood depends on the reliability of the communication network.

The reason why the Anonymous Price Leveling dynamics does not converge to the optimal solution of the PEV charging problem is the following. Let us assume that for some k, y^k is such that $u^k = u^\star$ for some $u^\star$ optimal. In order for the algorithm to converge to $u^\star$, it should happen that y^k is a fixed point of the algorithm under any

iteration, i.e., $y^k = g(y^k, I_k)$, for any $I_k \subset I$. However, if $I_k \subsetneq I$, there is a q such that $p_q^k = p(D_q + \sum_{i \in I_k} u_{i,q}^k) < p(D_t + x_t^\star)$, for some $t \notin \operatorname{argmin}_{s \in \tau} p(D_s + x_s^\star)$. Then, an iteration of the Anonymous Price Leveling algorithm aims to make u^{k+1} optimal with respect to $J_{I_k}(u) = \sum_{t \in \tau} p(D_t + \sum_{i \in I_k} u_{i,t})(D_t + x_t)$. Since in general u^k is not optimal with respect to J_{I_k}, then u^{k+1} will not coincide with u^k.

2.4 Simulations

The first part of this section, we present the results of a simulation carried out using the Non-Anonymous Price Leveling algorithm. In the second part we use the Anonymous Price Leveling algorithm with a different scenario.

Part 1: In order to demonstrate the Non-Anonymous Price Leveling algorithm performance, we carry out simulations for a 24 hour scenario, starting at 12:00 pm and ending at the same time the next day. This time window is divided in 48 time slots with equal duration, that is, each slot is half-an-hour long. We use a non-PEV forecasted demand profile that follows a realistic behavior, but with some scaling factor, and a population of 20 PEVs. This population is divided into four groups according to their forecasted usage schedules: the first group contains four PEVs that will be used from 4 : 30 pm to 7 : 30 pm, requiring during that period a normalized amount of energy between 0.6 and 0.9, depending on each $i \in I$. They are also required to have a normalized amount of energy between 0.25 and 1 by the end of the day, depending on each $i \in I$. The second group contains six PEVs that will be used from 2 : 00 pm to 4 : 00 pm, requiring the battery to be fully charged for that period. They are also required to have a normalized amount of energy of 0.8 by the end of the day. The third group only has two PEVs which will have two usage periods: the first goes from 12 : 00 pm to 1 : 00 pm with 0.3 normalized energy consumption, and a second usage period between 4 : 30 am and 6 : 00 am with energy requirement of 0.8 and 0.9 respectively. They are required to have a normalized amount of energy of 0.8 and 1 respectively, by the end of the day. The last group has eight PEVs which will be used between 4 : 30 am and 7 : 00 am. Their normalized consumption during that time is $5/9$, and by the end of the day they are expected to have a normalized energy of 0.8.

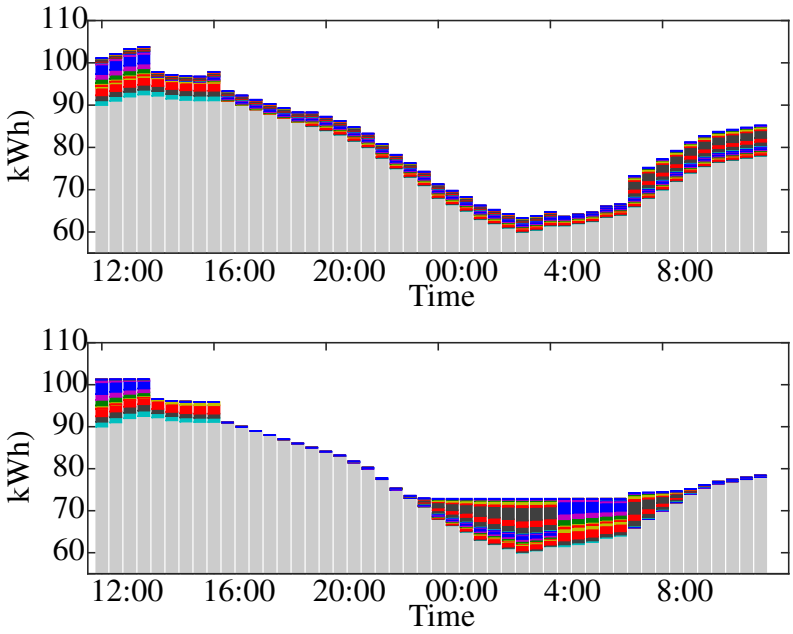

Figure 2.4: Top figure: initial estimated demand profile. Bottom: optimal estimated demand profile. The non-PEV demand is shown in light grey. Other colors show each PEV's charging profile.

Each PEV has a different battery capacity, and a different charging rate bound. For this particular example, PEVs 11 and 12 have been given a charging rate bound such that it is not possible to obtain enough energy for the requirement at the end of the day by only charging after the second usage lapse.

We consider a function p given by $p(x) = x^2$, $x \geq 0$. The functions ψ_i are linear functions of the form $\psi_i(x) = r^i_\psi x$, with $r^i_\psi = r_\psi$ for all $i \in I$. Even though our theoretical result presents an upper bound on the Lipschitz constant for the functions ψ_i, we use constants larger than the bound, which still leads to convergence in a smaller number of iterations. The number of iterations may vary depending on the initial conditions and the parameters D, $\{W_i\}$, $\{w_i\}$ for all $i \in I$.

In our scenario, failures are modeled as IID random variables, which represent the active links at iteration k. This means that for each iteration k, each link has a failure probability $1 - P$. In Figure 2.4, the non-PEV demand profile is shown in in light grey color, while the PEV demand is shown in dark colors. The plot on the top corresponds to the initial estimated charging strategies for all PEVs, along with the estimated non-PEV demand profile. More specifically, these initial charging strategies

satisfy the forecasted usage, while the estimated non-PEV demand profile is the one that is used for the computation of the optimal charging strategy for each PEV. The bottom plot shows the optimal estimated demand profile after the execution of the Non-Anonymous Price Leveling algorithm. It can be seen that the solution provided by the Non-Anonymous Price Leveling algorithm is converging to a strategy that levels prices as much as possible, given the problem constraints. The load profile we show corresponds to the truncation of the algorithm execution at iteration 150. This is a valley-filling-like solution for the charging problem, while respecting their usage constraints, along with the physical battery constraints. We have also verified in this simulation that the solution provided by the algorithm does not vary along the execution. This means that the algorithm converges to a specific optimal charging profile as opposed to having agents constantly switching between optimal charging profiles. We emphasize this fact, since we were not able to prove convergence of the algorithm to a point in the invariant set but just asymptotic convergence to the set of optimal solutions. This is a consequence of applying a LaSalle-like type of invariance result like ours. Nevertheless, previously proposed algorithms are not able to guarantee this property either, because standard LaSalle results are used there as well. Figure 2.5 shows how the overall energy cost decreases along the algorithm execution, for two different values of P, namely 0.3 and 0.8. As it is expected, the cost decreases faster when the communication links are more reliable ($P = 0.8$).

In Figure 2.6, we show the demand profile for an implementation of the optimal charging strategies provided by the Non-Anonymous Price Leveling algorithm on a realistic scenario. To this end, i) we take the estimated non-PEV demand profile and we corrupt it with additive white Gaussian noise to simulate the actual non-PEV demand profile, and ii) we randomly introduce mild modifications on the estimated usage schedule for the PEVs in order to simulate the actual usage schedules. The modifications simply correspond to extending or shortening the usage intervals W_i^ℓ by one or two time slots, and randomly generating additional usage requirements in case the usage interval is extended. Notice that in the actual scenario, usage schedules might not be satisfied by the computed charging strategies, therefore we assume that each PEV has some ancillary energy supply (e.g., fuel as in the case of a hybrid vehicle) that can take care of the

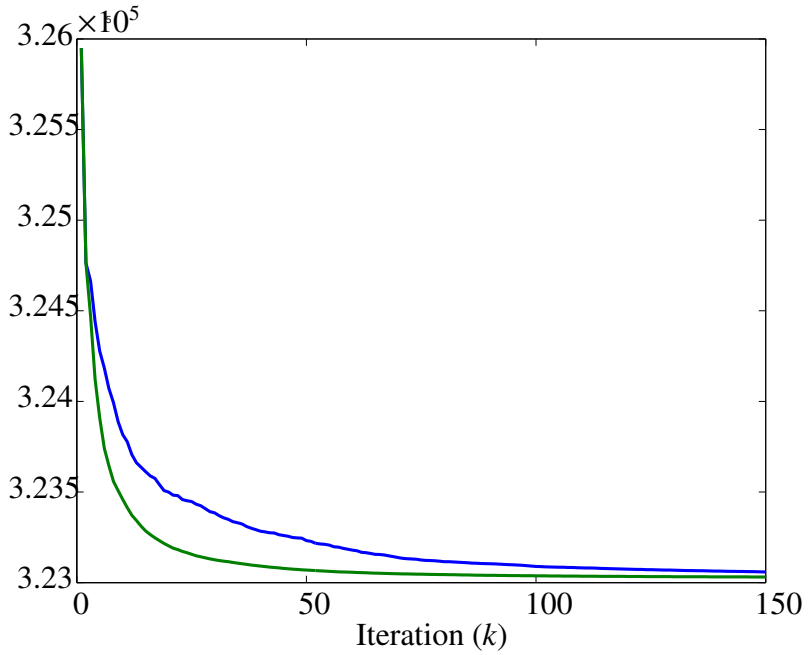

Figure 2.5: Evolution of the estimated cost function for the NON-ANONYMOUS PRICE LEVELING with $P = 0.3$, dashed line and $P = 0.8$, solid line.

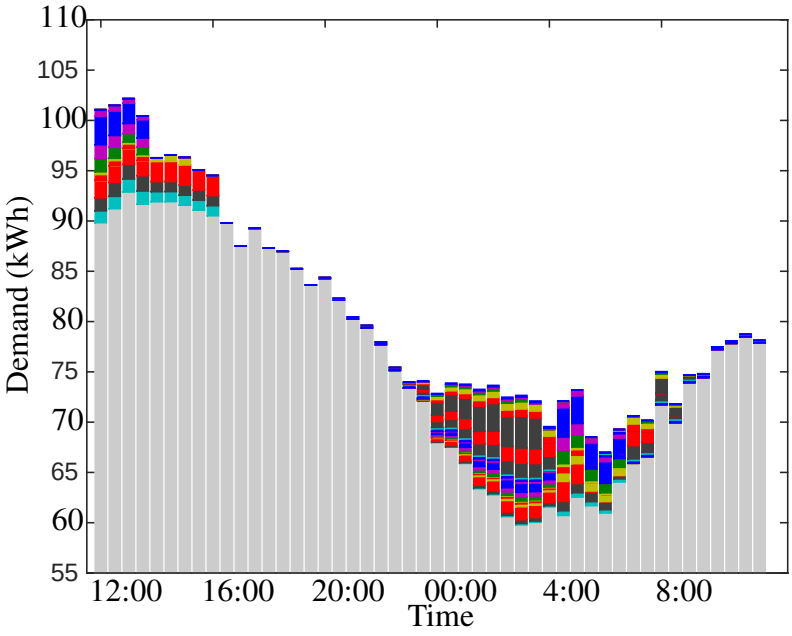

Figure 2.6: Actual demand profile using the optimal charging strategy. The non-PEV demand is shown in light grey. Other colors show each PEV's charging profile.

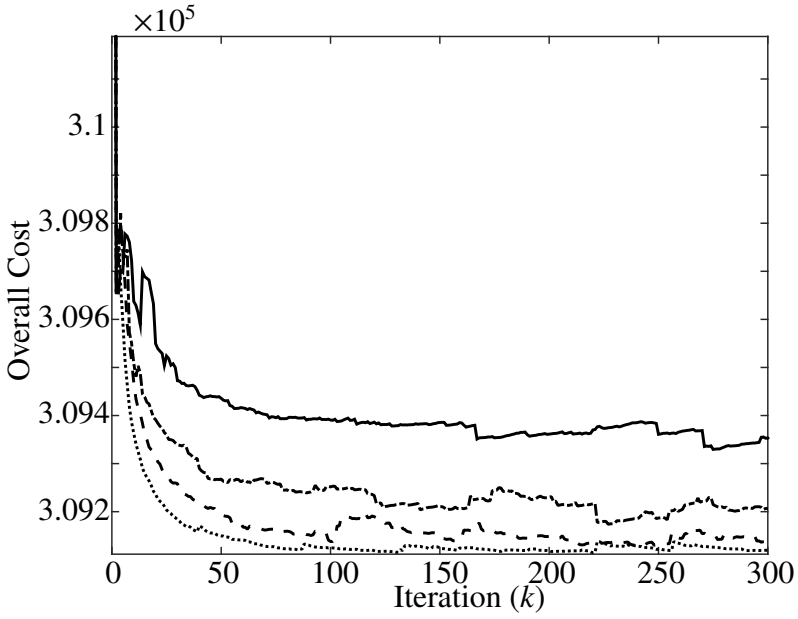

Figure 2.7: Overall cost of the consumed energy (PEV and non-PEV) for: $P = 0.2$ solid line, $P = 0.4$ dashdot line, $P = 0.6$ dashed line, $P = 0.8$ dotted line.

extra need.

It can be seen that even though the demand profile does not have a valley-filling-like shape due to the uncertainty, the PEV demand does tend to relocate in times when the non-PEV demand is lower.

Part 2: In order to show the behavior of the Anonymous Price Leveling algorithm, we define a simpler scenario, with 20 PEVs, the same values for D, but generating two groups of PEVs: the first group, given by PEVs from 1 to 6, has $W_i = \emptyset$, and $w_{i,T+1} = 1$ for all $i \in \{1, \ldots, 6\}$. The second group, i.e., PEVs from 7 to 20 are such that have a unique usage lapse from $2:00$ am to $4:30$ am. For each time slot in this period, $w_{i,t} = 0$, for all $i \in \{7, \ldots, 20\}$. It means that PEVs do not charge, but they do not consume energy on that time. It also holds for all $i \in \{7, \ldots, 20\}$, that $w_{i,T+1} = 1$. In Figure 2.7, we show how the cost evolves for different values of the link reliability parameter P. It can be seen that the higher P is, the lower average cost can be achieved.

2.5 Auxiliary section: invariance theory

Consider a discrete-time dynamical system given by the difference inclusion

$$z^{k+1} \in F(z^k), \qquad k \geq 0, \tag{2.19}$$

where the state z^k belongs to a compact manifold $\mathcal{M}$ of $\mathbb{R}^n$, and $F : \mathcal{M} \to \mathcal{M}$ is an upper-semicontinuous, closed set-valued map with non-empty values. Since $\mathcal{M}$ is compact, $F(z)$ is compact for all $z \in \mathcal{M}$ by definition. We denote by $\phi(k, z^0)$, $k \geq 0$, a solution starting from the initial condition $z^0 \in \mathcal{M}$. Note that any solution of (2.19) will be bounded, hence compact.

Definition 2.1 (Limit point, Omega-limit set)**.** *Consider a solution of* (2.19)*, $\phi(\cdot, z^0)$, with initial condition z^0. A point p is said to be a limit point of ϕ if there exists a sequence $\{k_j\}_{j=0}^{\infty}$, with $k_j \to \infty$ as $j \to \infty$, such that $\lim_{j\to\infty} \phi(k_j, x^0) = p$. The omega-limit set of ϕ denoted as $\Omega(\phi)$ is the set of all limit points of ϕ.*

Since $\mathcal{M}$ is compact and $\phi \subset \mathcal{M}$, the omega-limit set of ϕ is nonempty, and closed.

Assumption 2.3 (Height function on $\mathcal{S}$)**.** *Assume that:*

- $\Omega(\phi)$ *is contained in a submanifold $\mathcal{S} \subseteq \mathcal{M}$.*
- *There exists a compact neighborhood K of $\Omega(\phi)$ in $\mathcal{M}$, such that $O = \text{int}(K)$ is an open neighborhood of $\Omega(\phi)$.*
- *There is a continuous function $W : K \to \mathbb{R}$ such that $\zeta \leq 0$ for all $\zeta \in [W \circ F - W](z)$, for all $z \in \mathcal{S} \cap O$. Let E be the set defined as $E = \{z \in \mathcal{S} \cap O \mid 0 \in [W \circ F - W](z)\}$. Then $\zeta < 0$, for all $\zeta \in [W \circ F - W](z)$, $z \in (\mathcal{S} \cap O) \setminus E$.*

Assume that the continuous function W satisfies $\zeta \leq 0$ for all $\zeta \in [W \circ F - W](z)$ for all $z \in \mathcal{S} + B_\epsilon(0)$, $\epsilon > 0$. Since $\lim_{k\to\infty} \text{dist}(\phi(k, z^0), \Omega(\phi)) = 0$, it follows that there is some $k_1 < \infty$ such that $\phi(k, z^0) \in \mathcal{S} + B_\epsilon(0)$. Then, it is easy to conclude that ϕ converges to the largest weakly positively invariant set contained in $\{z \in \mathcal{S} + B_\epsilon(0) \mid 0 \in [W \circ F - W](z)\}$. However, if there are $\zeta \in [W \circ F - W](z)$ for some $z \in (\mathcal{S} + B_\epsilon(0)) \setminus \mathcal{S}$,

such that $\zeta > 0$, it is not possible to conclude a similar result, since the set $\mathcal{S}$ can be reached only in infinite time. The following results circumvent this problem, by using properties of omega-limit sets, and the continuity of W.

Lemma 2.7. *Let C be a compact set. Define $G : C \rightrightarrows \mathbb{R}$, as a closed, bounded, upper-semicontinuous set-valued map. If $G(z) < 0$ for all $z \in C$, then, there is a $\delta < 0$ such that for all $z \in C$, it holds that $\xi \leq \delta < 0$, for all $\xi \in G(z)$.*

Proof. Let us proceed by contradiction. Assume that there is no such δ. Then, there is a sequence $z_k \to z$, $z_k \in C$, with $\xi_k \to \xi$ such that $\xi_k \in G(z_k)$ and $\lim_k \xi_k = 0$. Since G is closed and upper-semicontinuous, it holds that $\xi \in G(z)$. Since C is compact, we have that $z \in C$, thus, $\xi < 0$ for all $\xi \in G(z)$, a contradiction. □

The following results are inspired by [23]. The invariance Lemma 2.3 admits an immediate generalization to systems which do not evolve in compact manifolds, but simply have trajectories that remain bounded. Another generalization of Lemma 2.3, is E being contained in a countable number of level sets of W with no accumulation points as in [23]. Nonetheless, we state a version that is sufficient to prove our main result.

Lemma 2.8. *Let Assumption 2.3, on the existence of a height function on $\mathcal{S}$ containing $\Omega(\phi)$, hold. Then, it must be that $\Omega(\phi) \cap E \neq \emptyset$.*

Proof. Assume that $\Omega(\phi) \cap E = \emptyset$ to reach a contradiction. By the compactness of $\Omega(\phi)$, it holds that the continuous function W attains its minimum in $\Omega(\phi)$. Since $\Omega(\phi) \cap E = \emptyset$, by Lemma 2.7 it also holds that $W(\zeta) - W(z) \leq \delta < 0$, $\zeta \in F(z)$, for each $z \in \Omega(\phi)$.

Since $\Omega(\phi)$ is weakly positively invariant under the dynamics in (2.19), then for each $z^0 \in \Omega(\phi)$, there is a solution φ starting at z^0 that remains in $\Omega(\phi)$ for all k. Consider a solution $\varphi(k, z^0)$, with $z^0 \in \Omega(\phi)$. We have $W(\varphi(K, z^0)) - W(z^0) = \sum_{k=0}^{K-1} \left(W(\varphi(k+1, z^0)) - W(\varphi(k, z^0)) \right) \leq \sum_{k=0}^{K-1} \delta$. This implies that $\lim_{K \to \infty} W(\varphi(K, z^0)) = -\infty$, which contradicts that W attains a minimum in $\Omega(\phi)$. Then it must be that $\Omega(\phi) \cap E \neq \emptyset$. □

Theorem 2.3 (Invariance Result)**.** *Let Assumption 2.3, on the existence of a height function on $\mathcal{S}$ containing $\Omega(\phi)$ hold. If E is contained in a single level set of W, then $\Omega(\phi)$ is contained in E.*

Proof. Suppose first that $\Omega(\phi)$ is contained in a single level set of W, say L. Since E lies in a single level set of W, as we have that $\Omega(\phi) \cap E \neq \emptyset$, by Lemma 2.8, then $E \subseteq L$. Given that $\Omega(\phi)$ is weakly positively invariant, there exists a solution φ starting at any point $z^0 \in \Omega(\phi) \setminus E$, such that $\varphi(k, z^0) \in \Omega(\phi) \subseteq L$ for all k. But if $z^0 \notin E$, then $W(\varphi(1, z^0)) - W(\varphi(0, z^0)) < 0$, which means that $W(\varphi(1, z^0)) < W(\varphi(0, z^0))$, therefore $\varphi(1, z^0) \notin L$, a contradiction. Then, it must be that $z^0 \in E$ for all $z^0 \in \Omega(\phi)$.

We proceed again by contradiction, assuming that $\Omega(\phi)$ is not contained in a single level set of W. Given that W is continuous and $\Omega(\phi)$ is compact, then $W(\Omega(\phi)) \subseteq [\min, \max]$ where $[\min, \max]$ is an interval of the real line. Define $\omega_E = \Omega(\phi) \cap E \neq \emptyset$, and let w be the value of W on ω_E.

Consider $\Omega(\phi) + B_\epsilon(0)$. Given that $\Omega(\phi) \subset \Omega(\phi) + B_\epsilon(0)$, and the solution ϕ converges asymptotically to its omega limit set, which is the smallest closed attracting set of ϕ, there exists some time k_0 such that $\phi(k, z^0) \in \Omega(\phi) + B_\epsilon(0)$ for all $k \geq k_0$, and for all $\epsilon > 0$. Let us consider only those ϵ such that $\Omega(\phi) + B_\epsilon(0) \subset O$. Recall that $O = \text{int}(K)$ is an open neighborhood of $\Omega(\phi)$. Let us define $\mathcal{U}_1$ as an open neighborhood of E in $\mathcal{M}$, and for a given ϵ, $\mathcal{U}_1(\epsilon) = \mathcal{U}_1 \cap (\Omega(\phi) + B_\epsilon(0))$. Denote $\underline{b}_1 = \inf_{z \in \mathcal{U}_1(\epsilon)} W(z)$, $\overline{b}_1 = \sup_{z \in \mathcal{U}_1(\epsilon)} W(z)$. Since $E \cap \Omega(\phi)$ is closed, we can choose some point $p \in \Omega(\phi) \setminus E$ (which exists by the assumption that $\Omega(\phi)$ is not contained in a single level set), with a neighborhood $\mathcal{U}_p \subset O$. Define, for a given ϵ, $\mathcal{U}_p(\epsilon) = \mathcal{U}_p \cap B_\epsilon(\Omega(\phi))$ and consider $\underline{b}_p = \inf_{z \in \mathcal{U}_p(\epsilon)} W(z)$, $\overline{b}_p = \sup_{z \in \mathcal{U}_p(\epsilon)} W(z)$. There are two possible cases: i) $W(p) < w$ or ii) $W(p) > w$. Let us consider the first case. Recall that $\omega_E \subset \mathcal{U}_1$ is the only subset of $S \cap O$ for which there exists some $\zeta \in F(z)$ such that $W(\zeta) - W(z) = 0$, and for all $z \in (S \cap O) \setminus E$, it holds that $W(\zeta) - W(z) < 0$, for all $\zeta \in F(z)$. By continuity of W and given that F is upper-semicontinuous, $W \circ F - W$ is upper-semicontinuous [24, Proposition 1.4.14]. Since $F(z)$ is compact for all z, then $[W \circ F - W](z)$ is also compact for all z.

Using this fact, along with Lemma 2.7, there is a $\delta_1 < 0$ such that for all $z \in (S \cap O) \setminus \mathcal{U}_1$, it holds that $W(\zeta) - W(z) \leq \delta_1 < 0$, for all $\zeta \in F(z)$.

Next, by upper-semicontinuity of $W \circ F - W$, for any δ_2, there is ϵ such that for every $y \in B_\epsilon(z)$, it holds that $[W \circ F - W](y) \subseteq [W \circ F - W](z) + B_{\delta_2}(0)$. Let us fix some $\delta_2 \in (\delta_1, 0)$. Then, we can choose ϵ such that:

$$\max[W \circ F - W](y) \leq \max B_{|\delta_2|}([W \circ F - W](z)) \leq \delta_1 + |\delta_2|, \tag{2.20}$$

for all $y \in B_\epsilon(\Omega(\phi) \setminus \mathcal{U}_1)$. It immediately implies that:

$$\sup_{z \in B_\epsilon(\Omega(\phi)) \setminus \mathcal{U}_1(\epsilon)} \left[\max_{\xi \in F(z)} (W(\xi) - W(z)) \right] \leq \delta_1 - \delta_2 = \delta < 0. \tag{2.21}$$

Therefore, if the set $K = \{z \in B_\epsilon(\Omega(\phi)) \mid \exists a \in [W \circ F - W](z), a > 0\}$ is nonempty, it must be that $K \subset \mathcal{U}_1(\epsilon)$. Let us fix $\mathcal{U}_1$, $\mathcal{U}_p$, satisfying $\underline{b}_1 > \overline{b}_p$. Since p is a limit point of ϕ, there is a k_1 such that $\phi(k_1, z^0) \in \mathcal{U}_p(\epsilon)$. Pick a point $q \in E \cap \Omega(\phi) \neq \emptyset$. Then, as q is a limit point of ϕ, there is a first $k_2 > k_1$ such that $\text{dist}(\phi(k_2, z^0), q) < \epsilon$ and $\phi(k_2, z^0) \in \mathcal{U}_1(\epsilon)$. However, due to (2.21) $W(\phi(k_2, z^0)) - W(\phi(k_1, z^0)) \leq \sum_{l=k_1}^{k_2} \delta < 0$, which means that $W(\phi(k_2, z^0)) \leq \overline{b}_p < \underline{b}_1$, a contradiction with the fact that $\phi(k_2, z^0) \in \mathcal{U}_1(\epsilon)$.

For the second case, i.e., $W(p) > w$ let us choose $\mathcal{U}_1$, $\mathcal{U}_p$ and ϵ as follows: $\mathcal{U}_1$ small enough so that $\max[W \circ F - W](v) < \epsilon_2$, $\epsilon_2 > 0$, for all $v \in \mathcal{U}_1$ (it can be done by upper-semicontinuity of $W \circ f - W$, in a similar way as we reached the bound δ in (2.21)), $\mathcal{U}_1$, $\mathcal{U}_p$ such that $\underline{b}_p > \overline{b}_1 + \epsilon_2$. Next, let us choose ϵ small enough so that (2.21) holds. Since $\Omega(\phi) \cap E \neq \emptyset$, there is a k_1 such that $\phi(k_1, z^0) \in \mathcal{U}_1(\epsilon)$. However, we have that $W(\phi(k_1 + 1, z^0)) - W(\phi(k_1, z^0)) < \epsilon_2 < \underline{b}_p - \overline{b}_1$, and since $\phi(k_1, z^0) \in \mathcal{U}_1(\epsilon)$, it holds that $W(\phi(k_1, z^0)) \leq \overline{b}_1$, therefore $W(\phi(k_1 + 1, z^0)) < \underline{b}_p$, which implies that $\phi(k_1 + 1, z^0) \notin \mathcal{U}_p(\epsilon)$. Hence, there are two possibilities: either $\phi(k_1 + 1, z^0) \in \mathcal{U}_1(\epsilon)$, in which case we can repeat out previous analysis to conclude that $\phi(k_1 + 2, z^0) \notin \mathcal{U}_p$, or $\phi(k_1 + 1, z^0) \in B_\epsilon(\Omega(\phi)) \setminus (\mathcal{U}_1(\epsilon) \cup \mathcal{U}_p(\epsilon)) \subseteq B_\epsilon(\Omega(\phi)) \setminus \mathcal{U}_1(\epsilon)$. Since $\mathcal{U}_p(\epsilon)$ contains a limit point of $\phi(k, z^0)$, there has to be a first $k_1 + n > k_1$ such that $\phi(k_1 + n, z^0) \in B_\epsilon(\Omega(\phi)) \setminus (\mathcal{U}_1(\epsilon) \cup \mathcal{U}_p(\epsilon))$ but $\phi(k_1 + n + 1, z^0) \in \mathcal{U}_p(\epsilon)$. In other words, by the above process we may construct a sequence $\{\phi(k_1 + l, z^0)\}_{l=1}^{n-1} \subseteq \mathcal{U}_1(\epsilon) \setminus \mathcal{U}_p(\epsilon)$, with $\phi(k_1 + n, z^0) \in B_\epsilon(\Omega(\phi)) \setminus (\mathcal{U}_1(\epsilon) \cup \mathcal{U}_p(\epsilon))$, but $\phi(k_1 + n + 1, z^0) \in \mathcal{U}_p(\epsilon)$. However, in this case we have that $W(\phi(k_1 + n + 1, z^0)) - W(\phi(k_1 + n, z^0)) \leq \delta_2 < 0$ by (2.21), which implies that W decreases. Therefore, $\phi(k_1 + n + 1, z^0) \notin \mathcal{U}_p(\epsilon)$ since the above implies that $W(\phi(k_1 + n + 1, z^0)) < W(\phi(k_1 + n, z^0)) < \overline{b}_1 < \underline{b}_p$, a contradiction. Then it must be that $\Omega(\phi)$ is in a single level set of W. Hence, we follow the proof of the first case in this lemma to show that $\Omega(\phi) \subseteq E$. □

2.6 Details of the proof of Theorem 2.1

Our proof is divided in three parts: in Part 1, we show that $V_l(y) \geq 0$ for all $y \in \mathcal{E}_{l-1}$, and in fact $V_l(y) = 0$ if and only if $y \in \mathcal{E}_l$, for $l \in \{1, \ldots, m\}$, with V_l and $\mathcal{E}_l$ introduced in the proof sketch of Theorem 2.1. In Part 2, we show that V_l is monotonically non-increasing along any solution of the system $y^{k+1} = g(y^k)$ given by the Price Leveling algorithm, starting at some point in $\mathcal{E}_{l-1}$, for all $l \in \{1, \ldots, m\}$. Finally, we conclude the convergence of the solutions towards the set of optimal charging strategies, using Theorem 2.3, which can be found in Section 2.5.

Part 1: We have as a direct consequence of Lemma 2.5 that there must exist $q \in \Upsilon_l$ such that $D_q + x_q^\star - D_q + x_q \geq 0$, for all $y \in \mathcal{E}_{l-1}$. Since $D_q + x_q^\star = \min_{t\in\Theta_l}(D_t + x_t^\star)$, and by definition of $\mathcal{E}_{l-1}$, $\min_{t\in\Theta_l}(D_t + x_t) \leq (D_q + x_q)$, it follows that $V_l(y) \geq 0$, for all $y \in \mathcal{E}_{l-1}$. If $y \in \mathcal{E}_{l-1} \setminus \mathcal{E}_l$, then, i) there is $q \in \Upsilon_l$ such that $\min_{t\in\Theta_l}(D_t + x_t^\star) > D_q + x_q$, then it follows that $V_l(y) > 0$, or ii) there exists some $q \in \Upsilon_l$ such that $\min_{t\in\Theta_l}(D_t + x_t^\star) < D_q + x_q$, then, from Lemma 2.5, there must exist $\bar{t} \in \Upsilon_l$ such that $\min_{t\in\Theta_l}(D_t + x_t^\star) > D_{\bar{t}} + x_{\bar{t}}$, hence, $V_l(y) > 0$.

Next, by definition of the set $\mathcal{E}_l$, if $y \in \mathcal{E}_l$, it holds that for $t \in \Upsilon_l \subset \Theta_l$, $D_t + x_t = \min_{q\in\Theta_l}(D_q + x_q)$, and $D_t + x_t = D_t + x_t^\star$, for all $t \in \bigcup_{r=1}^{l} \Upsilon_r$. Recall that by definition of the partition $\{\Upsilon_r\}_{r=1}^{m+1}$, if $t \in \Upsilon_l$, $D_t + x_t^\star \leq D_q + x_q^\star$ for all $q \in \Theta_l$, then $D_t + x_t^\star = \min_{q\in\Theta_l}(D_q + x_q^\star) = \min_{q\in\Theta_l}(D_q + x_q)$. Hence, if $y \in \mathcal{E}_l$, $V_l(y) = 0$.

Part 2: Next, let us show that each function V_l is monotonically non-increasing along the system solutions inside the set $\mathcal{E}_{l-1}$, and it is decreasing for all $y \in \mathcal{E}_{l-1} \setminus \mathcal{E}_l$. Assume that $y^k \in \mathcal{E}_{l-1}$. Consider the dynamics in (2.14). Then, we can write x_t^{k+1}, by summing in both sides over all $\ell \geq \mathbf{n}(i,t)$, $i \in \{j \in I \mid t \in Z_i^\ell\}$ as follows:

$$\begin{aligned} x_t^{k+1} =& x_t^k + \sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\mathcal{Z}_i^\ell} \Delta_{i,\ell}^k(q,t)\psi_i(\max\{p_q^k - p_t^k, 0\}) \\ & - \sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\mathcal{Z}_i^\ell} \Delta_{i,\ell}^k(t,q)\psi_i(\max\{p_t^k - p_q^k, 0\}), \end{aligned} \tag{2.22}$$

for all $t \in \tau$. Let us analyze what happens with any x_t^k, $t \in \Theta_l$, when $y^k \in \mathcal{E}_{l-1}$. Notice that by definition of $\mathcal{E}_{l-1}$, $u_{i,s}^k = u_{i,s}^\star$, for some $u^\star$ optimal, for all $s \in \bigcup_{r=1}^{l-1} \Upsilon_r$, for all $i \in I$. Also note that for a fixed $j \in I$, depending on W_j, w_j, we can have i) $\sum_{q\in W_j\cup\{T+1\}} w_{j,q} =$

$\frac{\alpha_j}{\beta_j}\sum_{s\in\bigcup_{r=1}^{l-1}\Upsilon_r} u^\star_{j,s} + \vartheta_{j,0}$, meaning that the PEV j optimally gets all the battery charge necessary for the usage schedule, during time slots in $\bigcup_{r=1}^{l-1}\Upsilon_r$, or ii) $\sum_{q\in W_j\cup\{T+1\}} w_{j,q} > \frac{\alpha_j}{\beta_j}\sum_{s\in\bigcup_{r=1}^{l-1}\Upsilon_r} u^\star_{j,s} + \vartheta_{j,0}$. In the first case, we have that for every $t\in\Theta_l$, $u^k_{j,t}=0$, therefore, the terms in the negative sum in (2.22) associated to j (i.e., $y^k_{j,t,\ell}$, for $\ell\in\{1,\dots,m_j\}$) are zero. If the second case holds, from Remark 2.3, in order for u^k to be feasible, there must exist some Υ_r, $r\geq l$ such that $u^k_{j,s}>0$ for some $s\in\Upsilon_r$. Then, we can fix $t\in\Theta_l$, for which $u^\star_{i,t}>0$. Let ℓ be such that $t\in Z^\ell_j$. Then, we have three different types of slots $q\in\bigcup_{r=1}^{l-1}\Upsilon_r$:

B1: First, for each $q\in\bigcup_{r=1}^{l-1}\Upsilon_r$ such that $q\in Z^n_j$, $n\in\{1,\dots,\ell-1\}$, by *C1* in Lemma 2.4 it must hold that either i) $u^k_{j,q}=u^\star_{j,q}=u_{j,\max}$, then $\gamma^k_1(j,q)=0$, or ii) the *maximum battery capacity constraint* is active some $r\in\{n,\dots,\ell-1\}$, i.e., there is $r\in\{n,\dots,\ell-1\}$ such that $\vartheta_{j,0}+\frac{\alpha_j}{\beta_j}\sum_{s\in\mathcal{Z}^r_j}u^k_{j,s}-\sum_{s\in\mathcal{W}^{r-1}_j}w_{j,s}=1$, then $\gamma^k_2(j,q,t)=0$. In any of these cases, the terms in the negative sum of (2.22) associated to $j\in I$ and $q\in\bigcup_{r=1}^{l-1}\Upsilon_r$ are zero.

B2: Now, for each $q\in\bigcup_{r=1}^{l-1}\Upsilon_r$ such that $q\in Z^n_j$, $n\in\{\ell+1,\dots,m_j\}$, it must hold by *C3* in Lemma 2.4 that either i) $u^\star_{j,q}=u_{i,\max}$, then, since $u^k_{i,q}=u^\star_{i,q}$ for some $u^\star$ optimal, it follows that $\gamma^k_1(j,q)=0$, or ii) the *user satisfaction constraint* is active for some $r\in\{\ell+1,\dots,n\}$, i.e., $\frac{\beta_j}{\alpha_j}\left(\sum_{q\in\mathcal{W}^{r-1}_j}w_{j,q}-\vartheta_{j,0}\right)=\sum_{q\in\mathcal{Z}^{r-1}_j}u^k_{j,q}$, where $u^k_{i,q}=u^\star_{i,q}$, for some $r\in\{\ell+1,\dots,n\}$. The second case implies that $\sum_{s\in\mathcal{W}^{r-1}_j}w_{j,s}=\frac{\alpha_j}{\beta_j}\sum_{s\in\mathcal{Z}^{r-1}_j}u^k_{j,s}+\vartheta_{j,0}$, for some $r\in\{\ell+1,\dots,n\}$. Recall that since $u^k_{j,s}=\sum_{\sigma=1}^{m_i}y^k_{j,s,\sigma}$, we have $\frac{\alpha_j}{\beta_j}\sum_{s\in\mathcal{Z}^{r-1}_j}\sum_{\sigma=1}^{m_i}y^k_{j,s,\sigma}+\vartheta_{j,0}=\sum_{s\in\mathcal{W}^{r-1}_i}w_{j,s}$. In addition, notice that summing on both sides of the expression in (2.14) over $\mathcal{Z}^\sigma_i$, for $i=j$ and any $\sigma\in\{1,\dots,m_j\}$, we obtain $\sum_{s\in\mathcal{Z}^\sigma_j}y^{k+1}_{j,s,\sigma}=\sum_{s\in\mathcal{Z}^\sigma_j}y^k_{j,s,\sigma}$, for all $k\in\mathbb{N}$. Then, given the selection of initial conditions of the Price Leveling algorithm, i.e., y^0 such that, $\frac{\alpha_j}{\beta_j}\sum_{s\in\mathcal{Z}^{r-1}_j}\sum_{\sigma=1}^{r-1}y^0_{j,s,\sigma}=\sum_{s\in\mathcal{W}^{r-1}_j}w_{j,s}-\vartheta_{j,0}$, we obtain $\frac{\alpha_j}{\beta_j}\sum_{s\in\mathcal{Z}^{r-1}_j}\sum_{\sigma=1}^{r-1}y^k_{j,s,\sigma}+\vartheta_{j,0}=\sum_{s\in\mathcal{W}^{r-1}_j}w_{j,s}$. Therefore, $\sum_{\sigma=1}^{r-1}\sum_{s\in\mathcal{Z}^{r-1}_j}y^k_{j,s,\sigma}=\sum_{s\in\mathcal{Z}^{r-1}_j}u^k_{j,s}$, hence it must be that $y_{j,t,\sigma}=0$ for $(t,\sigma)\in\mathcal{Z}^r_j\times\{r+1,\dots,m_j\}$. In both cases the terms in the negative sum of (2.22) related to j and q are zero.

B3: Finally, consider the case when $q\in Z^\ell_j$. By *C2* in Lemma 2.4, we have $u^\star_{i,q}=u_{i,\max}$, hence $\gamma^k_1(j,q)=0$.

It is easy to see that since $y^k \in \mathcal{E}_{l-1}$, $p(D_t + x_t^k) > p(D_q + x_q^k)$ for all $t \in \Theta_l$, $q \in \bigcup_{r=1}^{l-1} \Upsilon_r$. Therefore, in the equation (2.22) for any $t \in \Theta_l$, the terms in the positive sum associated to $q \in \bigcup_{r=1}^{l-1} \Upsilon_r$ are zero.

With this analysis, we have shown that for $y^k \in \mathcal{E}_{l-1}$, x_t^{k+1} for each $t \in \Theta_l$, depends uniquely on values of $y_{i,q,\ell}^k$, with $q \in \Theta_l$. Thus, we can write (2.22) as:

$$\begin{aligned} x_t^{k+1} = x_t^k &+ \sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\mathcal{Z}_i^\ell\cap\Theta_l} \Delta_{i,\ell}^k(q,t)\psi_i(\max\{p_q^k - p_t^k, 0\}) \\ &- \sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\mathcal{Z}_i^\ell\cap\Theta_l} \Delta_{i,\ell}^k(t,q)\psi_i(\max\{p_t^k - p_q^k, 0\}), \end{aligned} \tag{2.23}$$

for each $t \in \Theta_l$. Define $\psi(z) = \max_i(r_\psi^i)z$, that is, a linear function which is zero at zero, with slope equal to the maximum Lipschitz constant of a ψ_i over all $i \in I$. Next, let us write a lower bound for x_t^{k+1} by replacing the negative sum as follows:

$$\begin{aligned} &\sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\mathcal{Z}_i^\ell\cap\Theta_l} \Delta_{i,\ell}^k(q,t)\psi_i(\max\{p_t^k - p_q^k, 0\}) \\ &\qquad \le \sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\Theta_l} \Delta_{i,\ell}^k(q,t)\psi(\max\{p_t^k - p_q^k, 0\}) \\ &\qquad \le \sum_{q\in\Theta_l} \frac{x_t^k}{T}\psi(\max\{p_t^k - p_q^k, 0\}). \end{aligned}$$

In the first inequality we simply replaced the sum of $q \in \mathcal{Z}_i^\ell \cap \Theta_l$, such that $q \in \Theta_l$, by the sum on $q \in \Theta_l$, and we upper bounded the functions ψ_i by ψ. As $\mathcal{Z}_i^\ell \cap \Theta_l \subset \Theta_l$, the inequality holds. For the second inequality, we dropped the minimum operator that determines $\Delta_{i,\ell}^k(q,t)$, and replaced it by $y_{i,t,\ell}^k$. Then, we exchanged the summation order to obtain x_t^k. Therefore, we have:

$$\begin{aligned} x_t^{k+1} \ge x_t^k &+ \sum_{i\in I} \sum_{\ell=\mathbf{n}(i,t)}^{m_i} \sum_{q\in\mathcal{Z}_i^\ell\cap\Theta_l} \Delta_{i,\ell}^k(q,t)\psi_i(\max\{p_q^k - p_t^k, 0\}) \\ &- \textstyle\sum_{q\in\Theta_l} \frac{x_t^k}{T}\psi(\max\{p_t^k - p_q^k, 0\}). \end{aligned} \tag{2.24}$$

Define $H_l(x) = \{t \in \tau \mid t \in \operatorname{argmin}_{q\in\Theta_l}(D_q + x_q)\}$. Recall that $L_{\min}^k = \min_{t\in\Theta_l}(D_t + x_t^k)$. Since p is strictly increasing, $p(L_{\min}^k) = \min_{t\in\tau} p(D_t + x_t^k)$. Take (2.24) for $t = t_1$, $t_1 \in H_l(x^{k+1})$, and sum $D_{t_1} - L_{\min}^k$ on both sides. Then, we obtain:

$$L_{t_1}^{k+1} - L_{\min}^{k} \geq L_{t_1}^{k} - L_{\min}^{k} - \sum_{q \in \Theta_l} \frac{x_{t_1}^{k}}{T} \psi(\max\{p_{t_1}^{k} - p_{q}^{k}, 0\}) + \sum_{i \in I} \sum_{\ell = \mathbf{n}(i,t_1)}^{m_i} \sum_{q \in \mathcal{Z}_i^{\ell} \cap \Theta_l} \Delta_{i,\ell}^{k}(q,t) \psi_i(\max\{p_{q}^{k} - p_{t_1}^{k}, 0\}).$$

Using the inequality in (2.15):

$$L_{t_1}^{k+1} - L_{\min}^{k} \geq L_{t_1}^{k} - L_{\min}^{k} - \frac{x_{t_1}^{k}}{x_{\max}} (L_{t_1}^{k} - L_{\min}^{k}) + \sum_{i \in I} \sum_{\ell = \mathbf{n}(i,t_1)}^{m_i} \sum_{q \in \mathcal{Z}_i^{\ell} \cap \Theta_l} \Delta_{i,\ell}^{k}(q,t) \psi_i(\max\{p_{q}^{k} - p_{t_1}^{k}, 0\}).$$

This allows us to conclude that at any iteration $y^{k+1} = g(y^k)$ we have $V_l(y^{k+1}) - V_l(y^k) \leq 0$, due to $L_{t_1}^{k} - L_{\min}^{k} \geq 0$, for $y^k \in \mathcal{E}_{l-1}$. Moreover, for any $y^k \in \mathcal{E}_{l-1} \setminus \mathcal{E}_l$, by Lemma 2.5 it stands that $L_{\min}^{k} < L_q^{\star}$ for $q \in \Upsilon_l$. Thus, we have two possibilities:

A1: $H_l(x^k) \cap H_l(x^{k+1}) = \emptyset$. Then, it must be that $L_{t_1}^{k} - L_{\min}^{k} > 0$, and $V_l(y^{k+1}) - V_l(y^k) < 0$.

A2: $t_1 \in H_l(x^k) \cap H_l(x^{k+1})$. It implies that $L_{t_1}^{k} - L_{\min}^{k} = 0$. Since $L_{\min}^{k} < L_q^{\star}$, for all $q \in \Theta_l$, there must be some $i \in I$ with $\{t, t_1\} \subseteq \Theta_l \cap Z_i$, $t \notin H_l(x^k)$ such that $u_{i,t}^{k} > 0$, and a small change in u^k where $\bar{u}_{i,t_1}^{k} = u_{i,t_1}^{k} + \epsilon$ and $\bar{u}_{i,t}^{k} = u_{i,t}^{k} - \epsilon$, $\epsilon > 0$ small enough, while keeping all other components of u^k unchanged leads to a feasible solution of the system. When this happens, we obtain:

$$\sum_{\ell = \mathbf{n}(i,t_1)}^{m_i} \sum_{q \in \mathcal{Z}_i^{\ell} \cap \Theta_l} \Delta_{i,\ell}^{k}(q,t) \psi_i(\max\{p_{q}^{k} - p_{t_1}^{k}, 0\}) \geq \Delta_{i,\ell}^{k}(t,q) \psi_i(\max\{p_{t}^{k} - p_{t_1}^{k}, 0\}) > 0.$$

Then, we can conclude that $V_l(y^{k+1}) - V_l(y^k) < 0$, whenever $y^k \in \mathcal{E}_{l-1} \setminus \mathcal{E}_l$. This implies that for any $y^{k+1} = g(y^k)$, $V_l(y^{k+1}) - V_l(y^k) < 0$, whenever $y^k \in \mathcal{E}_{l-1} \setminus \mathcal{E}_l$.

Take $y^k \in \mathcal{E}_l$. From Part 1, we have that $V_l(y^k) = 0$. Because we have that $V_l(y) \geq 0$, and $V_l(g(y)) - V_l(y) \leq 0$ for all $y \in \mathcal{E}_{l-1}$, and given that $\mathcal{E}_l \subset \mathcal{E}_{l-1}$, it follows that $V_l(y^{k+1}) - V_l(y^k) \leq 0$, hence $V_l(y^{k+1}) \leq 0$, then, it must be that $V(y^{k+1}) = 0$, which means that $V_l(y^{k+1}) - V_l(y^k) = 0$.

Part 3: Consider compact manifolds $\mathcal{E}_l$, for $l \in \{0, \ldots, m\}$. By definition $\mathcal{E}_m \subset \mathcal{E}_{m-1} \subset \cdots \subset \mathcal{E}_1 \subset \mathcal{E}_0$. Next, define the dynamics $y^{k+1} \in G(y^k)$, where $G : \mathbb{R}_{\geq 0}^{d} \rightrightarrows \mathbb{R}_{\geq 0}^{d}$

is given by $G(y) = \{g(y)\}$. Notice that since g is Lipschitz continuous on the domain $\mathcal{F}$, then, it directly follows that G is upper-semicontinuous, with nonempty and compact values. Therefore, we can use Theorem 2.3 for $y^{k+1} \in G(y^k)$, with height function V_1, $\mathcal{M}_1 = \mathcal{S}_1 = \mathcal{E}_0$. Note that $\mathcal{E}_0$ is strongly invariant under $y^{k+1} \in G(y^k)$. Then, we have that for any solution ϕ of $y^{k+1} \in G(y^k)$, $\Omega(\phi) \subset \{y \in \mathcal{E}_0 \mid V_1(g(y)) - V_1(y) = 0\}$. By Part 2 of this proof, we have that $\{y \in \mathcal{E}_0 \mid V_1(g(y)) - V_1(y) = 0\} = \mathcal{E}_1$. Then, $\Omega(\phi) \subset \mathcal{E}_1$.

Further, for $l = 2$, we use V_2 as a height function for Lemma 2.3, $\mathcal{M}_2 = \mathcal{E}_0$, and $\mathcal{S}_2 = \mathcal{E}_1$. Since it has been shown in Part 2, that $V_l(g(y)) - V_l(y) = 0$, only if $y \in \mathcal{E}_l$, it immediately follows that $0 \in [V_2 \circ G - V_2](y)$ only if $y \in \mathcal{E}_2$. Since we saw in Part 1 that $V_l(y) = 0$ for all $y \in \mathcal{E}_l$, then the set $E_2 = \{y \in \mathcal{E}_1 \mid 0 \in [V_2 \circ G - V_2](y)\}$ is contained in a single level set of V_2. Then, we have that $\Omega(\phi) \subset \mathcal{E}_2$. We repeat for all $l \in \{3, \ldots, m\}$ the analysis performed on $l = 2$, then we will obtain that $\Omega(\phi) \subset \mathcal{E}_m$, where $\mathcal{E}_m$ is by definition the set of optimal solutions of the PEV charging problem.

2.7 Summary

In this work we define the PEV charging problem under usage schedule constraints, where PEVs have to compute charging strategies that allow them to fulfill charging deadline requirements at different moments of the day. This must be carried out optimally, in the sense of minimizing the aggregate energy price over the day. In order to solve this problem in a hierarchical way, we introduce a modification of the Price Leveling algorithm that has been presented in [12], which accounts for the usage constraints. Further, we introduce a mild modification of this algorithm: the Non-Anonymous Price Leveling algorithm, to solve the same problem, assuming that communication failures can occur. In order to show convergence in both cases, we introduce an invariance result for discrete-time systems that are represented by a difference inclusion. We also discuss the Anonymous Price Leveling algorithm, an anonymous setting under communication failures, which does not converge to an optimal solution of the PEV charging problem. It could be possible to design the aggregator for the Anonymous Price Leveling algorithm as an observer where $\sum_{i \in I_k} u_i$ corresponds to the measure of the output, and $\sum_{i \in I} u_i$ corresponds to the real output, so that the system performance could be improved. How

to do this while computing bounds on performance is the subject of our future work. Simulations illustrate the algorithm performance under communication failures.

Acknowledgements

This chapter contains work previously published in:

- A. Cortés and S. Martínez, "Optimal plug-in electric vehicle charging with schedule constraints," in the proceedings of Allerton (2013), 262-266.
- A. Cortés and S. Martínez, "A Hierarchical Algorithm for Optimal Plug-in Electric Vehicle Charging with Usage Constraints," submitted to Automatica, (2014), revised 2015.

Chapter 3

A hierarchical V2G protocol for PEVs

In the previous chapter, we discussed the challenges of the expected large penetration of PEVs in the power grid. We also presented a V1G approach to address such challenge.

In this chapter we work on the paradigm that PEVs can deliver/absorb power to the grid in order to provide ancillary services [25]. To this end, PEVs' batteries are used during inactivity periods to absorb the generation excess, if any, or to inject power power if there is demand excess. In this manuscript, we propose a hierarchical architecture in which intermediate aggregators coordinate with PEVs an optimal vehicle-2-grid (V2G) charging strategy. To do so, intermediate pricing and aggregated load signals are employed, which helps with privacy preservation goals.

The literature for V2G includes [26], where a centralized optimization problem is solved using simulated annealing and ant-colony optimization algorithms. Further, [27] presents a purely centralized optimal control algorithm to solve a V2G problem with uncertainty. A V2G game-theoretic formulation is given in [28], where PEVs are modeled as batteries that aim to inject to or draw from an aggregator a certain amount of energy in order to meet a desired energy state in such aggregator. In [28], the authors do not consider battery dynamics.

In this chapter, we present a hierarchical V2G problem formulation for the coordination of a fleet of PEVs connected to different points of the distribution side of the power grid. For simplicity, we assume that the distribution feeder follows a tree topology, which is consistent with the actual topology of existing settings. Then, we model

the power grid as a rooted tree graph, where the root node is the generation/transmission section of the grid, including the distribution substations. Other nodes represent buses on the distribution feeder, and PEVs are modeled as leaves of the tree. We exploit this hierarchy to define the communication structure of our coordination algorithm. Each bus has a non-PEV load that must be satisfied, and lines between buses have a maximum transmission capacity. Then, we formulate the V2G HIERARCHICAL algorithm. In this approach, all PEVs solve a local optimization problem to compute their charging/discharging profile over a finite discrete-time horizon. Then, they communicate it to the bus at which they are connected, which aggregates its PEV and its non-PEV load and sends it to the next bus up in the hierarchy. The aggregation is performed in a cascaded manner until the overall load reaches the root node, which then uses such information to compute a control signal that is down-streamed through the tree. As this signal passes through each bus, it is modified to account for the capacity limitations of the transmission lines that carry power to such bus. Finally, a modified signal reaches each PEV in the system, which employs it to recompute their charging/discharging profile. In this approach, PEVs do not have to provide directly private information on their usage habits to the root, but to intermediate buses, which helps with privacy preservation. Additionally, almost the entire computational load falls on the PEVs, while all buses in the network only act as aggregators. This results into good scalability properties of the algorithm. A convergence analysis is performed and simulations show the V2G HIERARCHICAL algorithm performance on various scenarios.

This chapter is organized as follows: Section 3.1 presents some preliminary notation on graph theory for this chapter. Section 3.2 introduces the V2G control problem to solve. Section 3.3 presents the formulation of the V2G problem as an optimal control problem. Section 3.4 introduces the hierarchical approach to the solution of the problem. Simulations and discussion are shown in Section 3.5.

3.1 Preliminary notation for this chapter

Consider an undirected graph $\mathcal{G} \triangleq (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is the set of nodes and $\mathcal{E}$ is the set of edges. A *path* $\mathcal{P}(i, j)$, on $\mathcal{G}$ for nodes $j, j \in \mathcal{V}$, is defined as a sequence

of nodes $\{n_1, \ldots, n_q\}$ such that $n_1 = i$, $n_q = j$ and $(n_\ell, n_{\ell+1})$ is an edge of $\mathcal{G}$, for all $\ell \in \{1, \ldots, q-1\}$. A graph $\mathcal{G}$ is a tree if there is a unique path between any two nodes $i, j \in \mathcal{V}$. The distance from node i to node j in $\mathcal{G}$ is given by the number of edges in the path $\mathcal{P}(i, j)$. For an undirected tree, any $\mathbf{r} \in \mathcal{V}$ can be called a *root* of $\mathcal{G}$. Then, $\mathcal{G}$ with root $\mathbf{r}$ is a *rooted tree*. For a node $j \in \mathcal{V}$, the set of *children* $\mathrm{ch}(j)$ is composed by all nodes that are connected by a single link to j, and whose distance to $\mathbf{r}$ is larger than that of j. The set of *descendants* of j, $\mathrm{des}(j)$, is the set of all nodes $i \in \mathcal{V} \setminus \{j\}$ such that $j \in \mathcal{P}(\mathbf{r}, i)$. Similarly, the *parent* of j, denoted as $\mathrm{pr}(j)$ is the unique node that is connected to j by a single edge, and belongs to $\mathcal{P}(\mathbf{r}, j)$. The set of *ancestors* of j, denoted by $\mathrm{an}(j)$, is the set of all nodes in $\mathcal{P}(\mathbf{r}, j) \setminus \{j\}$. A node $j \in \mathcal{V}$ is called a *leaf* of $\mathcal{G}$ if it is only connected to one node $l \in \mathcal{V}$.

3.2 Problem formulation

Consider a population of n plug-in electric vehicles (PEV) that is connected to the power grid. This population is spread over a large area. The objective of each PEV is to charge its battery in order to fulfill its user's needs. Additionally, each PEV is able to inject power back into the grid.

3.2.1 Structure of the power network

The power grid is composed of three easily discernible layers: generation, transmission, and distribution. We assume that the distribution side of the grid is composed of radial feeders only (with tree topology), and each feeder has a single connection point to the transmission grid. This is a reasonable assumption, as most of the existing feeders have this structure.

In our model, the generation/transmission side of the grid is condensed in the root node, $\mathbf{r}$, of a tree $\mathcal{T} \triangleq (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} \triangleq \{1, \ldots, \mathbf{r}\}$, and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. All the distribution trees that we consider in our model branch out of the root, and all buses in the distribution feeders are represented by nodes of the tree $\mathcal{T}$. In addition, PEVs attached to such buses are also represented by nodes of $\mathcal{T}$, however, they must be thought of as leaves of the tree. Without loss of generality, let us denote by $\mathcal{N} \triangleq \{1, \ldots, N\} \subset \mathcal{V}$ the

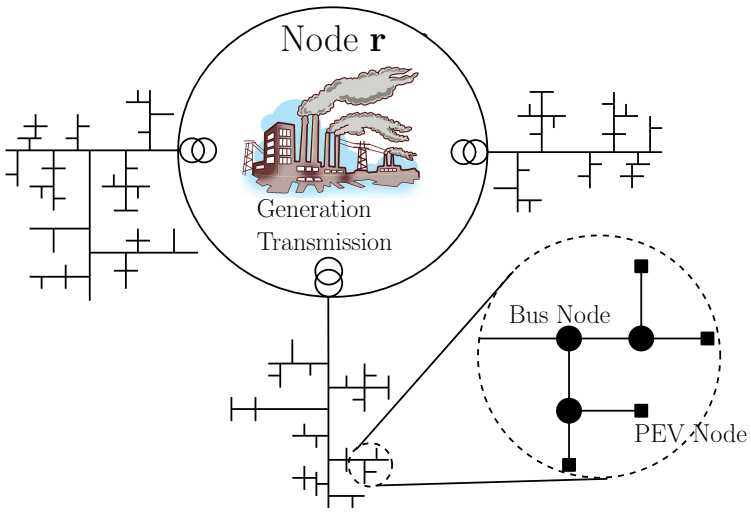

Figure 3.1: Graph model of the power grid. Node **r** represents the generation/transmission part of the grid, along with the connection points of the distribution feeders. In the zoomed area, circular nodes represent buses of the distribution feeder, while squared nodes represent PEVs.

subset of nodes of $\mathcal{T}$ that are PEVs connected to the system, while $\mathcal{M} \triangleq \{N+1, \ldots, \mathbf{r}\}$ consists of the buses of the distribution feeders of the system, along with the generation/transmission node **r**. Hence, all distribution buses belong to the set $\mathcal{M} \setminus \{\mathbf{r}\}$. In our setting, $\mathcal{M} \setminus \{\mathbf{r}\}$ does not contain nodes that represent connection points of the distribution feeders to the grid, as these are already encompassed within **r**. Buses have an associated non-PEV demand that must be satisfied.

As already mentioned, nodes representing PEVs have no children, i.e., $\text{ch}(i) = \emptyset$ for all $i \in \mathcal{N}$, while for the node **r**, $\text{pr}(\mathbf{r}) = \emptyset$. Each transmission line between two connected buses, i.e. $i, j \in \mathcal{M} \setminus \{\mathbf{r}\}$, for which there exists $(i, j) \in \mathcal{E}$, has an associated parameter $P_{ij}^{\max}(t)$ that corresponds to an upper bound on the amount of power that can go through the line (i, j). Likewise, there is a maximal amount of power $P_{\mathbf{r}}^{\max}(t)$ that the grid generation represented by node **r** can provide at each time t. Figure 3.1 shows a graphical explanation of the grid model.

3.2.2 PEV battery model

We assume that the battery of each PEV follows the dynamics:

$$z_{i,t} = z_{i,t-1} + \frac{\alpha_i^c}{\beta_i} u_{i,t} - \frac{1}{\alpha_i^d \beta_i} v_{i,t},$$

where $u_{i,t} \geq 0$ is the amount of energy that is charged into the battery during time interval $t \in \mathbb{N}$, $v_{i,t} \geq 0$ is the energy discharged from the battery during time $t \in \mathbb{N}$, $\alpha_i^c \in (0,1)$ is the battery system charging efficiency, $\alpha_i^d \in (0,1)$ is the battery system discharging efficiency, β_i stands for the battery capacity, and $z_{i,t}$ is the state of charge (SOC) at time $t \in \mathbb{N}$. The SOC must satisfy that $z_{i,t} \in [z_{i,\min}, z_{i,\max}]$, for $0 \leq z_{i,\min} < z_{i,\max} \leq 1$. In addition, some power bounds must be established in the battery charging/discharging, namely $u_{i,t} \leq u_{i,\max}$ and $v_{i,t} \leq v_{i,\max}$. Then, the charging/discharging action of each PEV, $i \in \mathcal{N}$, can be characterized by a demand profile $d_i \triangleq \{d_{i,t}\}_{t\in\mathbb{N}}$, where $d_{i,t} \triangleq u_{i,t} - v_{i,t}$.

3.2.3 Load buses in the distribution feeders

Each of the nodes $i \in \mathcal{M}$ represents a bus in a distribution feeder. The bus is characterized by a non-PEV load attached to it, denoted by $L_{i,t}$, which must be satisfied. Moreover, each node $i \in \mathcal{M}$ has a demand profile $d_i \triangleq \{d_{i,t}\}_{t\in\mathbb{N}}$ associated to it. This demand is given by all the power that is injected to all loads corresponding to node i or its children, i.e.:

$$d_{i,t} \triangleq \sum_{\ell \in \mathrm{ch}(i)} d_{\ell,t} + L_{i,t}. \tag{3.1}$$

By this definition of $d_{j,t}$, $j \in \mathcal{M}$, $d_{i,t}$ can be rewritten in terms of the descendants of i as follows:

$$d_{i,t} \triangleq L_{i,t} + \sum_{\ell \in \mathrm{des}(i)\cap\mathcal{M}} L_{\ell,t} + \sum_{j \in \mathrm{dN}(i)} d_{j,t}, \tag{3.2}$$

for all $i \in \mathcal{M}$, where $\mathrm{dN}(j) \triangleq \mathcal{N} \cap \mathrm{des}(j)$. Since PEVs may be descendants of node i, $d_{i,t}$ may be negative, which means that power is flowing upstream from node i toward its parent $\mathrm{pr}(i)$.

Remark 3.1. *Notice that the power flowing through line* (i, j)*,* $i, j \in \mathcal{M}$ *at time* $t \in \tau$ *is given by* $d_{i,t}$*, if* $j \in \text{pr}(i)$*, and* $d_{j,t}$ *if* $i \in \text{pr}(j)$*. This comes from the radial structure of all distribution feeders and the fact that the demand* $d_{i,t}$ *must be satisfied for all* $i \in \mathcal{V}$ *and for all* $t \in \tau$*. Given the rooted tree structure of the network, there is a one-to-one correspondence between nodes in* $\mathcal{M} \setminus \{\mathbf{r}\}$ *and the transmission lines in the distribution-side. Then, in order to account for the bounds in the transmission capacity, for all distribution lines, it suffices to pose the following constraints:*

$$|d_{j,t}| \leq P^{\max}_{j\,\text{pr}(j)}(t), \quad \forall t \in \tau, j \in \mathcal{M} \setminus \{\mathbf{r}\},$$

where, consistent with the notation, $P^{\max}_{j\,\text{pr}(j)}(t)$ *is the transmission capacity of the line between* j *and its parent. Therefore, for simplicity of notation, let us denote the capacity of the line between* j *and its parent by* $P^{\max}_{j}(t)$*, leading to:*

$$|d_{j,t}| \leq P^{\max}_{j}(t), \quad \forall t \in \tau, j \in \mathcal{M} \setminus \{\mathbf{r}\},$$

$\diamond$

3.2.4 The generation/pricing node

The node $\mathbf{r} \in \mathcal{V}$, referred to as the generation/price node models the behavior of the generation/transmission side of the power grid. For simplicity, we consider that the transmission lines in the transmission side of the grid do not have an upper limit on the amount of power they can carry. However, we do consider that there is an upper bound on the amount of energy that can be generated by the generation-side of the grid.

Therefore, the node $\mathbf{r}$ is solely represented by the maximum amount of power that it can provide to the distribution feeders, i.e., $P^{\max}_{\mathbf{r}}(t)$, and the generation cost for the energy supplied by the grid at time $t \in \mathbb{N}$. This generation cost is given by $C : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, which is a convex and increasing function, with $C(0) = 0$. The argument of this function corresponds to the aggregate power that is provided by the grid to the loads at time t. The function C models a market behavior in which the price varies according to the demand, and the whole demand must be satisfied. In case the i^{th} PEV is providing power at certain time, the function C will also determine the price to be paid to the

owner of the PEV.

Assumption 3.1 (Derivative of C is Lipschitz)**.** *The function C is such that C' is Lipschitz in its domain, with Lipschitz constant l_C.* ◇

3.3 Optimal control problem

The charging strategy is devoted to minimize a function corresponding to the total cost of the energy provided by the utility during a finite horizon $\tau \triangleq \{1, \dots, T\}$, subject to user needs and line capacity constraints. In addition, let $Z_i \subset \tau$, for all $i \in \mathcal{N}$ be a set of times at which the i^{th} vehicle has access to the power grid. This set can be used to model deadlines in the charging time for each PEV. Taking into account the consideration in Remark 3.1, we formulate the following optimization problem:

Problem 1: $\min_{u,v} J(u, v)$

subject to:

$$(u_i, v_i) \in \mathcal{F}_i, \quad \forall i \in \mathcal{N}, \tag{3.3a}$$

$$d_{i,t} = u_{i,t} - v_{i,t}, \quad \forall t, i \in \mathcal{N}, \tag{3.3b}$$

$$d_{j,t} = \sum_{\ell \in \text{ch}(j)} d_{\ell,t} + L_{j,t}, \quad \forall t, j \in \mathcal{M}, \tag{3.3c}$$

$$|d_{j,t}| \leq P_j^{\max}(t), \quad \forall t, j \in \mathcal{M} \setminus \{\mathbf{r}\}, \tag{3.3d}$$

$$d_{\mathbf{r},t} \leq P_{\mathbf{r}}^{\max}(t), \quad \forall t. \tag{3.3e}$$

Here,

$$J(u, v) = \sum_{t=1}^{T} C(d_{\mathbf{r},t}), \tag{3.4}$$

and $(u_i, v_i) \in \mathcal{F}_i$ if the following constraints hold:

$$z_{i,t} = z_{i,0} + \frac{1}{\beta_i} \sum_{\ell=1}^{t} \left(\alpha_i^c u_{i,\ell} - \frac{1}{\alpha_i^d} v_{i,\ell} \right), \quad t \in \tau, \tag{3.5a}$$

$$z_{i,\min} \leq z_{i,t} \leq z_{i,\max}, \quad t \in \tau, \tag{3.5b}$$

$$0 \leq u_{i,t} \leq u_{i,\max}, \quad t \in \tau, \tag{3.5c}$$

$$0 \leq v_{i,t} \leq v_{i,\max}, \quad t \in \tau, \tag{3.5d}$$

$$u_{i,t} = 0, \quad t \notin Z_i \subseteq \tau, \tag{3.5e}$$

$$v_{i,t} = 0, \quad t \notin Z_i \subseteq \tau, \tag{3.5f}$$

$$z_{i,T} = z_{i,\max}, \tag{3.5g}$$

$$u_{i,t} v_{i,t} = 0, \quad t \in \tau. \tag{3.5h}$$

Notice that the constraint (3.5h) is not convex. The following results allow us to relax the constraint without affecting the solution of the problem.

Lemma 3.1. *Let u, v be a feasible solution to the relaxed version of Problem 1. Then, there exists some $s \in \tau$, such that for all $t \in \{s, \ldots, T\} \subset \tau$, and for all $\ell \in \mathcal{M}$, $d_{\ell,s} \geq 0$.*

Proof. The result follows easily by observing that if u, v is feasible, $z_{i,T} = z_{i,\max}$, for all $i \in \mathcal{N}$. Then, for $d_{\ell,T}$ to be negative, there must be at least some $i \in \mathrm{dN}(\ell)$, such that $u_{i,T} - v_{i,T} < 0$. Then it must be that $z_{i,T-1} > z_{i,\max}$, which contradicts the feasibility of u, v. □

Lemma 3.2. *Let u, v be a feasible solution to the relaxed version of Problem 1, with some $i \in \mathcal{N}$ such that for some $t \in \tau$, $v_{i,t} > u_{i,t}$. Construct a new solution u^1, v^1 based on u, v, such that $u^1_{i,t} = \xi_1 u_{i,t}$, $v^1_{i,t} = \xi_1 v_{i,t}$, $\xi_1 \in (0,1)$, $u^1_{i,q} = \xi_2 u_{i,q}$, $v^1_{i,q} = \xi_2 v_{i,q}$ for all $q \in \{t+1, \ldots, T\}$, where $\xi_2 \in (0,1)$ is chosen as $\xi_2 \triangleq 1 + \frac{(1-\xi_1)(\alpha_i^c \alpha_i^d u_{i,t} - v_{i,t})}{\sum_{q=t+1}^{T} (\alpha_i^c \alpha_i^d u_{i,q} - v_{i,q})}$, while all other components of u^1, v^1 are identical to those of u, v. This solution is such that $u^1_j, v^1_j \in \bar{\mathcal{F}}_j$, for all $j \in \mathcal{N}$.*

Proof. Since $(u^1_j, v^1_j) = (u_j, v_j)$ for all $j \in \mathcal{N} \setminus \{i\}$, then it holds that $(u^1_j, v^1_j) \in \bar{\mathcal{F}}_j$. Moreover, since $(u^1_{i,q}, v^1_{i,q}) = (u_{i,q}, v_{i,q})$ for all $q \leq t$, all the constraints (3.5a) through (3.5f) hold for i until time $q = t - 1$. Then, to prove our statement, we only need to show that

the constraints (3.5a) through (3.5f) hold for i for all times $q \geq t$, and also that (3.5g) holds. Since $\xi_1 \in (0,1)$, it follows that $u^1_{i,t} \in [0, u_{i,\max}]$, $v^1_{i,t} \in [0, v_{i,\max}]$. Moreover, since $v_{i,t} > u_{i,t}$, it also holds that $0 > \xi_1(\alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t}) > \alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t}$. Then, given that $z^1_{i,t-1} = z_{i,t-1}$, and the fact that $0 > \xi_1(\alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t})$ it follows that $z_{i,\max} \geq z_{i,t-1} > z^1_{i,t} > z_{i,t} \geq z_{\min,i}$. Next, let us show that $\xi_2 \in (0,1)$. Notice that from the constraints (3.5a), (3.5g), we have that:

$$0 \leq z_{i,\max} - z_{i,t} = \frac{1}{\beta_i}(\alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t}) + \textstyle\sum_{q=t+1}^{T} \frac{1}{\beta_i}(\alpha^c_i u_{i,q} - (\alpha^d_i)^{-1} v_{i,q}).$$

Since $\alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t} < 0$ by the assumption of the lemma, we move it to the left side of the inequality above and it follows that:

$$-\frac{1}{\beta_i}(\alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t}) < \sum_{q=t+1}^{T} \frac{1}{\beta_i}(\alpha^c_i u_{i,q} - (\alpha^d_i)^{-1} v_{i,q}).$$

Then, after dividing both sides for the positive term on the right-hand side, it follows that $\xi_2 \in (0,1)$. It immediately implies that $u^1_{i,q} \in [0, u_{i,\max}]$ and $v^1_{i,q} \in [0, v_{i,\max}]$, for $q \in \{t+1, \ldots, T\}$. Next, let us show that the constraint (3.5b) holds for all $q \in \{t+1, \ldots, T\}$. From (3.5a) and the fact that $(u^1_{i,q}, v^1_{i,q}) = (\xi_2 u_{i,q}, \xi_2 v_{i,q})$, $q \in \{t+1, \ldots, T\}$, we have that:

$$z^1_{i,q} = z_{i,t-1} + \frac{\xi_1}{\beta_i}\left(\alpha^c_i u_{i,t} - \frac{1}{\alpha^d_i} v_{i,t}\right) + \frac{\xi_2}{\beta_i} \sum_{\ell=t+1}^{q} \left(\alpha^c_i u_{i,\ell} - \frac{1}{\alpha^d_i} v_{i,\ell}\right).$$

If the third summand in the expression above is less than zero, by negativity of $\alpha^c_i u_{i,t} - (\alpha^d_i)^{-1} v_{i,t}$, it holds that $z_{i,\max} \geq z_{i,t-1} > z^1_{i,q}$. Moreover, since $\xi_1, \xi_2 \in (0,1)$, it follows by definition of $z_{i,q}$ and $z^1_{i,q}$ that $z^1_{i,q} > z_{i,q} \geq z_{i,\min}$.

If the third summand in the expression above is greater or equal than zero, then we have:

$$z^1_{i,q} = z_{i,t-1} + \frac{\xi_3(1-\xi_1) + \xi_1}{\beta_i}\left(\alpha^c_i u_{i,t} - \frac{1}{\alpha^d_i} v_{i,t}\right) + \frac{1}{\beta_i} \sum_{\ell=t+1}^{q} \left(\alpha^c_i u_{i,\ell} - \frac{1}{\alpha^d_i} v_{i,\ell}\right), \tag{3.6}$$

where:

$$\xi_3 = \frac{\frac{1}{\beta_i}\sum_{\ell=t+1}^{q}\left(\alpha_i^c u_{i,\ell} - \frac{1}{\alpha_i^d}v_{i,\ell}\right)}{\frac{1}{\beta_i}\sum_{w=t+1}^{T}\left(\alpha_i^c u_{i,w} - \frac{1}{\alpha_i^d}v_{i,w}\right)}.$$

Clearly, $\xi_3 \in (0,1)$. Then, $\xi_3(1-\xi_1)+\xi_1 < 1$, and since $\alpha_i^c u_{i,t} - (\alpha_i^d)^{-1}v_{i,t} < 0$, it follows that $z_{i,q}^1 \leq z_{i,q}$, for all $q > t$, and from (3.6) we have:

$$\begin{aligned} z_{i,q}^1 \geq & z_{i,t-1} + \frac{1}{\beta_i}\left(\alpha_i^c u_{i,t} - \frac{1}{\alpha_i^d}v_{i,t}\right) + \frac{1}{\beta_i}\sum_{\ell=t+1}^{q}\left(\alpha_i^c u_{i,\ell} - \frac{1}{\alpha_i^d}v_{i,\ell}\right) \\ = & z_{i,q}. \end{aligned}$$

Then, $z_{i,t}^1 \geq z_{i,\min}$, for all $q > t$. The constraint (3.5g) can be easily verified by replacing ξ_2 in the battery dynamics. Then, $u_i^1, v_i^1 \in \bar{\mathcal{F}}_i$, and the result follows. □

Lemma 3.3 (Exact convex relaxation of Problem 1)**.** *The constraint $u_{i,t}v_{i,t} = 0$, for all $t \in \tau$, $i \in \mathcal{N}$ can be relaxed and the optimal solutions of the relaxed problem are exactly the optimal solutions for Problem 1.* ◇

Proof. For the proof of this result, we proceed by contradiction, by showing that if $u^\star, v^\star$ is an optimal solution to Problem 1, without the nonconvex constraint (3.5h), but there is some $i \in \mathcal{N}$, $t \in \tau$, such that $u_{i,t}^\star v_{i,t}^\star > 0$, then we can construct a solution that outperforms $u^\star, v^\star$. Assume that $u^\star, v^\star$ is an optimizer of Problem 1 without the nonconvex constraint (3.5h), and there is some $i \in \mathcal{N}$ and t such that $u_{i,t}^\star > 0$ and $v_{i,t}^\star > 0$. Consider a solution $\hat{u}^{\star,1}, \hat{v}^{\star,1}$, such that $(\hat{u}_{j,q}^{\star,1}, \hat{v}_{j,q}^{\star,1}) = (u_{j,q}^\star, v_{j,q}^\star)$ for all $(j,q) \neq (i,t)$, $t, q \in \tau$, $j \in \mathcal{N}$, and $\hat{u}_{i,t}^{\star,1} = \max\{0, u_{i,t}^\star - (\alpha_i^c\alpha_i^d)v_{i,t}^\star\}$, and $\hat{v}_{i,t}^{\star,1} = \max\{0, v_{i,t}^\star - (\alpha_i^c\alpha_i^d)u_{i,t}^\star\}$. Clearly, $\hat{u}^{\star,1}, \hat{v}^{\star,1}$ is such that $\hat{u}_j^{\star,1}, \hat{v}_j^{\star,1} \in \bar{\mathcal{F}}_j$ for all $j \in \mathcal{N}$. Moreover, since $u_{i,t}^\star, v_{i,t}^\star \geq 0$, it is easy to see that:

$$\max\{u_{i,t}^\star - (\alpha_i^c\alpha_i^d)^{-1}v_{i,t}^\star, (\alpha_i^c\alpha_i^d)u_{i,t}^\star - v_{i,t}^\star\} < u_{i,t}^\star - v_{i,t}^\star. \tag{3.7}$$

Therefore, the only constraints that $\hat{u}^{\star,1}, \hat{v}^{\star,1}$ may not satisfy are: $-P_\ell^{\max}(t) \leq \hat{d}_{\ell,t}^{\star,1}$ for some $\ell \in \text{an}(i)$.

If for all $\ell \in \mathcal{M}$, it holds that $-P_\ell^{\max}(t) \leq \hat{d}_{\ell,t}^{\star,1}$, then, $\hat{u}^{\star,1}, \hat{v}^{\star,1}$ is feasible to Problem 1 with relaxed nonconvex constraint, and from (3.7), it follows that $\hat{u}_{i,t}^{\star,1} -$

$\hat{v}^{\star,1}_{i,t} < u^\star_{i,t} - v^\star_{i,t}$, therefore, $\sum_{j\in\mathcal{N}}(\hat{u}^{\star,1}_{j,t} - \hat{v}^{\star,1}_{j,t}) < \sum_{j\in\mathcal{N}}(u^\star_{j,t} - v^\star_{j,t})$, and $\sum_{j\in\mathcal{N}}(\hat{u}^{\star,1}_{j,q} - \hat{v}^{\star,1}_{j,q}) = \sum_{j\in\mathcal{N}}(u^\star_{j,q} - v^\star_{j,q})$, for all $q \neq t$, $q \in \tau$. Given that the function C is convex and increasing, $J(\hat{u}^{\star,1}, \hat{v}^{\star,1}) < J(u^\star, v^\star)$, which contradicts the fact that $u^\star, v^\star$ is optimal.

Let us consider the opposite case, i.e., there is some $\ell \in \text{an}(i)$ such that $-P^{\max}_\ell(t) > \hat{d}^{\star,1}_{\ell,t}$. Let us fix such ℓ. Then, there is a nonempty set $\mathcal{D}(\ell, \hat{u}^{\star,1}, \hat{v}^{\star,1}) \subset \text{dN}(\ell)$ such that $\hat{u}^{\star,1}_{j,t} - \hat{v}^{\star,1}_{j,t} < 0$ for all $j \in \mathcal{D}(\ell, \hat{u}^{\star,1}, \hat{v}^{\star,1})$. Then, we construct a solution $\hat{u}^{\star,2}, \hat{v}^{\star,2}$ based on $\hat{u}^{\star,1}, \hat{v}^{\star,1}$ as described in Lemma 3.2, with $(\hat{u}^{\star,2}_{j,t}, \hat{v}^{\star,2}_{j,t}) = (\xi_1\hat{u}^{\star,1}_{j,t}, \xi_1\hat{v}^{\star,1}_{j,t})$, for all $j \in \mathcal{D}(\ell, \hat{u}^{\star,1}, \hat{v}^{\star,1})$, where ξ_1 is chosen as:

$$\xi_1 \triangleq 1 - \frac{(-P^{\max}_\ell(t) - \hat{d}^{\star,1}_{\ell,t})}{\sum_{s\in\mathcal{D}(\ell,\hat{u}^{\star,1},\hat{v}^{\star,1})}(\hat{u}^{\star,1}_{s,t} - \hat{v}^{\star,1}_{s,t})},$$

$(\hat{u}^{\star,2}_{j,q}, \hat{v}^{\star,2}_{j,q}) = (\xi_2\hat{u}^{\star,1}_{j,q}, \xi_2\hat{v}^{\star,1}_{j,q})$, $q \in \{t+1, \dots, T\}$, for all $j \in \mathcal{D}(\ell, \hat{u}^{\star,1}, \hat{v}^{\star,1})$, and ξ_2 as specified in Lemma 3.2. By definition of $\mathcal{D}(\ell, \hat{u}^{\star,1}, \hat{v}^{\star,1})$ and $\hat{d}^{\star,1}_{\ell,t}$, it holds that $\xi_1 \in (0, 1)$. Then, we choose all other components of $\hat{u}^{\star,2}, \hat{v}^{\star,2}$, equal to those of $\hat{u}^{\star,1}, \hat{v}^{\star,1}$, according to Lemma 3.2, to obtain $\hat{u}^{\star,2}, \hat{v}^{\star,2}$ such that $\hat{u}^{\star,2}_j, \hat{v}^{\star,2}_j \in \bar{\mathcal{F}}_j$ for all $j \in \mathcal{N}$. It can be seen that the choice of $\hat{u}^{\star,2}, \hat{v}^{\star,2}$ leads to $-P^{\max}_\ell(t) = \hat{d}^{\star,2}_{\ell,t}$, for all $j \in \text{dN}(\ell)$, and also $\hat{d}^{\star,2}_{\ell,q} \leq \hat{d}^{\star,1}_{\ell,q}$ for all $q \in \{t+1, \dots, T\}$. Next, if for all $q \in \{t+1, \dots, T\}$, $-P^{\max}_\ell(q) \leq \hat{d}^{\star,2}_{\ell,q}$, it means that $\hat{u}^{\star,2}, \hat{v}^{\star,2}$ is feasible for the relaxation of Problem 1. Hence, we can follow the same reasoning as in the case where $\hat{u}^{\star,1}, \hat{v}^{\star,1}$ was feasible to get $J(\hat{u}^{\star,2}, \hat{v}^{\star,2}) < J(u^\star, v^\star)$, which contradicts the optimality of $u^\star, v^\star$. If there is some $q \in \{t+1, \dots, T\}$ for which $\hat{d}^{\star,2}_{\ell,q} < -P^{\max}_\ell(q)$, we repeat the same procedure to find a solution $\hat{u}^{\star,3}, \hat{v}^{\star,3}$ such that $\hat{d}^{\star,2}_{\ell,q} = -P^{\max}_\ell(q)$. Then, since T is finite, and by Lemma 3.1, one can recursively construct solutions $\hat{u}^{\star,w}, \hat{v}^{\star,w}$ based on $\hat{u}^{\star,w-1}, \hat{v}^{\star,w-1}$, until eventually $\hat{u}^{\star,\bar{w}}, \hat{v}^{\star,\bar{w}}$, is feasible for the convex relaxation of Problem 1, for some $\bar{w} \leq T - t$. Given the construction of $\hat{u}^{\star,\bar{w}}, \hat{v}^{\star,\bar{w}}$, it outperforms $u^\star, v^\star$. Since we assume that $u^\star, v^\star$ is optimal, then we reach a contradiction. □

The next result is an adaptation of Theorem 1 in [11], and shows the uniqueness of the optimal demand profile generated by the optimizers of Problem 1.

Lemma 3.4. *(Uniqueness of the aggregate demand profile): Let $u^\star, v^\star$ and $\hat{u}, \hat{v}$ be optimizers of Problem 1. Then it holds that $\sum_{i\in\mathcal{N}}(u^\star_i - v^\star_i) = \sum_{i\in\mathcal{N}}(\hat{u}_i - \hat{v}_i)$.* ◇

The solution of this optimization problem is *valley filling* and *peak-shaving*, i.e., if u, v is optimal, the PEVs will try to provide as much energy as possible in the highest-price times and will try to obtain as much energy as possible in the lowest-price times.

In order to solve Problem 1, we use penalty functions to handle the coupling constraints. We formulate the following relaxation of the problem:

$$\text{Problem 2: } \min_{u,v} J(u,v) + \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \kappa_\ell \Phi_\ell(d_{\ell,t})$$

subject to:

$$(u_i, v_i) \in \bar{\mathcal{F}}_i, \quad \forall i \in \mathcal{N} \tag{3.8a}$$

$$d_{i,t} = u_{i,t} - v_{i,t}, \quad \forall t, i \in \mathcal{N} \tag{3.8b}$$

$$d_{j,t} = \sum_{\ell \in \text{ch}(j)} (d_{\ell,t} + L_{j,t}), \quad \forall t, j \in \mathcal{M}, \tag{3.8c}$$

where $\Phi_j : \mathbb{R} \to \mathbb{R}_{\geq 0}$ acts as a penalty function for the power constraint at node $j \in \mathcal{M}$, defined as:

$$\Phi_j(d_{j,t}) \triangleq (\max\{0, d_{j,t} - P_j^{\max}(t)\})^2,$$

and $\bar{\mathcal{F}}_i \triangleq \{(u,v) \in \mathbb{R}^{2nT} \mid \text{Eqns. (3.5a) –(3.5g) hold } \forall i \in \mathcal{N}\}$. Notice that:

$$\Phi_j'(d_{j,t}) = \max\{0, 2(d_{j,t} - P_j^{\max}(t))\},$$

then, $\Phi_j'(d_{j,t})$ is globally Lipschitz continuous with Lipschitz constant $l_B = 2$, for all $j \in \mathcal{M}$.

Remark 3.2. *Notice that we are using only a penalty function for inequalities $d_{\ell,t} \leq P_\ell^{\max}(t)$, for all $\ell \in \mathcal{M} \setminus \{\mathbf{r}\}$, $\ell \in \text{ch}(j)$, leaving aside $-P_\ell^{\max}(t) \leq d_{\ell,t}$. This is done for simplicity of presentation of the method and its analysis. The inclusion of a penalty function for the lower bound on $d_{\ell,t}$, can be treated in an analogous way.*

3.3.1 Analysis and design of the penalty method

From [29], it is known that for the penalty method to yield a feasible solution to the original problem, it is necessary to use non-differentiable penalty functions, except for selected cases, which our problem does not satisfy. Clearly, the quadratic penalty functions are continuously differentiable, therefore the solution to Problem 2 may not be feasible for Problem 1. However, It is also known that if $\kappa_\ell \to +\infty$, for all $\ell \in \mathcal{M}$, the solution to Problem 2 gets arbitrarily close to a solution to Problem 1.

Lemma 3.5. *If an optimal solution to Problem 2 satisfies the constraints* (3.3d)*, then such solution is also an optimal solution of Problem 1.* ◇

By the result above, we have that a solution to Problem 2 is not a solution to Problem 1 only if it violates at least one of the constraints given by (3.3d). Therefore, a suitable way to study how close the solution to Problem 2 is to a solution to Problem 1, is to analyze the maximum amount of constraint violation for a given value of the parameters κ_ℓ, for all $\ell \in \mathcal{M}$. In this way, one can design parameters κ_ℓ that lead to a desired tolerance on the constraint violation.

To this end, we introduce the following assumption.

Assumption 3.2 (Slater's Condition)**.** *There exists a feasible solution $u^\dagger, v^\dagger$ to Problem 1, such that:*

$$|d^\dagger_{\ell,t}| \leq P^{\max}_\ell(t) - \varepsilon, \quad \ell \in \mathcal{M},$$

for all $t \in \tau$, such that $P^{\max}_\ell(t) - \varepsilon > 0$, $\varepsilon > 0$ and does not depend on ℓ. ◇

The following result establishes how to choose all parameters κ_ℓ.

Proposition 3.1 (Characterization of κ_ℓ)**.** *Fix $\sigma \in (0, 1)$ and let :*

$$\kappa_\ell > \sqrt{|\mathcal{M}|T}\, J_{\max}/(\varepsilon\sigma \min_{\ell.t} P^{\max}_\ell(t)),$$

for all $\ell \in \mathcal{M}$, where

$$J_{\max} \triangleq \sum_{t=1}^{T} C(P^{\max}_{\mathbf{r}}(t)),$$

Let $u^\star, v^\star$ be a solution to Problem 2. Then, $u^\star, v^\star$ satisfies:

$$d^\star_{\ell,t} \leq P^{\max}_\ell(t)(1+\sigma),$$

for all $\ell \in \mathcal{M}$. ◇

Proof. First, consider an optimal solution $\hat{u}, \hat{v}$ to Problem 2, and an optimal solution $u^\star, v^\star$ to Problem 1. Notice that:

$$\sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \kappa_\ell \Phi_\ell(d^\star_{\ell,t}) = 0,$$

since $u^\star, v^\star$ satisfies all constraints (3.3d). Then, by optimality of $\hat{u}, \hat{v}$ in Problem 2, we have that:

$$J(\hat{u}, \hat{v}) + \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \kappa_\ell \Phi_\ell(\hat{d}_{\ell,t}) \leq J(u^\star, v^\star). \tag{3.9}$$

Next, define $\lambda^\star_{\ell,t}$, for all $\ell \in \mathcal{M}$, as the optimal Lagrange multipliers associated to the constraints in (3.3d) in Problem 1, which exist by Assumption 3.2 (Slater's condition). Let us define also the Lagrangian function for Problem 1 as follows:

$$\mathcal{L}(u, v, \lambda) \triangleq J(u, v) + \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \lambda_{\ell,t}(d_{\ell,t} - P^{\max}_\ell(t)).$$

By Duality Theory [30] and also Assumption 3.2 (Slater's condition), it follows that:

$$J(u^\star, v^\star) = \mathcal{L}(u^\star, v^\star, \lambda^\star) \leq \mathcal{L}(\hat{u}, \hat{v}, \lambda^\star). \tag{3.10}$$

Now, let us proceed by contradiction, to show that:

$$\max\{0, \hat{d}_{\ell,t} - P^{\max}_\ell(t)\} \leq \frac{\max_{\eta,t} \lambda^\star_{\eta,t}}{\min_\iota \kappa_\iota} \sqrt{|\mathcal{M}|T}, \tag{3.11}$$

for all $\ell \in \mathcal{M}$ and $t \in \tau$. For the sake of clarity, let us introduce the vector $\mathbf{g}(u, v) \in \mathbb{R}^s$, where $s \triangleq |\mathcal{M}|T$, whose components are $\max\{0, d_{\ell,t} - P^{\max}_\ell(t)\}$, for all ℓ, t.

Assume that $\hat{u}, \hat{v}$ is such that $\max\{0, \hat{d}_{\ell,t} - P^{\max}_\ell(t)\} > \sqrt{s}\xi$, where we define

$\xi \triangleq \max_{\eta,t} \lambda^\star_{\eta,t} / \min_\iota \kappa_\iota$, for some $\ell \in \mathcal{M}$, $t \in \tau$. It is equivalent to saying that $\|\mathbf{g}(\hat{u}, \hat{v})\|_\infty > \sqrt{s}\xi$. Since $\sqrt{s}\xi > 0$, we can multiply on both sides by $\|\mathbf{g}(\hat{u}, \hat{v})\|_1$ and move $\sqrt{s}$ to obtain $\|\mathbf{g}(\hat{u}, \hat{v})\|_1 \|\mathbf{g}(\hat{u}, \hat{v})\|_\infty / \sqrt{s} > \xi \|\mathbf{g}(\hat{u}, \hat{v})\|_1$. By properties of norms, we have that $\|\mathbf{g}(\hat{u}, \hat{v})\|_2 \geq \|\mathbf{g}(\hat{u}, \hat{v})\|_\infty$, and also $\|\mathbf{g}(\hat{u}, \hat{v})\|_2 \geq \|\mathbf{g}(\hat{u}, \hat{v})\|_1 / \sqrt{s}$. Then, it follows that $\|\mathbf{g}(\hat{u}, \hat{v})\|_2^2 \geq \|\mathbf{g}(\hat{u}, \hat{v})\|_1 \|\mathbf{g}(\hat{u}, \hat{v})\|_\infty / \sqrt{s} > \xi \|\mathbf{g}(\hat{u}, \hat{v})\|_1$, which implies that:

$$\begin{aligned}\sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \Phi(\hat{d}_{\ell,t}) &> \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \xi \max\{0, d_{\ell,t} - P^{\max}_\ell(t)\} \\ &\geq \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \frac{\max_{\eta,t} \lambda^\star_{\eta,t}}{\min_\iota \kappa_\iota} (d_{\ell,t} - P^{\max}_\ell(t)).\end{aligned}$$

The equation above implies that:

$$\sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \kappa_\ell \Phi(\hat{d}_{\ell,t}) > \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \lambda^\star_{\ell,t} (d_{\ell,t} - P^{\max}_\ell(t)),$$

which in turn implies that:

$$\mathcal{L}(\hat{u}, \hat{v}, \lambda^\star) < J(\hat{u}, \hat{v}) + \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \kappa_\ell \Phi_\ell(\hat{d}_{\ell,t}) \tag{3.12}$$

From Equations (3.9), (3.10), and (3.12), it follows that $J(u^\star, v^\star) < J(u^\star, v^\star)$, a contradiction, hence, (3.11) holds for all $\ell \in \mathcal{M}$, $t \in \tau$.

Now, from [31, Chapter 10], we have that:

$$\max_{\ell,t} \lambda^\star_{\ell,t} \leq \frac{1}{\gamma} (J(\bar{u}, \bar{v}) - \bar{J}),$$

where $\bar{J} = J(\tilde{u}, \tilde{v})$, $\tilde{u}, \tilde{v}$ is a solution to Problem 1, without the constraints (3.3d), $\bar{u}, \bar{v}$ is a Slater vector of Problem 1, and $\gamma \triangleq \min_{\ell,t} P^{\max}_\ell(t) - \bar{d}_{\ell,t}$. Note that by definition of Slater vector, $u^\dagger, v^\dagger$ as described in Assumption 3.2 is a Slater vector of Problem 1. By the same assumption, we have that $\gamma \geq \varepsilon$. This, together with the fact that $\bar{J}$ is nonnegative, leads us to:

$$\max_{\ell,t} \lambda^\star_{\ell,t} \leq \frac{1}{\varepsilon} J(\bar{u}, \bar{v}),$$

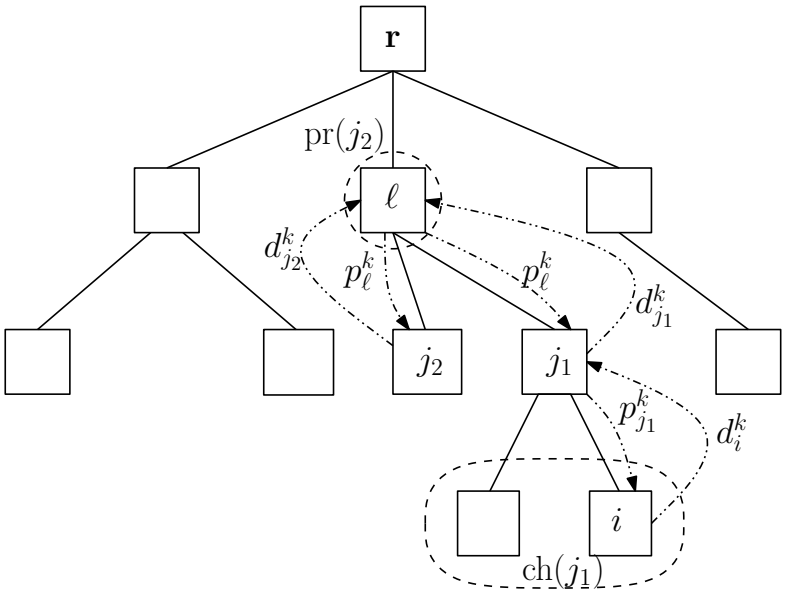

Figure 3.2: Grid hierarchical structure. Dashed lines indicate communication links while solid lines represent power links.

Finally, from the constraint in (3.3e), we have that:

$$J(\bar{u}, \bar{v}) \leq J_{\max}, \tag{3.13}$$

with $J_{\max}$ as defined in Lemma 3.1. Finally, from the choice of κ_ℓ presented in the statement of the Lemma, we have:

$$\frac{J_{\max}\sqrt{|\mathcal{M}|T}}{\min_\ell \kappa_\ell} \leq \sigma \min_{\iota,t} P_\iota^{\max}(t) \leq \sigma P_\ell^{\max}(t).$$

Then the result follows by combining it with (3.11) and (3.13). □

3.4 A hierarchical control architecture

Our main interest is to solve Problem 1 using a decentralized/hierarchical communication and control architecture that allows for distributed computation, scalability, and privacy. Next, we introduce a hierarchical approach for the solution of Problem 2. For now on, let us assume that the optimization problem is feasible.

3.4.1 V2G HIERARCHICAL algorithm

Our approach endows each congestible element in the grid with computation and communication capacity. Communications over the network follow the same tree topology as the power network. In this way, each element $i \in \mathcal{V} \setminus \{\mathbf{r}\}$ sends its parent the demand profile d_i and if $\text{ch}(i) \neq \emptyset$, implying that i is not a PEV, it sends its children a control signal which comes from the generation/pricing node $\mathbf{r}$, and the amount of violation on the maximum power constraints of i's ancestors.

The V2G HIERARCHICAL algorithm is inspired by the works presented in both [10, 11] but we modify the approach to account for the penalty functions of Problem 2. This is an iterative procedure in which at iteration $k \in \mathbb{N}$, each PEV generates a demand profile $d_{i,t}^k = u_{i,t}^k - v_{i,t}^k$, for all $t \in \tau$, that is feasible for its own battery constraints. Then, it transmits the profile to its parent, which in turn computes its own demand profile according to (3.1). This is done until the node $\mathbf{r}$ computes its demand profile; see Figure 3.2 for an illustration of the information flow over the communication network.

Based on this demand profile, the node $\mathbf{r}$ computes and transmits a coordination signal $p_{\mathbf{r}}^k \triangleq [p_{\mathbf{r},1}^k \dots, p_{\mathbf{r},T}^k]^\top \in \mathbb{R}^T$, such that:

$$p_{\mathbf{r},t}^k \triangleq \eta C'(d_{\mathbf{r},t}^k) + b_{\mathbf{r},t}^k,$$

for all $t \in \tau$, $\eta > 0$, with $b_{\mathbf{r},t}^k = \eta\kappa_{\mathbf{r}}\Phi'_{\mathbf{r}}(d_{\mathbf{r},t}^k)$, and Φ' as introduced with Problem 2, and $\kappa_j > 0$, for all $j \in \mathcal{M}$. Then, each node $j \in \mathcal{M} \setminus \{\mathbf{r}\}$ computes the control signal $p_j^k \triangleq [p_{j,1}^k \dots, p_{j,T}^k]^\top \in \mathbb{R}^T$:

$$p_{j,t}^k \triangleq p_{\text{pr}(j),t}^k + b_{j,t}^k,$$

for all $t \in \tau$, where:

$$b_{j,t}^k = \eta\kappa_j\Phi'_j(d_{j,t}^k).$$

Recall that $\text{pr}(j)$ denotes the node that is parent of node j. Further, as the signal reaches each PEV, it computes its next battery control by solving the following optimization

problem:

$$(u_i^{k+1}, v_i^{k+1}) = \operatorname{argmin}_{u_i, v_i} J_i(u_i, v_i) \tag{3.14}$$
$$\text{subject to:}$$
$$(u_i, v_i) \in \bar{\mathcal{F}}_i,$$

where:

$$J_i(u_i, v_i) = \sum_{t=1}^{T} p^k_{\mathrm{pr}(i),t}(u_{i,t} - v_{i,t}) + \frac{1}{2}\|u_i - v_i - d_i^k\|^2. \tag{3.15}$$

The procedure must be iterated until a stopping criterion is reached.

Theorem 3.1 (Convergence result)**.** *The* V2G HIERARCHICAL *algorithm converges to an optimizer of Problem 2 as $k \to \infty$, provided Assumption 3.1 (Derivative of C is Lipschitz) holds and $\eta < \min\{\eta_1, \eta_2\}$, where:*

$$\eta_1 \triangleq \min\{(N l_C(1 + |\operatorname{an}(i)|))^{-1} \mid i \in \mathcal{N}\},$$
$$\eta_2 \triangleq \min\{(2\kappa_\ell |\operatorname{dN}(\ell)|(1 + |\operatorname{an}(i)|))^{-1} \mid i \in \mathcal{N}, \ell \in \operatorname{an}(i)\}.$$

Proof. Let us choose the Lyapunov function:

$$V(u, v) = \sum_{t=1}^{T} \left(C(d_{\mathbf{r},t}) + \sum_{\ell \in \mathcal{M}} \kappa_\ell \Phi(d_{\ell,t}) \right).$$

Then, we aim to show that $V(u^{k+1}, v^{k+1}) \leq V(u^k, v^k)$, for all $k \in \mathbb{N}$ and $V(u^{k+1}, v^{k+1}) = V(u^k, v^k)$ only if (u^k, v^k) is a fixed point of the algorithm. Finally we show that a fixed point of the algorithm is an optimizer of Problem 2.

From the convexity of C we have that:

$$C(d^{k+1}_{\mathbf{r},t}) \leq C(d^k_{\mathbf{r},t}) + C'(d^{k+1}_{\mathbf{r},t})(d^{k+1}_{\mathbf{r},t} - d^k_{\mathbf{r},t}).$$

Similarly, we have that:

$$\Phi_\ell(d_{\ell,t}^{k+1}) \leq \Phi_\ell(d_{\ell,t}^{k}) + \Phi'_\ell(d_{\ell,t}^{k+1})(d_{\ell,t}^{k+1} - d_{\ell,t}^{k}),$$

for all $\ell \in \mathcal{M}$. Then, it follows that:

$$\begin{aligned} V(u^{k+1},v^{k+1}) \leq & \sum_{t=1}^{T} \left(C(d_{\mathbf{r},t}^{k}) + C'(d_{\mathbf{r},t}^{k+1})(d_{\mathbf{r},t}^{k+1} - d_{\mathbf{r},t}^{k}) \right) \\ & + \sum_{t=1}^{T} \sum_{\ell \in \mathcal{M}} \kappa_\ell \left(\Phi_\ell(d_{\ell,t}^{k}) + \Phi'_\ell(d_{\ell,t}^{k+1})(d_{\ell,t}^{k+1} - d_{\ell,t}^{k}) \right). \end{aligned} \tag{3.16}$$

Using the fact that C' is Lipschitz continuous, it holds that:

$$C'(d_{\mathbf{r},t}^{k+1})(d_{\mathbf{r},t}^{k+1} - d_{\mathbf{r},t}^{k}) \leq C'(d_{\mathbf{r},t}^{k})(d_{\mathbf{r},t}^{k+1} - d_{\mathbf{r},t}^{k}) \quad + l_C |d_{\mathbf{r},t}^{k+1} - d_{\mathbf{r},t}^{k}|^2, \tag{3.17}$$

likewise, since Φ' is also Lipschitz continuous with Lipschitz constant equal to 2:

$$\kappa_\ell \Phi'_\ell(d_{\ell,t}^{k+1})(d_{\ell,t}^{k+1} - d_{\ell,t}^{k}) \leq \frac{1}{\eta} b_{\ell,t}^{k}(d_{\ell,t}^{k+1} - d_{\ell,t}^{k}) \quad + 2|d_{\ell,t}^{k+1} - d_{\ell,t}^{k}|^2. \tag{3.18}$$

In the last expression we replaced $\kappa_\ell \Phi'(d_{\ell,t}^{k})$ according to the definition of $b_{\ell,t}^{k}$, $\ell \in \mathcal{M}$. Also we use the expression in (3.2) and the fact $L_{j,t}$ does not depend on k, for all $j \in \mathcal{M}$, to obtain:

$$d_{\ell,t}^{k+1} - d_{\ell,t}^{k} = \sum_{i \in \mathrm{dN}(\ell)} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}). \tag{3.19}$$

for all $\ell \in \mathcal{M}$. Then, using (3.17) and (3.18), to upper bound (3.16), and then plug-

ging (3.19) into the result, we obtain:

$$
\begin{aligned}
V(u^{k+1},v^{k+1}) \leq\ & V(u^k, v^k) \\
&+ \sum_{t=1}^{T} l_C \left| \sum_{i\in\mathcal{N}} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \right|^2 \\
&+ \sum_{t=1}^{T} \sum_{i\in\mathcal{N}} \frac{1}{\eta} p_{\mathbf{r},t}^{k} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \\
&+ \sum_{t=1}^{T} \sum_{\ell\in\mathcal{M}\setminus\{\mathbf{r}\}} \sum_{i\in \mathrm{dN}(\ell)} \frac{1}{\eta} b_{\ell,t}^{k} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \\
&+ 2 \sum_{t=1}^{T} \sum_{\ell\in\mathcal{M}} \kappa_\ell \left| \sum_{i\in \mathrm{dN}(\ell)} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \right|^2 .
\end{aligned}
\tag{3.20}
$$

In (3.20) we have used again (3.2) to write the whole expression in terms of u and v, and we have also written $C'(d_{\mathbf{r},t}) + \kappa_{\mathbf{r}}\Phi'(d_{\mathbf{r},t})$ as $\frac{1}{\eta} p_{\mathbf{r},t}^{k}$, for all $t \in \{1,\dots,T\}$.

Further, from the definition of the dynamics, (u_i^{k+1}, v_i^{k+1}) fulfills the optimality condition of Lemma 3.6 for the local PEV problem. Then, we obtain:

$$
\sum_{t=1}^{T} \left(p_{\mathbf{r},t}^{k} + \sum_{\ell\in \mathrm{an}(i)\setminus\{\mathbf{r}\}} b_{\ell,t}^{k} \right) (u_{i,t}^{k} - v_{i,t}^{k} - u_{i,t}^{k+1} + v_{i,t}^{k+1}) - \|u_i^{k+1} - v_i^{k+1} - u_i^{k} + v_i^{k}\|^2 \geq 0, \tag{3.21}
$$

where we have written out the definition of $p_{\mathrm{pr}(i),t}^{k}$. Now, we sum both sides of (3.21) over all $i \in \mathcal{N}$ and then we use Lemma 3.7 to change the summation indices, to obtain:

$$
\begin{aligned}
&\sum_{t=1}^{T} \sum_{i\in\mathcal{N}} p_{\mathbf{r},t}^{k} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \\
&+ \sum_{t=1}^{T} \sum_{\ell\in\mathcal{M}\setminus\{\mathbf{r}\}} \sum_{i\in \mathrm{dN}(\ell)} b_{\ell,t}^{k} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \\
&\qquad\qquad \leq - \sum_{i\in\mathcal{N}} \|u_i^{k+1} - v_i^{k+1} - u_i^{k} + v_i^{k}\|^2 .
\end{aligned}
\tag{3.22}
$$

On the other hand, from Hölder's inequality we have that:

$$
\left| \sum_{i\in\mathcal{I}} (u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}) \right|^2 \leq |\mathcal{I}| \sum_{i\in\mathcal{I}} |u_{i,t}^{k+1} - v_{i,t}^{k+1} - u_{i,t}^{k} + v_{i,t}^{k}|^2, \tag{3.23}
$$

for any subset of $\mathcal{I} \subset \mathcal{V}$. Our next step is to use (3.23) to bound the second and last summands of (3.20), with $\mathcal{I} = \mathcal{N}$ and $\mathcal{I} = \mathrm{dN}(\ell)$ respectively, then use the bound from (3.22) on the third and fourth summands of (3.20), and finally apply Lemma 3.7 on the last summand of (3.20). This yields:

$$
\begin{aligned}
V(u^{k+1}, v^{k+1}) \leq\ & V(u^k, v^k) \\
& + \sum_{i \in \mathcal{N}} l_C N \|(u_i^{k+1} - v_i^{k+1} - u_i^k + v_i^k)\|^2 \\
& - \sum_{i \in \mathcal{N}} \frac{1}{\eta} \|u_i^{k+1} - v_i^{k+1} - u_i^k - v_i^k\|^2 \\
& + \sum_{i \in \mathcal{N}} \sum_{\ell \in \mathrm{an}(i)} |\,\mathrm{dN}(\ell)| 2\kappa_\ell \|(u_i^{k+1} - v_i^{k+1} - u_i^k + v_i^k)\|^2.
\end{aligned}
\tag{3.24}
$$

Note that in (3.24) we used the fact that $\sum_{t=1}^{T} |z_t|^2 = \|z\|^2$, for any vector $z \in \mathbb{R}^T$. Now, it is easy to see that:

$$
\begin{aligned}
\frac{1}{\eta} \|u_i^{k+1} - v_i^{k+1} - u_i^k - v_i^k\|^2 =\ & \\
& \frac{1}{\eta(1 + |\,\mathrm{an}(i)|)} \|u_i^{k+1} - v_i^{k+1} - u_i^k - v_i^k\|^2 \\
& + \sum_{\ell \in \mathrm{an}(i)} \frac{1}{\eta(1 + |\,\mathrm{an}(i)|)} \|u_i^{k+1} - v_i^{k+1} - u_i^k - v_i^k\|^2,
\end{aligned}
$$

for all $i \in \mathcal{N}$. This follows from the fact that the term inside the sum does not depend on the index of such sum. Then, replacing the expression above in (3.24), it follows:

$$
\begin{aligned}
V(u^{k+1}, v^{k+1}) \leq\ & V(u^k, v^k) \\
& + \sum_{i \in \mathcal{N}} \left(N l_C - \frac{1}{\eta(1 + |\,\mathrm{an}(i)|)} \right) \|(u_i^{k+1} - v_i^{k+1} - u_i^k + v_i^k)\|^2 \\
& + \sum_{i \in \mathcal{N}} \sum_{\ell \in \mathrm{an}(i)} \left(|\,\mathrm{dN}(\ell)| 2\kappa_\ell - \frac{1}{\eta(1 + |\,\mathrm{an}(i)|)} \right) \times \cdots \\
& \quad \|(u_i^{k+1} - v_i^{k+1} - u_i^k + v_i^k)\|^2.
\end{aligned}
\tag{3.25}
$$

This means that $V(u^{k+1}, v^{k+1}) \leq V(u^k, v^k)$, if:

$$\frac{1}{\eta(1 + |\operatorname{an}(i)|)} > 2\kappa_\ell |\operatorname{dN}(\ell)|, \quad \forall \ell \in \operatorname{an}(i), i \in \mathcal{N}$$
$$\frac{1}{\eta(1 + |\operatorname{an}(i)|)} > N l_C, \quad \forall i \in \mathcal{N}.$$

In fact, $V(u^{k+1}, v^{k+1}) < V(u^k, v^k)$ whenever $u_i^{k+1} - v_i^{k+1} \neq u_i^k - v_i^k$, then the V2G HIERARCHICAL algorithm converges to the set of points $\mathcal{S} \triangleq \{(u^{k+1}, v^{k+1}) \in \mathbb{R}^{2nT} \mid u_i^{k+1} - v_i^{k+1} = u_i^k - v_i^k,$ and u^{k+1}, v^{k+1} are given by Equation (3.14), $\forall i \in \mathcal{N}\}$. Finally, if we use the optimality condition in Lemma 3.6 for the local problem in (3.14), then we sum over all $i \in \mathcal{N}$ and next we use the fact that $u_i^{k+1} - v_i^{k+1} = u_i^k - v_i^k$, we recover the optimality condition for Problem 2, which implies that any point in $\mathcal{S}$ is an optimizer of Problem 2, completing the proof. □

3.5 Simulations and discussion

Our simulation scenario consists in the rooted tree shown in Figure 3.3, where the circles represent the PEVs and the squares represent the nodes in $\mathcal{M}$, with $\mathcal{V} = \{1, \ldots, 25\}$, $\mathcal{N} = \{1, \ldots, 20\}$, $\operatorname{pr}(i) = 21$, for $i \in \{1, \ldots, 5\}$, $\operatorname{pr}(i) = 22$, for $i \in \{6, \ldots, 10\}$, $\operatorname{pr}(i) = 23$, for $i \in \{11, \ldots, 15\}$, and $\operatorname{pr}(i) = 24$, for $i \in \{16, \ldots, 20\}$. The initial conditions, efficiency and battery capacities have been chosen to be different for all the PEVs. We establish bounds for the power to go through lines $N+1$ to $N+3$ and $N+2$ to $N+3$ 21 as: $P_{N+1}^{\max}(t) = 8.5$ and $P_{N+2}^{\max}(t) = 14.5$, respectively, for all $t \in \tau$. There is no bound for other lines in the distribution feeders. The function C is chosen to be $C(x) = x^2$. All non-PEV demand, as well as the initial conditions, efficiency and battery capacities can be found at `http://fausto.dynamic.ucsd.edu/andres`.

Figure 3.4 shows the optimal aggregate demand for all the PEVs in $\mathcal{N}$, for a centralized solution of the exact problem, i.e., without penalty functions (red curve), and the aggregate demand given by the V2G HIERARCHICAL algorithm after 3000 iterations (black curve). It can be observed that the hierarchical solution almost matches the aggregate given by the centralized benchmark. Figure (3.5) show the aggregate PEV and non-PEV demand for the centralized solution (red) and for the hierarchical solution

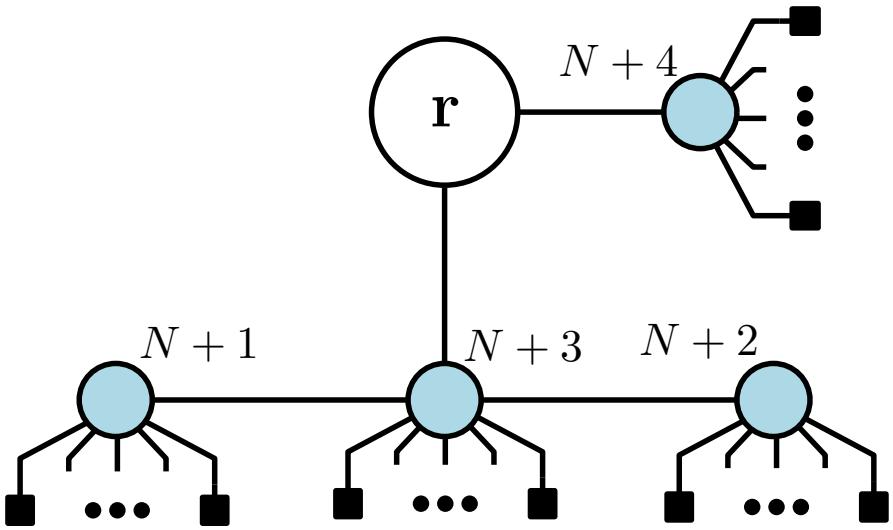

Figure 3.3: Topology for the simulation scenario. Black circles denote buses in distribution feeders, while squares denote PEVs.

(black). In addition, we show in blue the non-PEV demand. It can be seen that the optimal solution is peak-shaving and valley-filling, since PEVs tend to provide energy between 12:00 and 17:00, and they charge between 22:00 and 8:00. Figure 3.6 shows the demand curve for the node $N + 1$, with the corresponding transmission line bound (green) for both centralized (red) and hierarchical (black) cases. It can be seen that the solution of the V2G HIERARCHICAL algorithm satisfies the constraint for all time. Figure 3.7 shows the same features for the demand at node $N + 2$. In order to evaluate the impact of the number of PEVs on the V2G HIERARCHICAL algorithm performance, we have generated four scenarios with 20, 80, 140, and 200 PEVs respectively. In all cases, battery sizes as well as initial conditions and deadlines have been chosen to be random, but with comparable sizes. The distribution feeder configuration is the same for all scenarios, and corresponds to that shown in Figure 3.3. The top plot of Figure 3.8 shows the evolution of $J(u^k, v^k) - J(u^\star, v^\star)$ vs iterations of the V2G HIERARCHICAL algorithm. On the bottom plot, we show the optimal aggregate demand for each scenario. We can see on the top plot that the more PEVs we incorporate in the scenario, and consequently the larger demand that they generate, the higher the value of $J(u^k, v^k) - J(u^\star, v^\star)$. However, for all cases, we have that after 3000 iterations, $J(u^k, v^k) - J(u^\star, v^\star)$ lies within 4% of the optimal cost for each scenario.

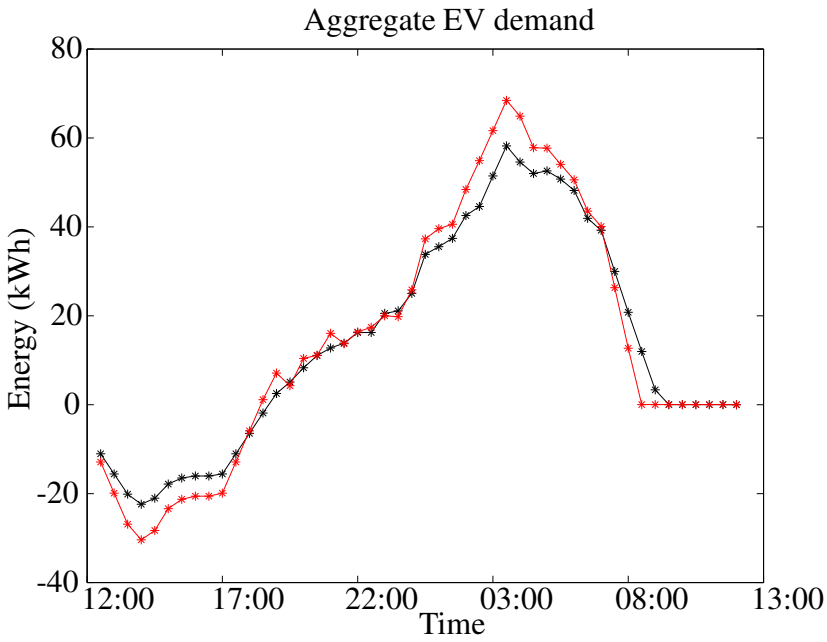

Figure 3.4: The optimal aggregate EV demand is shown in red. The aggregate EV demand from the V2G HIERARCHICAL algorithm is shown in black.

3.6 Auxiliary results for this chapter

The following result has been taken from [30].

Lemma 3.6. *For the feasible convex optimization problem:*

$$\begin{aligned} &\textit{minimize: } f(x), \\ &\textit{subject to:} \\ &x \in X, \end{aligned}$$

with $x \in \mathbb{R}^n$, $x^\star$ is an optimizer if and only if:

$$\nabla f(x^\star)^\top (x - x^\star) \geq 0,$$

for all $x \in X$.

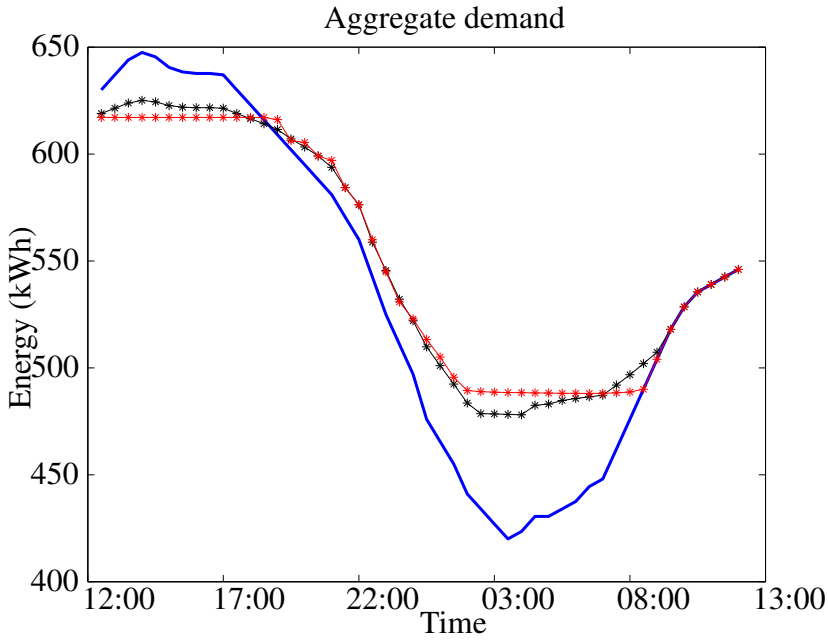

Figure 3.5: The non-PEV demand is shown in blue. The aggregate optimal demand is shown in red. The aggregate demand obtained through the V2G HIERARCHICAL algorithm is shown in black.

Lemma 3.7 (Counting of edges in $\mathcal{T}$)**.** *The equality:*

$$\sum_{\ell \in \mathcal{M}\setminus\{\mathbf{r}\}} \sum_{i \in \mathrm{dN}(\ell)} A_i B_\ell = \sum_{i \in \mathcal{N}} \sum_{\ell \in \mathrm{an}(i)\setminus\{\mathbf{r}\}} A_i B_\ell,$$

holds for any terms A_i, B_ℓ.

Proof. It is easy to show that each of the summands accounts for one path between each PEV and each of its ancestors. □

3.7 Summary

We present a hierarchical protocol for a vehicle-to-grid (V2G) system in which a fleet of plug-in electric vehicles must coordinate their charging/discharging strategies to minimize a cost function consisting in the price of the total energy provided by the utility during a finite discrete-time horizon.The power flow leaves the transmission side of the power grid and enters the distribution side. It is modeled as a rooted tree, where

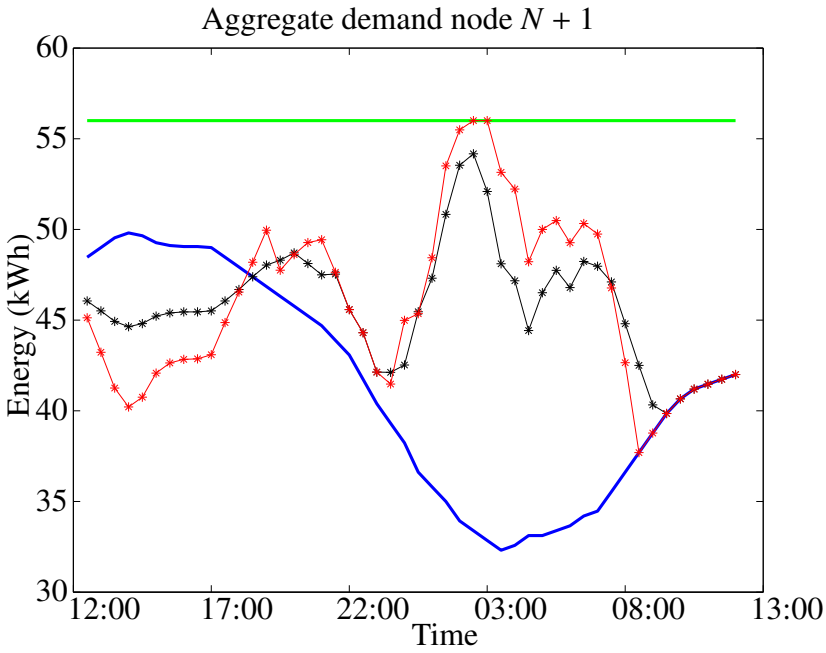

Figure 3.6: The non-PEV demand is shown in blue. The optimal aggregate demand is shown in red. The aggregate demand obtained through the V2G HIERARCHICAL algorithm is shown in black. The green line shows the upper bound for the power capacity.

nodes represent buses in distribution feeders, as well as PEVs. Our model also accounts for power capacity constraints in the distribution lines. In order to account for these constraints, we use penalty functions. We characterize the size of the constraint violation in terms of the penalty parameters and the parameters of the problem. This characterization provides a design methodology for the choice of the penalization parameters in terms of a desired performance. The presented V2G HIERARCHICAL algorithm does not require communication between PEVs, and the coordination signal is transmitted from the utility down the tree network, while being modified at each non-PEV node, until it reaches the PEVs. Then, each PEV uses it to iterate over its charging/discharging profile. We show that the V2G HIERARCHICAL algorithm converges to the optimizer of the cost function given the network constraints. Simulations show the system behavior for a particular testbed.

As a future direction, we aim to address the power constraints in the congestible elements using non-differentiable penalty functions that allow exact solutions of the original problem, but require a subgradient-based algorithm for the solution, with the

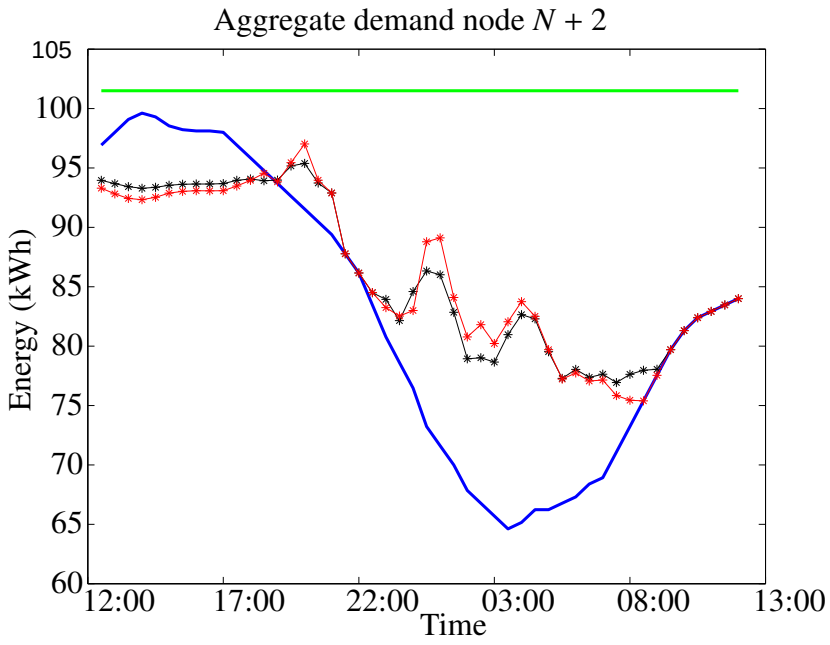

Figure 3.7: The non-PEV demand is shown in blue. The optimal aggregate demand is shown in red. The aggregate demand obtained through the V2G HIERARCHICAL algorithm is shown in black. The green line determines upper bound for the power capacity.

ensuing complications in the analysis.

Acknowledgments

Parts of this chapter have been published in the following works:

- A. Cortés and S. Martínez, "A Hierarchical Demand-Response Algorithm for Vehicle-to-Grid Integration," submitted to the IEEE Transactions on Control Systems Technology (2015).

- A. Cortés and S. Martínez, "A Hierarchical Demand-Response Algorithm for Optimal Vehicle-to-Grid Coordination," in the proceedings of the European Control Conference (2015).

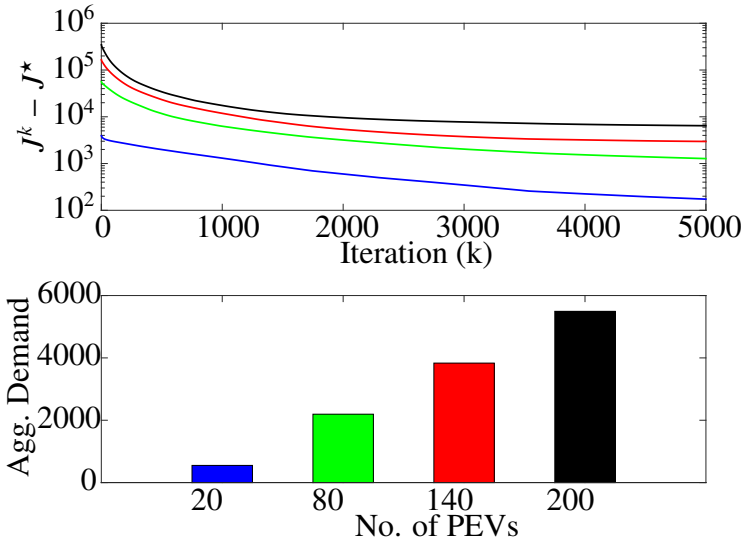

Figure 3.8: Top figure: Evolution of $J(u^k, v^k) - J(u^\star, v^\star)$ vs number of iterations, for four scenarios with different No. of PEVs. Bottom figure: Optimal aggregate demand for each scenario.

Chapter 4

Hierarchical demand response with on/off Loads

Demand response, by which a virtual reserve capacity can be created, is an idea under investigation that can help integrate renewables into the power grid. By modifying the state of flexible loads, the rapidly changing outcome of renewable generation can be matched by the demand and achieve power balancing required for stability. However, demand response can severely affect power-grid users, not only due to a potential power shortage and reduced quality of service, but also because they may require the disclosure of private information. Therefore, the development of smarter load control strategies is necessary to enable this technology in a beneficial way.

Model predictive control (MPC), which can provide suboptimal control strategies for modeled systems subject to uncertainty, is a promising tool in this regard. However, its practical implementation requires adaptation on two fronts. Firstly, its required computational time needs to be lowered. This becomes specially critical for flexible loads, most of which are on/off and lead to hard problems. Secondly, users' privacy must be preserved. This requires sufficiently fast decentralized algorithms which do not slow down convergence. Motivated by this, we propose here a decentralized demand response strategy, which is embedded in an MPC framework and which accounts for thermal and on/off loads.

Demand response is currently the How to control loads for demand response with the least impact on the users has also been studied in a significant number of works

under different sets of assumptions. In [1], the authors present decentralized algorithms for the dispatch of Distributed Energy Resources (DERs) and demand response. It computes the DER controls given a signal provided by an aggregator and by interacting with neighboring loads. In [2], the same authors address the DER control problem, by providing a decentralized solution of an optimization problem to match the grid balance objective. This method does not require an aggregator, but agents only communicate with each other to solve the optimization problem. In [12], the authors consider a decentralized optimization approach for electric vehicle charging coordination, under usage constraints. This leverages the storage capacity and flexibility of electric vehicle requirements to optimally use the grid generation resources. The use of the inherent storage capacity of some power loads, e.g., thermal loads, has been presented in [32, 14, 33]. In these works the authors use a centralized model predictive control (MPC) formulation to take into account the storage capacity of some loads and forecast information available at each time, to compute a control that holds the demand response objectives. None of the aforementioned approaches include on/off loads. In [34] the authors introduce a centralized MPC approach for thermal on/off loads, but they simply define a convex optimization problem and generate a on/off control using pulse width modulation (PWM). In our framework, doing so may lead to violating maximum power constraints. The paper [35] does consider on/off loads in the introduced setting, however, it is a centralized framework for a single household management. In [36], the load control problem is formulated for on/off loads, with the objective of minimizing the power generation cost. However, the authors do not explicitly address the integer constraints due to the on/off loads in their proposed solution. In [37], not only do the authors present a framework that considers on/off loads, but also they introduce an agile approach for dispatching those loads. It is shown through simulations that the performance is satisfactory. However, their agile dispatch is a centralized process in which an aggregator directly controls the loads associated to it. In order to solve problems with integer constraints, the job-scheduling literature provides some approaches. One of them is based on Lagrangian relaxation [38], in which the coupling constraint is relaxed using a Lagrange multiplier, and each agent solves local mixed-integer programs parameterized by the estimate of the optimal Lagrange multiplier, which is computed by the aggregator. This is

a slow iterative process, and convergence to the optimal solution is not guaranteed. The auction-based approach proposed in [39, 40] guarantees convergence to the optimizer to the problem, but also requires solutions to smaller mixed-integer programs. The intense computation effort required for these algorithms, makes them prohibitive to most real applications.

In this chapter, we present a decentralized load control approach that explicitly takes into account the on/off nature of available loads, in order to fulfill a demand response event (DRE). This formulation also accounts for the energy storage capacity exhibited by thermal loads. To this end, we use a thermal model that includes outside temperature as a disturbance. The problem is formulated as a mixed-integer program, with the objective of minimizing the effect of a demand response event (DRE) on the users' comfort. We propose an algorithm that provides a feasible solution with a reasonably good performance, using only computationally tractable methods. The algorithm solves a convex relaxation of the original problem, and uses the relaxed control input for all on/off loads as a measure of need for power. After this, the resources are assigned using a greedy approach that provides power to the loads that need it the most until the maximum available power is allocated. Both, the convex optimization and the on/off load assignment are carried out in a decentralized manner. The algorithms proposed to this end are guaranteed to converge to the solution of the corresponding centralized problem. Since the load models are subject to uncertainty due to forecast, e.g., outside temperature forecast, we present an MPC implementation of this algorithm, to mitigate the impact such uncertainty. Finally, we use a set of simulation cases to show the algorithm performance under different operating conditions. These simulations illustrate the cost of a suboptimal solution provided by our algorithm vs a lower bound on the optimal solution, as well as a comparison between the open-loop suboptimal solution vs the MPC implementation.

4.1 Management of demand response events

Demand-response events (DRE) result into the shaping of flexible power demand over a time horizon to provide different ancillary services to the power grid. We

associate a DRE with a coordinated action of a large amount of power loads that modify their power consumption during a certain time lapse $\mathcal{T}$, in order to maintain the generation/demand balance in the power grid. The amount of power that the DRE must provide/withdraw from the grid is generally established from a transaction in an energy market by a utility [41].

In order to implement this, we introduce a DRE manager, which is an entity in charge of a large group of buildings with certain flexibility in their electric loads. The DRE manager aims to drive all loads into satisfying the DRE requirements, while minimizing its impact on the users' comfort.

In this particular study, we consider thermal loads, such as air conditioners and heaters; memoryless loads, such as light bulbs; and non-flexible loads, that must be invariably active (or inactive) at certain times of the day. Moreover, most of these loads are on/off loads.

The control strategy is designed to take advantage of the inherent energy storage capacity of thermal loads, and also of the prior knowledge of variables such as temperature or natural illumination, from a previously determined forecast process.

Another major interest in the computation of load control is the users' privacy. In general, users may not want to share their comfort model with the DRE manager. This is why a control strategy that can be computed in a decentralized way, is a priority in the present work.

4.1.1 Modeling a demand response event

Let us consider a DRE manager in charge of a set $I \triangleq \{1, \dots, N\}$ of buildings. Each building $i \in I$ contains five different types of loads. Let $\mathrm{Th}_{\mathrm{on/off}}(i)$ be the set of thermal on/off loads, $\mathrm{Th}_{\mathrm{curt}}(i)$ be the set of thermal curtailable loads, $\mathrm{L}_{\mathrm{on/off}}(i)$ be the set of memoryless loads, and finally, let $\mathrm{L}_{\mathrm{curt}}(i)$ be the set of memoryless curtailable loads for building $i \in I$. We denote by $L(i) \triangleq \mathrm{Th}_{\mathrm{on/off}}(i) \cup \mathrm{Th}_{\mathrm{curt}}(i) \cup \mathrm{L}_{\mathrm{on/off}}(i) \cup \mathrm{L}_{\mathrm{curt}}(i)$ as the set of flexible loads in $i \in I$.

A DRE time lapse $\mathcal{T}$ is divided into T time slots with duration $\Delta t = \mathcal{T}/T$; we let τ denote the sequence $\tau \triangleq \{0, \dots, T-1\}$ all discrete slots associated with it.

The power consumption of each flexible load $j \in L(i)$, $i \in I$, is denoted by $u_{ij}(t)$,

where $t \in \tau$ is a discrete time instant. Since there is no feasible action for fixed loads in the buildings, we characterize them by a value $P_{\text{fix}}(t)$.

A thermal load $j \in \mathbf{Th}(i)$ is modeled by a discrete-time SISO linear system as follows:

$$\begin{aligned} x_{ij}(t+1) &= A_{ij}x_{ij}(t) + B^1_{ij}u_{ij}(t) + B^2_{ij}T^a_{ij}(t), \\ T_{ij}(t) &= C_{ij}x_{ij}(t), \\ x_{ij}(0) &= x^0_{ij}, \end{aligned} \tag{4.1}$$

where $x_{ij}(t)$ is the system state, $T_{ij}(t)$ is the temperature inside the room corresponding to the thermal load j, and $T^a_{ij}(t)$ is the outdoors temperature for the load at time $t \in \tau$. The vector $x_{ij}(0)$ represents the state at the beginning of the DRE. This discrete-time model may come from either an identification process using input-output data, or from the discretization of a continuous-time thermal model (e.g., an RC thermal model [42]) with time step Δt. Each load in a building has a discomfort value that is associated to its power input. For instance, the comfort value of a thermal load $j \in \mathbf{Th}(i)$ is given by:

$$\begin{aligned} f_{ij}(t) \triangleq & \kappa_{ij}(\max\{0, T_{ij}(t+1) - T^{\max}_{ij}\} \\ & + \max\{0, T^{\min}_{ij} - T_{ij}(t+1)\}), \end{aligned}$$

where $[T^{\min}_{ij}, T^{\max}_{ij}]$ is the temperature interval in which the users of the j^{th} thermal load in the i^{th} building are most comfortable, and κ_{ij} the users' tolerance to discomfort. Note the time shift in the temperature value, which is consistent with the fact that $T_{ij}(t+1)$ directly depends on $u_{ij}(t)$ for all $j \in \mathbf{Th}(i)$, $i \in I$. For memoryless loads we introduce a (generally nonnegative) discomfort function defined as:

$$f_{ij}(t) = \alpha_{ij}(t)u_{ij}(t) + \beta_{ij}(t),$$

where $\alpha_{ij}(t), \beta_{ij}(t) \in \mathbb{R}$, for all $t \in \tau$.

Remark 4.1. *The parameters $\alpha_{ij}(t), \beta_{ij}(t) \in \mathbb{R}$ have been chosen to be time-varying, to model events such as the change of natural illumination inside a room during the DRE. This event may change the impact of load j in its user's comfort.* ◇

The DRE itself is modeled as an upper (lower) bound on the amount of energy the whole set of buildings can use. This information is provided by an Independent System Operator to the DRE manager, and is based on the load forecast on the power grid. For simplicity, we represent this bound by the constraint:

$$\sum_{i=1}^{N} \sum_{j \in L(i)} \hat{u}_{ij}(t) \leq P_{\max}(t), \quad \forall t \in \tau,$$

for all $t \in \tau$.

Remark 4.2. *In this study, we only consider DREs where a positive power compensation is required, i.e., the demand must be shortened. In the opposite case, the entire procedure can be adapted analogously.* ◇

4.1.2 Optimal DRE control problem formulation

Here, we formulate an optimal control problem that results into the minimization of the general discomfort among the users of all the loads during the DRE. An algorithm to solve this problem is proposed in Section 4.2.

The DRE manager will aim to solve the following optimization problem:

$$\mathcal{P}1 : \text{minimize}_u \sum_{t=0}^{T-1} \sum_{i=1}^{N} \sum_{j \in L(i)} f_{ij}(t) \tag{4.2a}$$

subject to:

$$\text{Equation (4.1)}, \quad \forall j \in \mathbf{Th}(i), \forall i \in I, \forall t \in \tau, \tag{4.2b}$$

$$u_{ij}(t) \in [0, u_{ij}^{\max}], \quad \forall j \in \mathbf{Curt}(i), \forall i \in I, \forall t \in \tau, \tag{4.2c}$$

$$u_{ij}(t) \in \{0, u_{ij}^{\text{on}}\}, \quad \forall j \in \mathbf{Onoff}(i), \forall i \in I, \forall t \in \tau, \tag{4.2d}$$

$$\sum_{i=1}^{N} \sum_{j \in L(i)} u_{ij}(t) \leq P_{\max}(t), \quad \forall t \in \tau. \tag{4.2e}$$

Notice that Notice that the state of the thermal loads at time $t = 0$ corresponds to the system state immediately prior to the beginning of the DRE.

The previous problem presents a convex cost function, however, constraints described by (4.2d) are binary, hence, the problem becomes is a mixed integer program.

Since mixed integer programs are NP-complete, there is no algorithm that can solve it in polynomial time, and the solution time grows exponentially as the amount of integer variables grows.

Remark 4.3. *Notice that using PWM driven by the solution to the convex relaxation to the problem to generate an on/off load control may lead to the constraint* (4.2e) *to be violated at some time instants, since there is no synchronization among all loads.* ◇

4.2 Solution approach

Privacy is a major objective for our load management solution. Thus, a centralized approach in which the DRE manager knows the model of all loads and discomfort functions of users may not be acceptable. Moreover, for large amounts of buildings or loads, the problem to be solved in a centralized way could grow too large to be manageable. Hence, the solution approach we consider must be susceptible of being executed in a decentralized or parallel manner.

Our solution approach consists of two steps: convex optimization and thresholding. The thresholding step is devoted to use the result from the convex relaxation of $\mathcal{P}1(1)$, that is not feasible to the problem and generate a feasible solution to it, without deteriorating the service provided to the users.

We describe the overall execution of these steps in the following, and leave the specific details for Subsection 4.2.1 and 4.2.2. Then, we propose a decentralized implementation of these steps in Section 4.3.

We consider the problem $\mathcal{P}2(0)$; see Subsection 4.2.1, which is a convex relaxation of the problem $\mathcal{P}1$ in (4.2), with the only difference that we replace the constraints (4.2d) by $u_{ij}(t) \in [0, u_{ij}^{\text{on}}]$, for all on/off loads. Let $v^{\star,0}$ be an optimal solution of this relaxed problem. If we compute $y_{ij}^{\star,0}(t) \triangleq v_{ij}^{\star,0}(t)/u_{ij}^{\text{on}} \in [0, 1]$, the result can be interpreted as the *level of urgency* that load j in building i has at time $t \in \tau$. For the sake of clarity, consider the time $t = 0$. The value of $y_{ij}(0)$ for all on/off loads can be used to establish the relative priority of these loads, and thus determine what loads should be on, based on the limited available power resources.

A threshold variable $\theta(0) \in [0, 1]$ is introduced to decide on the state of on/off

loads. Thus, all on/off loads for which $y_{ij}^{\star,0}(0) \in (0,1)$, $y_{ij}^{\star,0}(0) < \theta(0)$, must turn off, while those for which $y_{ij}^{\star,0}(0) \geq \theta(0)$ must turn on, i.e.:

$$\hat{u}_{ij}(0) = \begin{cases} u_{ij}^{\text{on}} & \text{if } y_{ij}^{\star,0}(0) \geq \theta(0) \\ 0 & \text{otherwise,} \end{cases}$$

for all $j \in \mathbf{Onoff}(i)$, $i \in I$, where $\hat{u}_{ij}(0)$ is defined as the control input to load $j \in L(i)$, $i \in I$. The threshold variable can always be chosen in such a way that after turning on and off the corresponding loads, the constraint on the maximum allowed demand is satisfied (see Lemma 4.1). Once the value for the on/off controllers $\hat{u}_{ij}(0)$ for all loads have been chosen, we proceed to solve the problem $\mathcal{P}2(0)$ as described above, but including the constraints:

$$u_{ij}(0) = \hat{u}_{ij}(0), \qquad \forall j \in \mathbf{Onoff}(i), \forall i \in I.$$

This computation will perform two tasks: i) to refine the computed values of $\hat{u}_{ij}$ for all loads $j \in \mathbf{Curt}(i)$, improving the use of resources at time $t = 0$, and ii) to provide a computation of the level of urgency for all on/off loads at time $t = 1$. Then, a threshold can be computed for $t = 1$, leading to the control values for all on/off loads at such time.

In this way, we increasingly fix the values of all on/off loads for each time $t \in \{1, \ldots, T-1\}$, given the previously computed control values control values for such loads at all times $q \in \{0, \ldots, t-1\}$.

4.2.1 Step 1: convex optimization

More precisely, the proposed relaxation is defined next.

$$\mathcal{P}2(t) : \text{minimize}_u \sum_{q=0}^{T-1} \sum_{i=1}^{N} \sum_{j \in L(i)} f_{ij}(Q) \tag{4.3a}$$

subject to:

$$\text{Equation (4.1)}, \quad \forall j \in \mathbf{Th}(i), \forall i \in I, \forall q \in \tau, \tag{4.3b}$$

$$u_{ij}(q) \in [0, u_{ij}^{\max}], \forall j \in \mathbf{Curt}(i), \forall i \in I, \forall q \in \{0, \ldots, T-1\}, \tag{4.3c}$$

$$u_{ij}(q) \in [0, u_{ij}^{\text{on}}], \forall j \in \mathbf{Onoff}(i), \forall i \in I, \forall q \in \{t, \ldots, T-1\}, \tag{4.3d}$$

$$u_{ij}(q) = \hat{u}_{ij}(q), \forall j \in \mathbf{Onoff}(,) \forall i \in I, \forall q \in \{0, \ldots, t-1\}, \tag{4.3e}$$

$$\sum_{i=1}^{N} \sum_{j=1}^{L(i)} u_{ij}(q) \leq P_{\max}(q), \quad \forall q \in \{t, \ldots, T-1\}, \tag{4.3f}$$

for all $t \in \{1, \ldots, T-1\}$. Note that the problem $\mathcal{P}2(t)$ simply consists in relaxing the integer constraints for all on/off loads for all times $q \in \{t, \ldots, T-1\}$, fixing the previous computed control values for all loads at times $q \in \{0, \ldots, t-1\}$. Define $v^{\star,t} \triangleq \{v_{ij}^{\star,t}(q)\}_{j \in L(i), i \in I, q \in \tau}$ as the solution of the problem $\mathcal{P}2(t)$, for all $t \in \{1, \ldots, T-1\}$. This will be used to perform the thresholding procedure for the computed control values at time $t \in \tau$.

The following result establishes the existence of a suitable threshold to compute a feasible solution for all times $t \in \tau$.

Proposition 4.1. *Let $v^{\star,t}$ be a solution to the convex relaxation $\mathcal{P}2(t)$, for all $t \in \{0, \ldots, T-1\}$. Then, there exists a $\theta(t) \in [0, 1]$ such that if:*

$$\hat{u}_{ij}(t) = \begin{cases} u_{ij}^{\text{on}} & \text{if } y_{ij}^{\star,t}(t) \geq \theta(t) \\ 0 & \text{otherwise}, \end{cases} \tag{4.4}$$

the constraint $\sum_{i=1}^{N} \sum_{j \in L(i)} \hat{u}_{ij}(t) \leq P_{\max}(t)$ holds.

Proof. The result follows immediately by noting that if $\theta(t) = 1$, only those loads for

which $y_{ij}^{\star,t} = 1$ will be turned on. Then, clearly, for all $i \in I$:

$$\begin{aligned} P_{\max}(t) &\geq \sum_{i=1}^{N} \sum_{j \in L(i)} v_{ij}^{\star,t}(t) \\ &= \sum_{i=1}^{N} \left(\sum_{\substack{j \in \mathbf{Onoff}(i) \\ y_{ij}^{\star,t}(t)=1}} u_{ij}^{\text{on}} + \sum_{j \in \mathbf{Curt}(i)} \hat{u}_{ij}(t) \right). \end{aligned}$$

□

4.2.2 Step 2: threshold Computation

So far, we have explained how the thresholding procedure can be used to satisfy the constraint on the maximum available power. Now, we establish a way of choosing $\theta(t)$ so that the overall cost is minimized.

In general, we consider that the maximum available power is a scarce resource that must be split among all loads. Then, we propose a *greedy* strategy in which the "optimal" threshold is chosen in the following way.

Definition 4.1. *An optimal threshold $\theta^{\star}(t)$ is one such that if we choose $\theta^{\text{new}}(t) \triangleq \max\{y_{ij}^{\star,t}(t) > \theta^{\star}(t), \forall y_{ij}^{\star,t}(t) \mid j \in \textbf{\textit{Onoff}}(i), i \in I\}$, and the thresholding process is carried out using $\theta^{\text{new}}(t)$, then the solution does not satisfy $\sum_{i=1}^{N} \sum_{j \in L(i)} \hat{u}_{ij}(t) \leq P_{\max}(t)$.* ◇

Algorithm 2 Approximation algorithm

for $t = 0$ to $T - 1$ **do**

- Compute $v^{\star,t}$ as an optimizer of $\mathcal{P}2(t)$.
- Compute threshold $\theta(t)$ and $\hat{u}_{ij}(t)$, for all $j \in \mathbf{Onoff}(i)$, $i \in I$, according to Equation (4.4).

end for

Compute $\hat{u}_{ij}(t)$ for all $j \in \mathbf{Curt}(i)$, $i \in I$, $t \in \tau$, by solving $\mathcal{P}1$ with the constraint $u_{ij}(q) = \hat{u}_{ij}(q)$, $q \in \{0, \ldots, T-1\}$, $j \in \mathbf{Onoff}(i)$, $i \in I$.

Remark 4.4. *The following drawback may affect the solution performance. A large amount of on/off loads—possibly all of them—can result in an identical value of $y_{ij}^{\star,t}(t) =$*

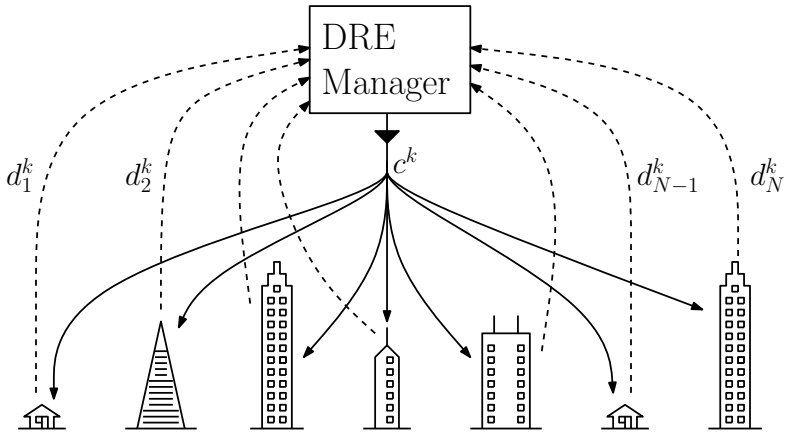

Figure 4.1: Communication architecture for the decentralized implementation of the algorithm. See Subsections 4.3.1 4.3.2 for the definition of d_i^k and c^k in each case.

$\bar{y}(t)$, for some $t \in \tau$. This means that if all loads are on and the solution is infeasible, the optimal threshold $\theta^\star(t)$ as introduced above is such that $\theta^\star(t) < \bar{y}(t)$, leading to all loads to be off at time t. A simple way to break the symmetry is the introduction of a small perturbation such that the value $y_{ij}^{\star,t}(t) = v_{ij}^{\star,t}(t)/u_{ij}^{\text{on}} + \epsilon_{ij}(t)$, where $\epsilon_{ij}(t)$ is a random variable with uniform distribution and very low variance. With this disturbance, the probability that two loads have exactly the same value $y_{ij}^{\star,t}(t)$ is zero, and the thresholding approach can be carried out without significant modification. Nevertheless, it is very unlikely that, under a large number of loads, a scenario like the above occurs in practice. ◇

4.3 Decentralized solution

Our solution approach has been structured in a way that all computations are amenable to decentralization. This means that the calculation of the control inputs for all loads can be made by the agents associated to each building $i \in I$ as we describe next.

Recall that our solution approach consists of two separate steps, namely, i) solution of a convex relaxation and ii) thresholding. Then, we use two algorithms, one for each step, that are executed in an iterative fashion, via an information exchange between the building agents and the DRE manager. Figure 4.1 shows the communication structure and the information exchange of this network. At each algorithm, each agent $i \in I$ provides the DRE manager a usage signal d_i^k, while the DRE manager returns a

coordination signal c^k which depends on d_i^k. In order to respect users' privacy, d_i^k does not include comfort parameters or load models.

The definition of d_i^k and c_i^k will be introduced after the explanation of each step.

4.3.1 Step 1: Decentralized convex optimization

The Dual Decomposition method [31], can be used to provide a decentralized optimization algorithm for a given problem type as $\mathcal{P}2(t)$ and a communication network as in 4.1. Even though convergence guarantees may be established for this method, it results in very slow convergence rates in general.

In order to overcome such drawback, we employ an augmented-Lagrangian methodology from [43] which is adapted to our setting. The Distributed Augmented Lagrangian Method (**ADAL**) of [43] is a provable-correct algorithm under the assumption of coupling equality constraints. In order to apply this algorithm with the same guarantees, we modify $\mathcal{P}2(t)$ as follows:

$$\mathcal{O}1 \text{ :minimize} \sum_{i=1}^{N+1} f_i(z_i)$$

subject to:

$$z_i \in Z_i, \forall i \in I \cup \{N+1\}, \tag{4.5a}$$

$$\sum_{i=1}^{N+1} F_i z_i = b, \tag{4.5b}$$

where each entry of the vector z_i corresponds to a decision variable associated to building i, either $\{x_{ij}(q), T_{ij}(q)\}$ for some load $j \in \mathbf{Th}(i)$, $q \in \tau$, $u_{ij}(q)$ for some load $j \in \mathbf{Curt}(i)$, $q \in \tau$, or $u_{ij}(q)$ for some load $j \in \mathbf{Onoff}(i)$, $q \in \{t, \ldots, T-1\}$, for all $i \in I$. Likewise, $f_i(z_i) \triangleq \sum_{j \in L(i)} \sum_{q \in \tau} f_{ij}(q)$, with $f_{ij}(t)$ as defined in Subsection 4.1.1, for all $i \in I$. After this, $Z_i \triangleq \{z_i \mid \text{Local constraints in } \mathcal{P}2(t) \text{ hold}\}$, for all $i \in I$, where Local constraints in $\mathcal{P}2(t)$ are given by Equations (4.3b) through (4.3e). In addition, $F_i z_i \in \mathbb{R}^T_{\geq 0}$, corresponds to the vector with components $(F_i z_i)_\ell \triangleq \sum_{j \in L(i)} u_{ij}(q)$, where $(F_i z_i)_\ell$ is the ℓ^{th} component of $F_i z_i$. This implies that $F_i z_i$ is the aggregate demand profile of building i, given by the relaxed problem $\mathcal{P}2(t)$. By the non-negativity of $u_{ij}(q)$ for all j, q and $i \in I$, it is

evident that $F_i z_i \succeq 0$, where the symbols $\succeq, \preceq$ indicate component-wise inequalities. Also, $b \in \mathbb{R}^T_{\geq 0}$ is such that the ℓ^{th} component of b corresponds to $P_{\max}(\ell - 1)$. Notice that with these definitions, the inequality constraints in problem $\mathcal{P}2(t)$ correspond to $\sum_{i=1}^{N} F_i z_i \preceq b$. The new variable $z_{N+1} \in \mathbb{R}^T_{\geq 0}$ is introduced simply as a slack variable to turn the inequality coupling constraints of $\mathcal{P}2(t)$ into equality constraints. Then, we define $f_{N+1}(z_{N+1}) = 0$, $F_{N+1} z_{N+1} = z_{N+1}$. Since $F_i z_i \succeq 0$ for all $i \in I$, it holds that $z_{N+1} \in Z_{N+1} \triangleq \{y \in \mathbb{R}^T \mid 0 \preceq y \preceq b\}$. Notice that the problem $\mathcal{P}1$ can be formulated as described above for $\mathcal{P}2(t)$. By the **ADAL** algorithm, agents and DRE manager execute the following iteration:

$$\begin{aligned} \hat{z}_i^k &= \operatorname{argmin}_{z_i \in Z_i} \mathcal{L}_i(z_i, \lambda^k) + \frac{\rho}{2} \|F_i z_i + \xi^k - F_i z_i^k\|^2, \\ z_i^{k+1} &= (1 - \gamma) z_i^k + \gamma \hat{z}_i^k, \end{aligned} \tag{4.6}$$

for all $i \in I \cup \{N+1\}$, where $\mathcal{L}_i(z_i, \lambda) \triangleq f_i(z_i) + \lambda^\top F_i z_i$ and $\xi^k \triangleq \sum_{i=1}^{N+1} F_i z_i^k - b$, and:

$$\lambda^{k+1} = \lambda^k + \rho \gamma \xi^{k+1}, \tag{4.7}$$

where $\lambda^k \in \mathbb{R}^T$ is a Lagrange multiplier estimate.

Theorem 4.1. *The* **ADAL** *algorithm in* (4.6) - (4.7) *converges to an optimal solution of* $\mathcal{O}1$ *for* $0 < \gamma < 1/(N+1)$.

Proof. From [43], it follows that the **ADAL** algorithm converges to the optimal solution of the problem $\mathcal{O}1$ if: i) the problem satisfies the Slater's condition, ii) Z_i is compact for $i \in I \cup \{N+1\}$, and iii) the parameter γ is positive and smaller than the inverse of the maximum number of agents involved in each coupling equality constraint. Notice that for the problem $\mathcal{O}1$, since $u_{ij}(q)$ is bounded for all $j \in L(i)$, i$\in I$, $q \in \tau$, Z_i is bounded, for all $i \in I$. Since the feasible sets are closed, Z_i is compact, for all $i \in I$. By definition of $Z_{N+1} = \{y \in \mathbb{R}^T \mid 0 \preceq y \preceq b\}$, it is compact. Finally, since the problem is convex and all constraints are affine, the problem satisfies the Slater's condition, and the result follows. □

Now, let us describe the decentralized implementation of the previous algorithm

with the communication structure of Figure 4.1. At each iteration k, the values ξ^k, z^k_{N+1}, and λ^k are first computed by the DRE manager, which submits the signal $c^k = (\xi^k, \lambda^k)$ to buildings. After this, each building computes z^k_i and $d^k_i = F_i z^k_i$, for all $i \in I$. Then, d^k_i is sent to the DRE manager for the next iteration. Recall that $F_i z^k_i$ corresponds to the aggregate demand profile for the i^{th} building, for $i \in I$, which means that the users' privacy is preserved. The iteration is run until the constraint violation fulfills certain tolerance value.

4.3.2 Step 2: Decentralized thresholding

In order to compute the threshold θ for time $t \in \tau$ in a decentralized manner, we propose an iterative process as follows: first, the DRE manager sends an estimate of the optimal threshold for $\theta(t)$; then, based on the estimate and the solution of the relaxed optimization problem, the building agents compute the control inputs for all their on/off loads. Next, all buildings submit their aggregate load to the DRE manager, who updates the threshold estimate based on the latest information. The updating rule for the threshold estimate is given by:

$$x^{k+1}(t) = \begin{cases} \begin{bmatrix} \frac{x^k_1(t)+x^k_2(t)}{2} \\ x^k_2(t) \end{bmatrix} & \text{if } \sum_{i=1}^{N} \sum_{j \in L(i)} \hat{u}^k_{ij}(t) \leq P_{\max}(t) \\ \begin{bmatrix} x^k_1(t) \\ \frac{x^k_1(t)+x^k_2(t)}{2} \end{bmatrix} & \text{otherwise.} \end{cases} \tag{4.8}$$

$$\theta^k(t) = x^k_1(t),$$

with $x^0_1(t) = 1$ and $x^0_2(t) = 0$, for all $t \in \tau$. Recall that $\hat{u}^k_{ij}(t)$ is computed using the expression (4.4), with threshold $\theta^k(t)$, for all $j \in \textbf{Onoff}(i)$. Observe this is a bisection-like search approaching asymptotically the optimal threshold $\theta^\star(t)$.

From Proposition 4.1, we have that $\theta^0(t)$ provides a feasible solution for the optimization problem $\mathcal{P}1$. Furthermore, given the threshold recursion (4.8), it is easy to see that θ^k provides a feasible solution to the problem, for all $k \in \mathbb{N}$.

This algorithm can be run until the error $\|\theta^\star(t) - \theta^k(t)\| < \varepsilon$, for some $\varepsilon \ll 1$.

Since this is a bisection-based algorithm, it is clear that $\|\theta^\star(t) - \theta^k(t)\| \leq \|x_1^k(t) - x_2^k(t)\| \leq (1/2)^k$, for all $k \in \mathbb{N}$. Thus, the stopping criterion can be recast as $k \geq -\log_2 \varepsilon$.

Following the communication architecture from Figure 4.1, in order to estimate $\theta^\star(t)$, for $t \in \tau$, we have that $c^k = \theta^k(t)$, while $d_i^k = \sum_{j \in L(i)} \hat{u}_{ij}^k(t)$.

Remark 4.5. *Notice that the thresholding process aims to assign all the available power $P_{\max}(t)$ for the DRE at each time. This approach is not the best if the amount of power that all loads need at time t is less than $P_{\max}(t)$. Some thermal loads could be on in spite of being better off switched down. A simple way to solve this problem is to compute the power assignment using the introduced optimization/thresholding approach, and then using a lower-level local on/off controller for those thermal loads that were assigned power at time t. Such controller sets the input to zero if the load is hitting the upper bound in the comfort range of temperature.* ◇

4.4 Model predictive control implementation

The model we use to construct the optimization problem $\mathcal{P}1$ is subject to several sources of uncertainty. The outside temperature $T_{ij}^a(t)$ comes from a forecast process that presents error. The thermal models themselves are not necessarily a perfect representation of the thermal loads. There can be unmodeled disturbances that affect the system performance. The parameters $\alpha_{ij}(t)$, $\beta_{ij}(t)$ may also come from forecast processes, e.g., if they are related to natural illumination in a room.

A way of addressing this uncertainty is via a Model Predictive Control (MPC) methodology [44]. By means of this, an optimization problem is solved at the beginning of each time slot $t \in \tau$, in which the initial conditions for the thermal systems are measured, and the forecast of those unknown variables is updated based on the latest information available at the moment. Then, from the computed control input for all time steps in $\{q \in \tau \mid q \geq t\}$, only the values corresponding to time t are applied on the plant.

In order to compute the control input, we define the problem:

$$\mathcal{M}1(t) : \text{minimize}_u \sum_{q=t}^{T-1} \sum_{i=1}^{N} \sum_{j \in L(i)} f_{ij}(q) \tag{4.9a}$$

subject to:

Dynamics in Equation (4.1), with i. c. $x_{ij}(t) = x_{ij}^t$,

$$\forall j \in \mathbf{Th}(i), \forall i \in I, \forall q \in \{t, \ldots, T-1\}, \tag{4.9b}$$

$$u_{ij}(t) \in [0, u_{ij}^{\max}], \forall j \in \mathbf{Curt}(i), \forall i \in I, \forall q \in \{t, \ldots, T-1\}, \tag{4.9c}$$

$$u_{ij}(t) \in \{0, u_{ij}^{\text{on}}\}, \forall j \in \mathbf{Onoff}(i), \forall i \in I, \forall q \in \{t, \ldots, T-1\}, \tag{4.9d}$$

$$\sum_{i=1}^{N} \sum_{j \in L(i)} u_{ij}(t) \leq P_{\max}(t), \quad \forall q \in \{t, \ldots, T-1\}, \tag{4.9e}$$

for each $t \in \tau$, where i.c. stands for initial conditions, and they are measured (or estimated) using the measured information at time $t \in \tau$. A feasible solution $\hat{u}^t$ of this problem can be computed in a decentralized manner using the methodology presented in Section 4.3. Moreover, since we only use the first step of the control input for each load, i.e., $\hat{u}_{ij}^t(t)$ into the system, we do not need to run all the computations described in Section 4.3. At each time $t \in \tau$, we simply need to:

- Solve the convex relaxation of the problem $\mathcal{M}1$.
- Run a thresholding process to compute $\hat{u}_{ij}(t)$ for all $j \in \mathbf{Onoff}(i)$, $i \in I$.
- Refine the solution for all $\hat{u}_{ij}(t)$, $j \in \mathbf{Curt}(i)$, $i \in I$, by solving the convex relaxation of the problem $\mathcal{M}1$, including the constraint $u_{ij}(t) = \hat{u}_{ij}(t)$, for all $j \in \mathbf{Onoff}(i)$, $i \in I$.

4.5 Simulations

In this section, we aim to show and discuss the performance of our decentralized technique for demand response management.

First, we compare the **ADAL** approach with the widely known Dual Decomposition algorithm. To this end, we generate a random DRE scenario with 10 buildings. The maximum flexible load is near 2 MW and the DRE maximum power constraint is 500

kW. Then, we aim to solve the convex relaxation of the problem $\mathcal{P}1$, with $T = 12$. Recall that T is the discrete duration of the DRE. Figure 4.2 shows the maximum amount

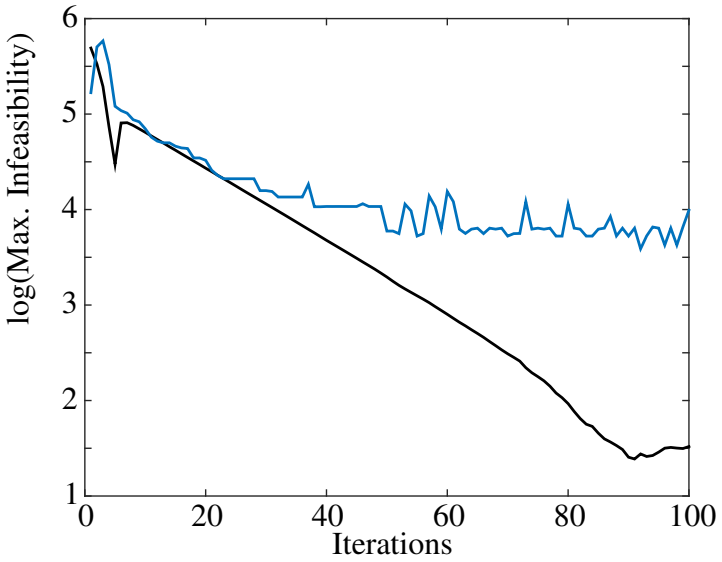

Figure 4.2: $\log_{10}$ of maximum constraint violation vs iterations for: the Dual Decomposition algorithm (blue) and for the extension of the **ADAL** algorithm (green).

of constraint violation vs the number of executed iterations. For convenience, the vertical axis is presented in a logarithmic scale. It is evident that the **ADAL** significantly outperforms the Dual Decomposition algorithm in speed of convergence. Both methods converge to the same primal solution of the problem.

Next, we run randomly generated scenarios with 10, 15, 20, 25 and 30 buildings, where each building corresponds to an average load of 0.2 MW. In addition, for each scenario, it is considered that the maximum available power for the entire set of buildings at each time is 30% of the maximum flexible load of the scenario. Figure 4.3 shows in red the cost (discomfort) for different scenarios for our suboptimal solution approach, while the blue bars represent a lower bound in the optimal cost. This lower bound consists in the solution to the convex relaxation of $\mathcal{P}1$ for each scenario. Recall that the solution to the convex relaxation of $\mathcal{P}1$ is not feasible for controlling on/off loads, while our approach does provide a feasible control for on/off loads. Figure 4.4 shows the normalized amount of energy that is provided to each type of load during the DRE duration, by our suboptimal approach (red), and by the solution to the convex relaxation

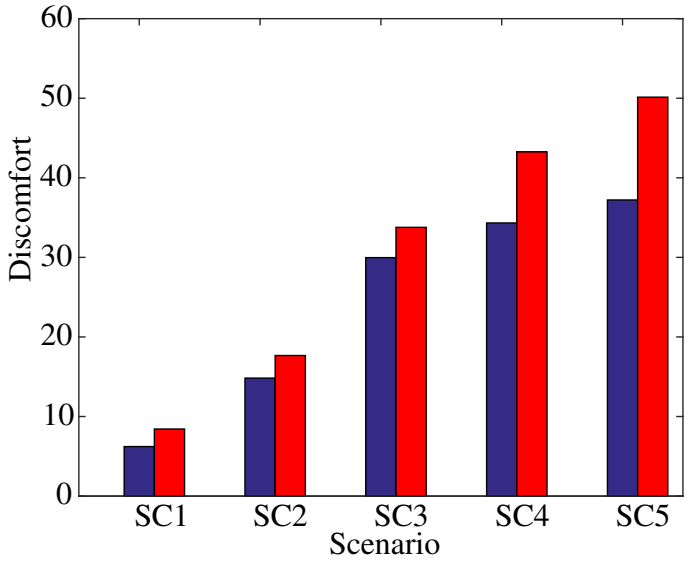

Figure 4.3: In blue: discomfort (cost) for the solution to the convex relaxation of the problem $\mathcal{P}1$ (infeasible). In red: normalized discomfort for solution to $\mathcal{P}1$ provided by our proposed method (feasible). Simulation over 5 different scenarios.

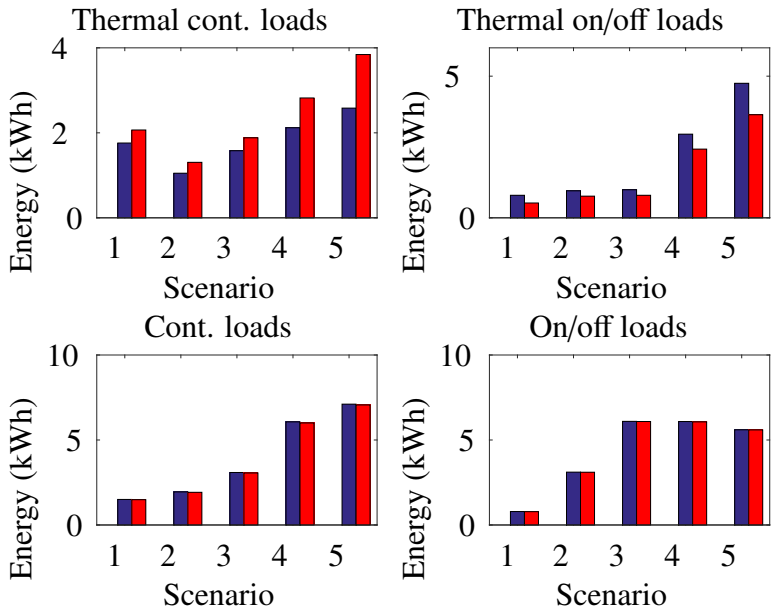

Figure 4.4: In blue: energy dispatched according to the solution to the convex relaxation of the problem $\mathcal{P}1$, for each type of flexible load. In red: energy dispatched according to our proposed method for each type of load. Simulation over 5 different scenarios.

of $\mathcal{P}1$ (blue). Notice that the solution to the convex relaxation of $\mathcal{P}1$ is not feasible to implement and it is only used as a benchmark. It can be seen that the loads in $\text{Th}_{\text{curt}}(i)$ for all $i \in I$ tend to have more power availability with our approach than with the convex relaxation of $\mathcal{P}1$, while the loads in $\text{Th}_{\text{on/off}}(i)$ for all $i \in I$ tend to receive less power. This is consistent with the fact that on/off loads cannot be partially activated, which is not taken into account in the convex relaxation of $\mathcal{P}1$. All non-thermal loads seem to receive slightly less energy availability with our approach that in the convex relaxation of $\mathcal{P}1$. This happens because for this simulation exercise, thermal loads are on average 50 times larger than non-thermal ones. This implies that the suboptimality of our approach affects more the larger loads than the smaller ones. A more detailed description of the simulation scenario may be found in the extended version of this manuscript [45].

Finally, we compare the MPC implementation of our management strategy, vs an open-loop computation at the beginning of the DRE, where there is uncertainty in the problem parameters. We introduce normally distributed model uncertainty in the thermal load models, and we contaminate the data on outside temperature with noise. The variance of this noise increases as the number of steps ahead that the variable is forecasted increases. We have used 5 randomly generated scenarios with 10, 20, 30, 40 and 50 buildings, where each building has an average load of 0.2 MW. Figure 4.5 shows how with the MPC approach the cost (discomfort) decreases for all scenarios. This improvement occurs due to the feedback that is inherent to the MPC approach, and the fact that the up-to-date forecast of a variable contains less error than the simple propagation of an outdated forecast.

4.6 Summary

We propose a decentralized method for coordination of loads in a demand response event (DRE). The method takes explicitly into account the fact that some loads are on/off, and aims for the suboptimal solution of a mixed-integer program. The objective of this problem is the minimization of user discomfort due to the DRE. The solution approach consists in solving a convex relaxation of the mixed-integer program, combined with a decentralized greedy dispatch of the available power. While the convex

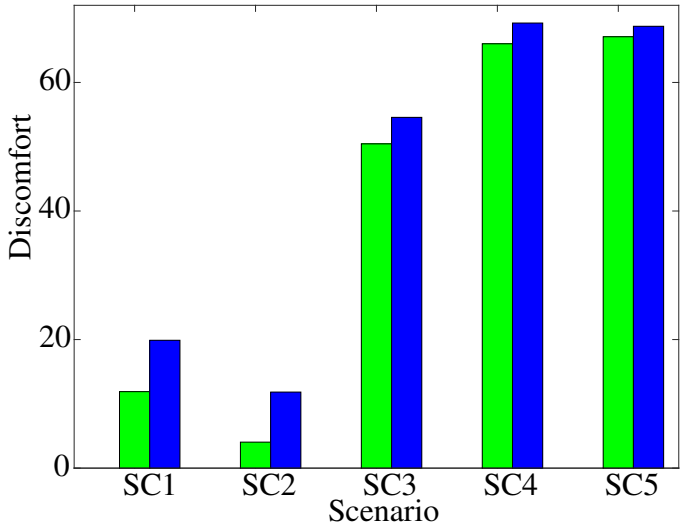

Figure 4.5: In green: normalized discomfort (cost) for the solution provided by our method. In blue: normalized discomfort for the MPC implementation of our method.

relaxation provides a measure of the urgency level for power of each load in the DRE, a DRE manager must compute an urgency threshold to decide on the on/off loads that must receive power at each time. Additionally, we present an MPC implementation of our approach, in order to mitigate uncertainties and disturbances of the model. Simulations show that the loss of optimality of our approach is acceptable, and the possibility of computing the solution in a tractable time makes it a tool that can be used for problems with similar structure.

As future directions, we aim to establish analytic suboptimality bounds for our approach.

Acknowledgements

This chapter has been submitted for publication in the following work:

- A. Cortés and S. Martínez, "Decentralized Management of Demand Response Events with On/Off Loads," submitted to the American Control Conference (2016).

Chapter 5

Distributed control of user-side resources in microgrids

In order to solve problems related to voltage stability and variable load satisfaction, several solutions are being contemplated. Additional backup generation plants could help compensate fluctuations, but their deployment would incur significantly higher costs for both the utilities and the users. Large-scale storage systems are proposed to shift the energy generated during low-demand hours or that corresponding to high-solar generation times to those of high demand. Smart inverters placed at the PV themselves can be leveraged to inject reactive power for voltage regulation and optimization of the network performance. In particular, since PV systems and batteries can be locally owned and distributed throughout the network, distributed optimization algorithms can be used to achieve the power network objectives in a faster and more robust way. These objectives can be, to some extent, expressed as the well known *Optimal Power Flow* (OPF) problem.

The OPF problem [46] is a non-convex, hard optimization problem that has received wide attention in the literature. Most of the works on the OPF propose centralized solutions that consider voltage as a decision variable [47, 48]. Recently, the papers [48, 49, 50] circumvent the nonconvexity of the OPF through a relaxation that ensures a zero-duality gap under some general conditions. The work [51, 52] studies an OPF setting with storage integration in which battery systems are added to some nodes in a grid. The paper [51] addresses an OPF problem where voltages and battery charg-

ing/discharging rates are taken as decision variables, simplified by assumptions such as small-angles and infinite charging/discharging capacities. Electricity prices are assumed to vary over time, and the objective function is the cost of the energy provided by the utility. The paper [52] presents a more general setting in which the assumptions of [51] are avoided and exploits the aforementioned zero-duality-gap approach.

The reactive power control problem, which consists of providing for losses without producing excess heating or incurring voltage drops, has been addressed in a significant number of works. In the 1980s, banks of capacitors and transformer taps were used for reactive power compensation. In [53] a problem of optimal sizing for capacitors is solved using a relaxation of the power flow equations over radial networks. An optimal reactive power generation algorithm is introduced in [54], in order to minimize active power losses and improve voltage regulation. More recently, research has focused on the control of smart inverters which allow changes in the generated reactive power. In [55], a convexification of the OPF problem is presented for grids that fulfill some assumptions on the input voltage and the impedance in the transmission lines. The objective is to minimize transmission losses in the grid by varying the injected reactive power at all generators in a distributed way. This work does not consider voltage regulation constrains, which are addressed in [6]. Although the convexification idea of [6] remains identical to that of [55], a different communication structure is employed.

Here, we propose an algorithm to compute both the optimal reactive power generation and storage control strategies for a microgrid over a given time horizon. This computation is intended for use in a model-predictive control scheme that can incorporate forecasts on generation and load. In order to present a distributed algorithm based on the dual decomposition method, we employ a convexification approach that exploits high voltages at the connection point of the microgrid. The type of microgrid we consider is endowed with generation and/or storage capacity at certain nodes. We consider a discrete-time horizon, which, at each instant of time, has an associated electricity cost per kWh, a forecasted active generation, and node-wise load. The algorithm utilizes measurements and predictions of voltage in the microgrid, which have implicit information on the power injected at the nodes, in order to choose the optimal reactive power that must be injected by each generator, and the optimal charging/discharging rate of

the storage at each node. The goal is to minimize a cost function that weighs in the grid transmission losses with the overall cost of the active power provided by the utility during the time horizon. Finally, the solution provided by the algorithm is meant to respect voltage regulation constraints. The algorithm convergence analysis is performed by characterizing the behavior of an upper bound algorithm which can be studied via Lyapunov theory. In particular, we conclude that trajectories converge to the unique value of reactive power injection and charging/discharging rate for each node of the microgrid. Finally, we describe a novel way of approximately predicting voltage in order to implement the proposed algorithm.

This chapter is organized as follows. Section 5.2 presents the microgrid model and the optimization problem we aim to solve. Section 5.3 describes the dual decomposition algorithm to solve the optimization problem. In Section 5.4 we introduce a different algorithm that presents better convergence rate. Section 5.5 introduces a solution to voltage prediction required by both algorithms. Simulations and comparison of the rate of convergence for both algorithms are presented in Section 5.6.

5.1 Notions of graph theory for this chapter

Let $\mathbb{G} = (\mathcal{V}, \mathcal{E}, Y)$ be an undirected weighted graph with a set $\mathcal{V} = \{0, \ldots, N-1\}$ of N vertices, a set of edges $\mathcal{E}$ and a weight matrix $Y \in \mathbb{C}^{N\times N}$. Each edge in $\mathcal{E}$ is expressed as (h, l), for $h, l \in \mathcal{V}$. Consider some labeling of the set $\mathcal{E}$ with the set of indices $\{1, \ldots, |\mathcal{E}|\}$. In addition, let us assign an arbitrary direction to each edge $(h, l) \in \mathcal{E}$. The *incidence matrix* A of $\mathbb{G}$ is a matrix in $\{0, \pm 1\}^{|\mathcal{E}|\times N}$, which depends on the arbitrary direction associated with each edge of $\mathbb{G}$, such that:

$$A_{dl} = \begin{cases} -1, & \text{if } \varepsilon_d \in \mathcal{E} \text{ is an outgoing edge of } l \in \mathcal{V}, \\ 1, & \text{if } \varepsilon_d \in \mathcal{E} \text{ is an incoming edge of } l \in \mathcal{V}, \\ 0, & \text{otherwise.} \end{cases}$$

Consider a diagonal matrix $C \in \mathbb{C}^{|\mathcal{E}|\times|\mathcal{E}|}$ such that C_{ii} is the weight of $\varepsilon_i = (h, l) \in \mathcal{E}$ given by Y_{hl}. Then, the *Laplacian* of the undirected graph associated with $\mathbb{G}$ is given by

$L = A^\top CA$. Consider the graph $\mathbb{G}$. A *path* $\mathcal{P}(l, h)$, for nodes $l, h \in \mathcal{V}$, is defined as a sequence of nodes $\{n_1, \ldots, n_\ell\}$ such that $n_1 = l$, $n_\ell = h$ and (n_i, n_{i+1}) is an edge of $\mathbb{G}$, for all $i \in \{1, \ldots, \ell - 1\}$.

5.2 Problem statement

Consider a microgrid which is connected to the grid at a single point. The microgrid is modeled as an undirected weighted graph $\mathbb{G} = (\mathcal{V}, \mathcal{E}, Y)$, $\mathcal{V} \triangleq \{0, \ldots, N-1\}$, $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$, $Y \in \mathbb{C}^{N \times N}$, where nodes in $\mathcal{V}$ represent buses and edges in $\mathcal{E}$ represent the interconnection lines. Weights are given by the matrix Y and correspond to the line admittances of such interconnection lines. We consider a microgrid with generation and storage capacities. This microgrid has three different types of nodes: i) a single connection point to the grid, represented by node 0, which acts as a slack node, with fixed voltage and unlimited power generation, ii) a subset of nodes $\mathcal{G}$ with generation and storage capacity, and iii) a set $\mathcal{M} = \mathcal{V} \setminus (\{0\} \cup \mathcal{G})$ of nodes with neither generation nor storage capacity. All nodes in $\mathcal{V} \setminus \{0\}$ have a load that must be satisfied.

Let us consider a discrete, finite-time window τ, with T time slots, i.e., $\tau \triangleq \{1, \ldots, T\}$. Each time slot t has an electricity cost $c(t)$ associated with it, which is the price per kWh, given by a map $c : \tau \to \mathbb{R}_{\geq 0}$, and depends on the overall demand satisfied by the utility. The value of $c(t)$ for each $t \in \tau$ is assumed to be known to nodes of the grid, since utility companies make it publicly available. Similarly, for each time slot $t \in \tau$ there is an amount of active power that each generator can provide during the whole time slot t, called $p_{\text{in},l}(t) > 0$, for each $l \in \mathcal{G}$. We assume that this value cannot be controlled, as it is the case with renewable generation, but we can estimate it using forecasting techniques. In order to make the problem slightly more general, the load at each node may also depend on the time slot $t \in \tau$. For each $l \in \mathcal{V} \setminus \{0\}$, let $p_{\text{load},l}(t) + jq_{\text{load},l}(t)$, $p_{\text{load},l}(t) \geq 0$, be power that must be supplied to a load placed at the node, for all $t \in \tau$.

The problem we would like to solve consists in finding the optimal policy for the reactive power injection, and for the storage charging/discharging rates, that optimizes the microgrid operation defined by the combined objective of minimizing active power

generation cost and distribution losses. Before stating the problem more formally, we next describe the model we use to represent a microgrid and the elements involved in it.

5.2.1 Mathematical model of a microgrid

Let $u_l(t) \in \mathbb{C}$ be the voltage at node $l \in \mathcal{V}$, at time $t \in \tau$. Let $\iota_l(t) \in \mathbb{C}$ be the current flowing through node $l \in \mathcal{V}$. Let A be the incidence matrix associated with $\mathbb{G}$, based on some given ordering on $\mathcal{E}$. For node 0, let $u_0(t)$ be fixed for all $t \in \tau$ and described as $u_0(t) = U_0 e^{j\phi}$. Let $s_l(t) = u_l(t)\hat{\iota}_l(t)$ be the complex apparent power drawn or supplied to node l at time t. Recall that $s_l(t) = P_l(t) + jQ_l(t)$, where $P_l(t)$, $Q_l(t)$ are the active and reactive power at node $l \in \mathcal{V}$, at time $t \in \tau$.

Next, we introduce some compact-form notation. Define $u(t) \triangleq [u_0, u_G^\top(t), u_L^\top(t)]^\top \in \mathbb{C}^{N\times N}$ as the voltage vector for all nodes in the set $\mathcal{V}$, at time $t \in \tau$, where $u_G(t)$ is the vector of all voltages at nodes in $\mathcal{G}$, and $u_L(t)$ is the vector of all voltages in the set $\mathcal{M}$. Likewise, $\iota(t) \triangleq [\iota_0(t), \iota_G^\top(t), \iota_L^\top(t)]^\top$, $s(t) \triangleq [s_0(t), s_G^\top(t) s_L^\top(t)]^\top$, with $s(t) = P(t) + jQ(t)$, $P(t) \triangleq [P_0(t), P_G^\top(t), P_L^\top(t)]^\top$, and $Q(t) \triangleq [q_0(t), Q_G^\top(t), Q_L^\top(t)]^\top$.

The convention for the active power sign is that if $P_l(t) < 0$, power is injected to the l^{th} node, while $P_l(t) > 0$ means that power is drawn from the l^{th} node, for all $l \in \mathcal{V}$, $t \in \tau$. As an example, for a node with load only, it must hold that $P_l(t) \leq 0$ for all $t \in \tau$. Likewise, for a node l with only generation, but neither storage nor load, $P_l(t) \geq 0$ for all $t \in \tau$. Given the type of nodes we consider in the microgrid, we have that $P_G(t) = p_{\text{in},G}(t) - p_{\text{load},G}(t) - v(t)$, where $v(t)$ is a vector whose l^{th} component $v_l(t)$ is the amount of power that is being supplied to the battery at such node, $l \in \{1, \ldots, |\mathcal{G}|\}$. Similarly, $P_L(t) = -p_{\text{load},L}(t)$, for all $t \in \tau$. Finally, we have that $Q_G(t) = q_{\text{in}}(t) - q_{\text{load},G}(t)$ and $Q_L(t) = -q_{\text{load},L}(t)$, for all $t \in \tau$. The vector $q_{\text{in}}(t) \in \mathbb{R}^{|\mathcal{G}|}$ represents the reactive power that is supplied by each generator in $\mathcal{G}$ at time $t \in \tau$.

Then, using the Kirchoff's current and voltage laws, the relation between voltage and current is given by:

$$\begin{aligned} A^\top \mathcal{I}(t) + \iota(t) &= 0 \\ Au(t) + Z\mathcal{I}(t) &= 0, \end{aligned} \tag{5.1}$$

where $\mathcal{I}(t) \in \mathbb{C}^{|\mathcal{E}|}$ is a vector with the values of current at each edge in $\mathcal{E}$ and time $t \in \tau$, and Z is the diagonal matrix in $\mathbb{C}^{|\mathcal{E}|\times|\mathcal{E}|}$ whose elements are the line impedances in the microgrid.

In the following, we will make the following assumptions on the microgrid parameters. These assumptions have already been used in [55, 6], and are accurate for actual operation conditions in real microgrids.

Assumption 5.1 (Large input voltage at the microgrid)**.** *The value of U_0 is very large as compared to the currents provided by the inverters and batteries, or supplied to the loads.*

Assumption 5.2 (Transmission lines' reactance/resistance ratio)**.** *The microgrid has transmission lines with the same reactance/resistance ratio. Therefore, for all edges $\ell \in \mathcal{E}$, the impedance z_ℓ can be written as $z_\ell = |z_\ell|e^{j\theta}$.*

5.2.2 Mathematical model of a battery

We model a battery with the following dynamics:

$$x_l(t) = x_l(t-1) + \frac{1}{\beta_l} v_l(t), \qquad \forall l \in \mathcal{G}, \quad \forall t \in \tau, \tag{5.2}$$

where $x_l(t) \in [0, 1]$ is the battery state, β_l corresponds to the battery capacity, divided by the length of the time slot. Recall that $v_l(t)$ denotes the power injected to or drained from the battery during the time slot $t \in \tau$. Clearly, the battery has physical constraints, formulated as $v_l \in [V^l_{\text{dis}}, V^l_{\text{ch}}]$, where $V^l_{\text{dis}} < 0$, $-V^l_{\text{dis}}$ is the maximum amount of power the battery can discharge during a time slot $t \in \tau$, and $V^l_{\text{ch}} > 0$ is the maximum amount the battery can charge. Note that this model assumes that the charger efficiency is one. Assume that the initial state for the battery at node $l \in \mathcal{G}$ is $x_l(0)$. Then, for each $t \in \tau$, the battery charge is given by:

$$x_l(t) = x_l(0) + \frac{1}{\beta_l} \sum_{\ell=1}^{t} v_l(\ell), \qquad \forall l \in \mathcal{G}, \quad \forall t \in \tau.$$

In compact form, denote by $v(t)$ the vector of battery charge/discharge rates at time t, for all nodes in the set $\mathcal{G}$.

5.2.3 Communication network

Generators and storage systems will coordinate operations by means of a communication network. The communication topology is based on the microgrid topology and the location of the generation/storage nodes in the microgrid.

Definition 5.1 (Communication network)**.** *The communication network is given by the undirected graph* $\mathbb{G}_G = (\mathcal{G} \cup \{0\}, \mathcal{E}_G)$*, where* $\mathcal{E}_G \subseteq \mathcal{G} \cup \{0\} \times \mathcal{G} \cup \{0\}$ *is defined as* $\mathcal{E}_G \triangleq \{(l,h) \in \mathcal{G} \cup \{0\} \times \mathcal{G} \cup \{0\} \mid \mathcal{P}(h,l) \cap (\mathcal{G} \cup \{0\}) = \{h,l\}\}$*. The set of* neighbors $\mathcal{N}_G(l)$ *of* $l \in \mathcal{G} \cup \{0\}$ *in the* communication network *is given by* $\mathcal{N}_G(l) \triangleq \{h \in \mathcal{G} \cup \{0\} \mid (l,h) \in \mathcal{E}_G\}$.

5.2.4 The MICROGRID CONTROL PROBLEM

Based on the microgrid and battery models we have presented above, actuation over the microgrid will be established through the decision variables $q_{\text{in},l}(t)$, which represents the reactive power supplied by generation systems, and $v_l(t)$, which represents the active power provided by the batteries, for all $l \in \mathcal{G}$, while the input voltage u_0, the energy cost $c(t)$, the loads $p_{\text{load},l}(t)$ and $q_{\text{load},l}(t)$ for $l \in \mathcal{V} \setminus \{0\}$, and the forecasted active power generation $p_{\text{in},l}(t)$, for $l \in \mathcal{G}$, are parameters of the problem, for all $t \in \tau$.

Our objective is to compute optimal reactive power generation and storage control profiles for the time horizon τ. This must be done in such a way that the voltage at each node in $\mathcal{G}$ is maintained within a desirable range and the storage control respects all physical constraints related to the batteries, while a cost is minimized. The cost we consider encompasses two possibly conflicting objectives: i) minimize the cost of active power from the grid and ii) minimize the transmission losses in the transmission lines present in the microgrid. Minimizing transmission losses is a significant objective in Optimal Power Flow and it has been considered in several works [48, 52, 55, 6]. Thus, the cost function is given by:

$$J(u) = \sum_{t \in \tau} (J_{\text{loss}}(u(t)) + \delta J_{\text{power}}(u(t))),$$

where $J_{\text{loss}}(u(t))$ represents the loss in the transmission lines of the microgrid at time slot t, and $J_{\text{power}}(u(t))$ is the overall cost of the active power provided by the utility at time slot

t. The trade-off trade-off between these possibly conflicting objectives, is parameterized by the nonnegative constant $\delta > 0$, which is used to modify the relative importance of $J_{\text{power}}(u(t))$ with respect to $J_{\text{loss}}(u(t))$ in the optimization. The loss in the transmission lines can be expressed as $J_{\text{loss}}(u(t)) = \hat{u}^\top(t)Lu(t)$, where $L = A^\top Z_{\text{mag}}^{-1}A$, $Z_{\text{mag}} = \|Z\|_{\mathbb{C}}$ [6], and the power cost is given by $J_{\text{power}}(u(t)) = -c(t)\,\text{Re}(s_0(t)) = -c(t)\,\text{Re}(\hat{\imath}_0(t)u_0(t))$. The negative sign follows the introduced convention for the active power sign. Since $\iota(t)$ can be approximated by Assumption 5.2 on the common transmission lines' reactance/resistance ratio as $\iota(t) = e^{-j\theta}Lu(t)$, then $J_{\text{power}}(u(t)) = -\,\text{Re}\{e^{j\theta}\hat{u}^\top(t)L\mathbf{e}_0\mathbf{e}_0^\top u(t)\}c(t)$.

The following results give us a convenient way of approaching the problem, by writing $u(t)$ as a linear function of the decision variables.

Lemma 5.1 (Existence of matrix X [55])**.** *There exists a unique symmetric, positive semidefinite matrix $X \in \mathbb{R}^{N\times N}$, which can be written as:*

$$X = \begin{bmatrix} 0 & 0 & 0 \\ 0 & W & F \\ 0 & F^\top & R \end{bmatrix},$$

such that $XL = I_N - \mathbf{1}(\mathbf{e}_0)^\top$ and $X\mathbf{e}_0 = 0$, where $W \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{G}|}$, $R \in \mathbb{R}^{|\mathcal{M}|\times|\mathcal{M}|}$, and $F \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{M}|}$.

The physical meaning of the matrix X is widely discussed in [55], [6]. One of the properties of X is that the product $(\mathbf{e}_h - \mathbf{e}_l)^\top X(\mathbf{e}_h - \mathbf{e}_l)$ corresponds to the effective impedance from node $h \in \mathcal{V}$ to node $l \in \mathcal{V}$. The following result provides a linearization of the relation between voltages and powers on the microgrid.

Lemma 5.2 (Microgrid voltage approximation [6])**.** *Consider* (5.1)*, along with $s_l(t) = u_l(t)\hat{\imath}_l(t)$. Then, the microgrid voltages satisfy:*

$$\begin{bmatrix} u_0(t) \\ u_G(t) \\ u_L(t) \end{bmatrix} = e^{j\phi}\left(U_0\mathbf{1} + \frac{e^{j\theta}}{U_0}\begin{bmatrix} 0 & 0 & 0 \\ 0 & W & F \\ 0 & F^\top & R \end{bmatrix}\begin{bmatrix} 0 \\ \hat{s}_G(t)) \\ \hat{s}_L(t) \end{bmatrix}\right) + o\left(\frac{1}{U_0}\right). \tag{5.3}$$

Notice that by Assumption 5.1 on the large magnitude of the input voltage, the relaxation given by Lemma 5.2 provides an accurate approximation, as the term $o\left(\frac{1}{U_0}\right)$ vanishes for large values of U_0.

The following result follows directly from Lemmas 3 and 4 in [6].

Lemma 5.3 (Matrix G)**.** *There exists a unique symmetric matrix* G*, such that:*

$$\begin{bmatrix} 0 & 0 \\ 0 & W \end{bmatrix} \mathsf{G} = I_{|\mathcal{G}|+1} - \mathbf{1}(\mathbf{e}_0)^\top, \qquad \mathsf{G}\mathbf{1} = 0.$$

Moreover, the matrix G *has a sparsity induced by the communication network graph, this is,* $\mathsf{G}_{ij} \neq 0$ *if and only if* $j \in \mathcal{N}_G(i)$*. The matrix* W *is a block of the matrix* X *described in Lemma 5.1.*

From the result above, it is immediately noted that the matrix W is invertible. We can also see that its inverse matrix corresponds to a block in the matrix G, which means that it also has a sparsity induced by the communication network graph.

Replacing the approximation in (5.3), in the cost function for the problem, it can be rewritten as a quadratic function of the decision variables q_{in} and v. Let us define $q_{\text{in}} \in \mathbb{R}^{|\mathcal{G}|T}$, $v \in \mathbb{R}^{|\mathcal{G}|T}$ as $q_{\text{in}} = [q_{\text{in},1}^\top, \ldots, q_{\text{in},|\mathcal{G}|}^\top]^\top$ and $v = [v_1^\top, \ldots, v_{|\mathcal{G}|}^\top]^\top$, where $q_{\text{in},\ell}$ is the vector whose entries are the generated reactive power of the ℓ^{th} generator for all times in the horizon τ, and $q_{\text{in}}(t)$ is the vector whose entries are the reactive power generated at each generator at time $t \in \tau$. Vectors v_ℓ and and $v(t)$ are defined similarly. After the algebraic procedure, and removing all constant terms that do not affect the problem solution, we obtain:

$$J(q_{\text{in}}, v) = \sum_{t=1}^{T} \Big(q_{\text{in}}^\top(t) W q_{\text{in}}(t) + 2\xi_q^\top(t) q_{\text{in}}(t) + v^\top(t) W v(t) + 2\xi_v^\top(t) v(t) \Big) + \delta \sum_{t=1}^{T} c(t) \mathbf{1}^\top v(t),$$

where $\xi_q(t) \in \mathbb{R}^{|\mathcal{G}|}$, $\xi_v(t) \in \mathbb{R}^{|\mathcal{G}|}$ are given by:

$$\begin{aligned} \xi_q(t) &= W q_{\text{load},G}(t) + F q_{\text{load},L}(t) \\ \xi_v(t) &= -W(p_{\text{in},G}(t) - p_{\text{load},G}(t)) + F p_{\text{load},L}(t). \end{aligned} \tag{5.4}$$

Having established the relaxed cost function that takes into account the linear relaxation of the power flow equations, we formulate the following valid approximation

of the OPF problem:

$$\min_{q_{\text{in}},v,u} J(q_{\text{in}}, v)$$

subject to:

$$U_{\min} \leq \|u_l(t)\|_{\mathbb{C}} \leq U_{\max}, \qquad l \in \mathcal{G}, t \in \tau, \tag{5.5a}$$

$$|q_{\text{in},l}(t)| \leq q_{\text{in,max}}(t), \qquad l \in \mathcal{G}, t \in \tau, \tag{5.5b}$$

$$V^l_{\text{dis}} \leq v_l(t) \leq V^l_{\text{ch}}, \qquad l \in \mathcal{G}, t \in \tau, \tag{5.5c}$$

$$0 \leq x_l(0) + \frac{1}{\beta_l} \sum_{s=1}^{T} v_l(s) \leq 1, \qquad l \in \mathcal{G}, t \in \tau. \tag{5.5d}$$

Clearly, the lower bound constraint on the voltage magnitude introduces a non-convex constraint to the problem. In order to follow the convexification idea in [6], we define $w_G(t)$, as a vector in $\mathbb{R}^{|\mathcal{G}|}$, whose components are the squares of the magnitudes of the complex voltages $u_l(t)$, normalized by U_0^2, for $l \in \mathcal{G}$. After some manipulation, from (5.3), we obtain:

$$w_G(t) = \mathbf{1} + \frac{2}{U_0^2}\Big(\cos\theta\Big(WP_G(t) + FP_L(t)\Big) \quad + \sin\theta\Big(WQ_G(t) + FQ_L(t)\Big)\Big) + o\left(\frac{1}{U_0^2}\right). \tag{5.6}$$

Thus, we can write the constraints on the voltage magnitude as:

$$W_{\min} = \frac{U^2_{\min}}{U_0^2} \leq w_l(t) \leq \frac{U^2_{\max}}{U_0^2} = W_{\max}, \tag{5.7}$$

for all $l \in \mathcal{G}$, $t \in \tau$. Clearly, as $U_{\min}$, $U_{\max}$ have in practice a similar order-of-magnitude than that of U_0, it holds that $U_0^{-2}U^2_{\min}$, $U_0^{-2}U^2_{\max}$ are close to one. For large values of U_0, (5.6) is affine in the decision variables $q_{\text{in}}(t)$, $v(t)$, hence the constraint above is convex.

5.3 Distributed reactive power and storage control

In order to solve the MICROGRID CONTROL PROBLEM in a distributed way, we propose an extension of the dual decomposition approach presented in [6]. Here we optimize not

only the reactive power injection, but also on the battery charge/discharge, considering the physical constraints of the battery control. The fact that we also consider a different cost function modifies the algorithm dynamics. The dual decomposition algorithm consists of performing a gradient ascent on the Lagrangian with respect to the dual variables of the problem, while an unconstrained optimization with respect to the primal variables, parameterized by the estimated dual variables is executed.

The Lagrangian for the optimization problem is given by:

$$\begin{aligned}\mathcal{L}(q_{\text{in}},v,\psi) = J(q_{\text{in}},v) &+ \sum_{t=1}^{T} \underline{\lambda}^{\top}(t)(W_{\min} - w_G(t)) + \sum_{t=1}^{T} \overline{\lambda}^{\top}(t)(w_G(t) - W_{\max}) \\ &+ \sum_{t=1}^{T} \underline{\eta}^{\top}(t)\mathbf{d}(\Delta V)^{-1}(V_{\text{dis}} - v(t)) + \sum_{t=1}^{T} \overline{\eta}^{\top}(t)\mathbf{d}(\Delta V)^{-1}(v(t) - V_{\text{ch}}) \\ &- \sum_{t=1}^{T} \underline{\mu}^{\top}(t)\left(x(0) + \mathbf{d}(\beta^{-1})\sum_{h=1}^{t} v(h)\right) \\ &+ \sum_{t=1}^{T} \overline{\mu}^{\top}(t)\left(x(0) + \mathbf{d}(\beta^{-1})\sum_{h=1}^{t} v(h) - 1\right),\end{aligned}$$

where $\underline{\lambda}(t), \overline{\lambda}(t), \underline{\eta}(t), \overline{\eta}(t), \underline{\mu}(t), \overline{\mu}(t) \geq 0$, for all $t \in \tau$, are Lagrange multipliers, and $\mathbf{d}(\beta^{-1}) \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{G}|}$ is a diagonal matrix such that $\mathbf{d}(\beta^{-1})_{ll} = \frac{1}{\beta_l}$, for all $l \in \{1,\dots,|\mathcal{G}|\}$. We introduce $\mathbf{d}(\Delta V) \in \mathbb{R}^{|\mathcal{G}|\times|\mathcal{G}|}$ as a diagonal matrix that works as a regularization parameter, such that $\mathbf{d}(\Delta V)_{ll} \triangleq V^l_{\text{ch}} - V^l_{\text{dis}}$, for all $l \in \{1,\dots,|\mathcal{G}|\}$. Notice that here we are not including the constraints $|q_{\text{in},\ell}| \leq q_{\text{in,max}}$ for brevity, however, including them in the procedure is very simple, by adding Lagrange multipliers.

Let us define as a compact representation of the dual variables:

- $\underline{\lambda} \triangleq [\underline{\lambda}^{\top}(1),\dots,\underline{\lambda}^{\top}(T)]^{\top} \in \mathbb{R}^{|\mathcal{G}|T}_{\geq 0}$, resp. $\overline{\lambda}$,
- $\underline{\eta} \triangleq [\underline{\eta}^{\top}(1),\dots,\underline{\eta}^{\top}(T)]^{\top} \in \mathbb{R}^{|\mathcal{G}|T}_{\geq 0}$, resp. $\overline{\eta}$,
- $\underline{\mu} \triangleq [\underline{\mu}^{\top}(1),\dots,\underline{\mu}^{\top}(T)]^{\top} \in \mathbb{R}^{|\mathcal{G}|T}_{\geq 0}$,resp. $\overline{\mu}$,
- $\psi(t) \triangleq [\overline{\lambda}^{\top}(t), \underline{\lambda}^{\top}(t), \overline{\eta}^{\top}(t), \underline{\eta}^{\top}(t), \overline{\mu}^{\top}(t), \underline{\mu}^{\top}(t)]^{\top}$,

and a variable: $\psi \triangleq [\overline{\lambda}^{\top}, \underline{\lambda}^{\top}, \overline{\eta}^{\top}, \underline{\eta}^{\top}, \overline{\mu}^{\top}, \underline{\mu}^{\top}]^{\top} \in \mathbb{R}^{r}_{\geq 0}$, where $r \triangleq 6T|\mathcal{G}|$. The dual decom-

position algorithm is given by:

$$(q_{\text{in}}^k, v^k) = \text{argmin}_{q_{\text{in}},v}(\mathcal{L}(q_{\text{in}}, v, \psi^k)), \tag{5.8}$$

$$\psi^{k+1} = \left[\psi^k + \gamma \nabla_\psi \mathcal{L}(q_{\text{in}}^k, v^k, \psi^k)\right]_+, \tag{5.9}$$

where $\nabla_\psi \mathcal{L}$ is the gradient of $\mathcal{L}$ with respect to the dual variable ψ, and $[\cdot]_+$ is the projection operator onto the positive orthant. The parameter γ is a small enough positive scalar to be characterized later.

Since the Lagrangian is a quadratic function of (q_{in}, v), a minimizer for $\mathcal{L}(q_{\text{in}}, v, \psi^k)$ can be found directly by solving $\nabla_{q_{\text{in}},v}\mathcal{L}(q_{\text{in}}, v, \psi^k) = 0$. Some algebraic manipulations lead to:

$$q_{\text{in}}^k(t) = q_{\text{load},G}(t) + W^{-1}F q_{\text{load},L}(t) - \sin\theta[\overline{\lambda}^k(t) - \underline{\lambda}^k(t)] \tag{5.10}$$

$$\begin{aligned} v^k(t) = &- p_{\text{load},G}(t) - p_{\text{in},G}(t) - W^{-1}F p_{\text{load},L}(t) + \cos\theta[\overline{\lambda}^k(t) - \underline{\lambda}^k(t)] \\ &- \frac{U_0^2}{2}W^{-1}\left(\mathbf{d}(\Delta V)^{-1}(\overline{\eta}^k(t) - \underline{\eta}^k(t)) + \sum_{h=t}^{T}\mathbf{d}(\beta^{-1})(\overline{\mu}^k(h) - \underline{\mu}^k(h)) + \delta c(t)\mathbf{1}\right), \end{aligned} \tag{5.11}$$

for all $t \in \tau$. Formulas (5.10), (5.11) are obtained as follows. First, the derivative of $\mathcal{L}$ is computed with respect to the variables $q_{\text{in}}(t)$ and $v(t)$, for all t. To this end, a chain rule is used by which $\mathcal{L}$ is differentiated with respect to $u(t)$ and $w_G(t)$, and in turn, $u(t)$ and $w_G(t)$ are differentiated with respect to $q_{\text{in},(}t)$ and $v(t)$. Next, terms of $o\left(\frac{1}{U_0^2}\right)$ are neglected. The remaining linear equations are set equal to zero and a solution in $q_{\text{in}}(t)$ and $v(t)$, for all $t \in \tau$, is found. This leads to expressions (5.10), (5.11).

Since the Lagrangian is linear in the dual variables, the derivative with respect to each of them is merely the expression representing the constraint associated with that

dual variable. Thus, the gradient ascent algorithm for the dual variables becomes:

$$
\begin{aligned}
\overline{\lambda}^{k+1}(t) &= \left[\overline{\lambda}^{k}(t) + \gamma(w_G^k(t) - W_{\max})\right]_+, \\
\underline{\lambda}^{k+1}(t) &= \left[\underline{\lambda}^{k}(t) + \gamma(W_{\min} - w_G^k(t))\right]_+, \\
\overline{\eta}^{k+1}(t) &= \left[\overline{\eta}^{k}(t) + \gamma \mathbf{d}(\Delta V)^{-1}(v^k(t) - V_{\text{ch}})\right]_+, \\
\underline{\eta}^{k+1}(t) &= \left[\underline{\eta}^{k}(t) + \gamma \mathbf{d}(\Delta V)^{-1}(V_{\text{dis}} - v^k(t))\right]_+, \\
\overline{\mu}^{k+1}(t) &= \left[\overline{\mu}^{k}(t) + \gamma\left(x(0) + \sum_{h=1}^{t} \mathbf{d}(\beta^{-1})v^k(h) - \mathbf{1}\right)\right]_+, \\
\underline{\mu}^{k+1}(t) &= \left[\underline{\mu}^{k}(t) - \gamma\left(x(0) + \sum_{h=1}^{t} \mathbf{d}(\beta^{-1})v^k(h)\right)\right]_+,
\end{aligned}
\tag{5.12}
$$

for all $t \in \tau$, with a common parameter γ. Following similar computations as in [6], one can obtain the following result:

Lemma 5.4 (Distributed algorithm)**.** *The expressions in* (5.10)*,* (5.11) *can be approximated by:*

$$
\begin{aligned}
q_{in}^k(t) = \operatorname{Im}\left(e^{-j\theta}\begin{bmatrix} 0 & \mathbf{d}(\overline{u}_G^{k-1}(t)) \end{bmatrix} \mathsf{G} \begin{bmatrix} u_0^{k-1}(t) \\ u_G^{k-1}(t) \end{bmatrix}\right) + q_{in}^{k-1}(t) + o\left(\frac{1}{U_0^2}\right) \qquad (5.13) \\
- \sin\theta[\overline{\lambda}^k(t) - \underline{\lambda}^k(t)],
\end{aligned}
$$

$$
\begin{aligned}
v^k(t) = \operatorname{Re}\left(e^{-j\theta}\begin{bmatrix} 0 & \mathbf{d}(\overline{u}_G^{k-1}(t)) \end{bmatrix} \mathsf{G} \begin{bmatrix} u_0^{k-1}(t) \\ u_G^{k-1}(t) \end{bmatrix}\right) + v^{k-1}(t) + \cos\theta(\overline{\lambda}^k(t) - \underline{\lambda}^k(t)) \qquad (5.14) \\
- \frac{U_0^2}{2} W^{-1}\left(\mathbf{d}(\Delta V)^{-1}(\overline{\eta}^k(t) - \underline{\eta}^k(t)) + \sum_{h=t}^{T} \mathbf{d}(\beta^{-1})(\overline{\mu}^k(h) - \underline{\mu}^k(h)) + \delta c(t)\mathbf{1}\right), \\
+ o\left(\frac{1}{U_0^2}\right), \qquad (5.15)
\end{aligned}
$$

for all $t \in \tau$.

The result above provides an update rule that can be executed by each of the nodes where some type of decision can be made, which in turn is distributed according to the sparsity of G. The proposed updating rule described by (5.12), (5.13), and (5.14)

is referred to as MICROGRID CONTROL ALGORITHM and is summarized in Algorithm 3.

Algorithm 3 The MICROGRID CONTROL ALGORITHM. Execution for node $l \in \mathcal{G}$

Set $w_l^{k-1}(t)$, $\overline{\eta}_l^{k-1}(t)$, $\underline{\eta}_l^{k-1}(t)$, $\overline{\mu}_l^{k-1}(t)$, $\underline{\mu}_l^{k-1}(t)$, $q_l^{k-1}(t)$, $v_l^{k-1}(t)$ (also $\overline{\lambda}_l^{k-1}(t)$, $\underline{\lambda}_l^{k-1}(t)$, if $l \in \mathcal{G}$), for all $t \in \tau$

for $t \in \{1, \dots, T\}$ **do**

$$\overline{\lambda}_l^k(t) = \left[\overline{\lambda}_l^{k-1}(t) + \gamma(w_l^{k-1}(t) - W_{\max})\right]_+$$
$$\underline{\lambda}_l^k(t) = \left[\underline{\lambda}_l^{k-1}(t) + \gamma(W_{\min} - w_l^{k-1}(t))\right]_+$$
$$\overline{\eta}_l^k(t) = \left[\overline{\eta}_l^{k-1}(t) + \gamma\tfrac{1}{\Delta V_l}(v_l^{k-1}(t) - V_{\text{ch}}^l)\right]_+$$
$$\underline{\eta}_l^k(t) = \left[\underline{\eta}_l^{k-1}(t) + \gamma\tfrac{1}{\Delta V_l}(V_{\text{dis}}^l - v_l^{k-1}(t))\right]_+$$
$$\overline{\mu}_l^k(t) = \left[\overline{\mu}_l^{k-1}(t) + \gamma\left(x_l(0) + \tfrac{\alpha_l}{\beta_l}\textstyle\sum_{h=1}^t v_l^{k-1}(h) - 1\right)\right]_+$$
$$\underline{\mu}_l^k(t) = \left[\underline{\mu}_l^{k-1}(t) + \gamma\left(-x_l(0) - \tfrac{\alpha_l}{\beta_l}\textstyle\sum_{h=1}^t v_l^{k-1}(h)\right)\right]_+$$

end for

Gather $\overline{\eta}_h^k(t)$, $\underline{\eta}_h^k(t)$, $\overline{\mu}_h^k(t)$, $\underline{\mu}_h^k(t)$, for all $t \in \tau$, for all $h \in \mathcal{N}_S(l) \setminus \{0\}$

Gather $u_h^{k-1}(t)$ for all $t \in \tau$, for all $h \in \mathcal{N}_S(l)$ (for all $h \in \mathcal{N}_S(l) \cup \mathcal{N}_G(l)$ if $l \in \mathcal{G}$)

for $t \in \{1, \dots, T\}$ **do**

$$v_l^k(t) = v_l^{k-1}(t) + \cos\theta(\overline{\lambda}_l^k(t) - \underline{\lambda}_l^k(t)) + \textstyle\sum_{h\in\mathcal{N}_G(l)} \mathsf{G}_{lh}\Big(\|u_l^{k-1}(t)\|_{\mathbb{C}}\|u_h^{k-1}(t)\|_{\mathbb{C}}\cos(\angle u_h^{k-1}(t) - \angle u_l^{k-1}(t) - \theta)\Big) + \tfrac{U_0^2}{2}\textstyle\sum_{h\in\mathcal{N}_G(l)\setminus\{0\}} \mathsf{G}_{lh}\Big(-\tfrac{1}{\Delta V_l}(\overline{\eta}_h^k(t) - \underline{\eta}_h^k(t)) - \tfrac{1}{\beta_h}\textstyle\sum_{b=t}^T(\overline{\mu}_h^k(b) - \underline{\mu}_h^k(b)) - \delta c(t)\Big)$$

$$q_{\text{in},l}^k(t) = q_{\text{in},l}^{k-1} - \sin\theta(\overline{\lambda}_l^k(t) - \underline{\lambda}_l^k(t)) + \textstyle\sum_{h\in\mathcal{N}_G(l)} \mathsf{G}_{lh}\Big(\|u_l^{k-1}(t)\|_{\mathbb{C}}\|u_h^{k-1}(t)\|_{\mathbb{C}}\sin(\angle u_h^{k-1}(t) - \angle u_l^{k-1}(t) - \theta)\Big)$$

end for

5.3.1 Convergence analysis of the algorithm

In order to analyze convergence of the algorithm, we introduce the following auxiliary results.

Lemma 5.5 (Optimizers and fixed points of the algorithm)**.** *The vector $\psi^\star$ is an optimizer of the dual of the* MICROGRID CONTROL PROBLEM*, and $(q_{in}^\star, v^\star)$ is an optimizer of the* MICROGRID CONTROL PROBLEM *if and only if $(\psi^\star, q_{in}^\star, v^\star)$ is a fixed point of the iteration that represents the* MICROGRID CONTROL ALGORITHM.

Proof. It is easy to show that if $(\psi^\star, q_{\text{in}}^\star, v^\star)$ is a fixed point of the dynamics in (5.12), then $(\psi^\star, q_{\text{in}}^\star, v^\star)$ satisfies the KKT conditions for the MICROGRID CONTROL PROBLEM in (5.5).

Since the cost J is continuously differentiable and the constraints are affine functions, the KKT conditions are sufficient and necessary, hence $(\psi^\star, q_{\text{in}}^\star, v^\star)$ is an optimizer of the primal-dual problem. Next, let us consider any $(\psi^\star, q_{\text{in}}^\star, v^\star)$, which is an optimizer of the MICROGRID CONTROL PROBLEM in (5.5). The KKT conditions must hold at this point. Using the complementary slackness condition onto (5.12), we can see that if $\psi^k = \psi^\star$, then $\psi^{k+1} = \psi^\star$. Further, it is easy to see from the expressions in (5.10), (5.11) that $q_{\text{in}}^{k+1} = q_{\text{in}}^k = q_{\text{in}}^\star$, and $v^{k+1} = v^k = v^\star$, meaning that $(\psi^\star, q_{\text{in}}^\star, v^\star)$ is a fixed point for the algorithm. □

Definition 5.2 (Kronecker product [56])**.** *Consider matrices $A \in \mathbb{R}^{m\times n}$, with entries a_{ij}, $B \in \mathbb{R}^{p\times q}$, with entries b_{ij}. The Kronecker product $C = A \otimes B$ returns the block matrix in $\mathbb{R}^{mp\times nq}$ such that the blocks are $(C)_{kl} = a_{kl}B$, for $k \in \{1,\dots,m\}$, $l \in \{1,\dots,n\}$.*

Definition 5.3 (Khatri-Rao product [56])**.** *Consider matrices $A \in \mathbb{R}^{m\times n}$, with entries a_{ij}, $B \in \mathbb{R}^{p\times q}$, with entries b_{ij}. Let A be block partitioned in blocks $(A)_{kl} \in \mathbb{R}^{m_k\times n_l}$, and B in blocks $(B)_{kl} \in \mathbb{R}^{p_k\times q_l}$, for $k \in \{1,\dots,K_1\}$, $l \in \{1,\dots,K_2\}$. The Khatri-Rao product $A * B$ is defined as $(A * B)_{kl} = (A)_{kl} \otimes (B)_{kl}$.*

Definition 5.4 (Matrix M)**.** *Define the matrix M as $M \triangleq M_1 * M_2 \in \mathbb{R}^{r\times r}$, where $M_1 \in \mathbb{R}^{r_1\times r_1}$, $r_1 \triangleq 6T$, is block-partitioned into 6×6 T-square blocks, $(M_1)_{kl} \in \mathbb{R}^{T\times T}$, for all $k,l \in \{1,\dots,6\}$, and $M_2 \in \mathbb{R}^{r_2\times r_2}$, with $r_2 \triangleq 6|\mathcal{G}|$, is block-partitioned into 6×6 square blocks of size $\mathcal{G}$. Let M_1 be defined as $M_1 \triangleq M_1^1 * M_1^2$, where $M_1^1 \triangleq \mathbf{1}_6(\mathbf{1}_6)^\top$ is partitioned in 2×2 blocks, i.e., with 3 block-rows and 3 block-columns, and M_1^2 is the block matrix:*

$$M_1^2 = \begin{bmatrix} I_T & I_T & \mathsf{U} \\ I_T & I_T & \mathsf{U} \\ \mathsf{L} & \mathsf{L} & \mathsf{C} \end{bmatrix}, \tag{5.16}$$

$\mathsf{U} \in \mathbb{R}^T$ is an upper triangular matrix with $\mathsf{U}_{ij} = 1$ for all $i \le j$, i.e., its diagonal entries are also one, $\mathsf{L} \in \mathbb{R}^T$ is a lower triangular matrix with $\mathsf{L}_{ij} = 1$ for all $i \ge j$, $\mathsf{C} = \mathsf{L}\mathsf{U}$. The matrix M_2 is defined as $M_2 \triangleq M_2^1 * M_2^2$, where:

$$M_2^1 \triangleq [\ 1\ \ -1\ \ -1\ \ 1\ \ -1\ \ 1\ ]^\top [\ 1\ \ -1\ \ -1\ \ 1\ \ -1\ \ 1\ ],$$

is partitioned into 2×2 *blocks, i.e., with* 3 *block-rows and* 3 *block columns, and* M_2^2 *is the block matrix:*

$$M_2^2 \triangleq \begin{bmatrix} \frac{2}{U_0^2} W & \cos\theta \mathbf{d}(\Delta V)^{-1} & \cos\theta \mathbf{d}(\beta^{-1}) \\ \cos\theta \mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2} \mathbf{d}(\Delta V)^{-1} W^{-1} \mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2} \mathbf{d}(\Delta V)^{-1} W^{-1} \mathbf{d}(\beta^{-1}) \\ \cos\theta \mathbf{d}(\beta^{-1}) & \frac{U_0^2}{2} \mathbf{d}(\beta^{-1}) W^{-1} \mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2} \mathbf{d}(\beta^{-1}) W^{-1} \mathbf{d}(\beta^{-1}) \end{bmatrix}. \tag{5.17}$$

Lemma 5.6 (Positive semidefiniteness of the Khatri-Rao product [56])**.** *Let A, B be compatibly partitioned positive semidefinite symmetric matrices, with square diagonal blocks. Then* $A * B$ *is positive semidefinite.*

Lemma 5.7 (Properties of the matrix M)**.** *For the matrix M in Definition 5.4. The following holds:*

- M *is positive semidefinite,*
- $\mathrm{null}(M)$ *has dimension* $4|\mathcal{G}|T$*,*
- $\mathrm{row}(M)$ *has dimension* $2|\mathcal{G}|T$*,*
- *there is a complete basis for* $\mathrm{null}(M)$ *given by:*

$$E \triangleq \begin{bmatrix} I_{|\mathcal{G}|T} & 0 & 0 & 0 \\ I_{|\mathcal{G}|T} & 0 & 0 & 0 \\ 0 & I_{|\mathcal{G}|T} & 0 & \mathsf{U} \otimes (\mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1})) \\ 0 & I_{|\mathcal{G}|T} & 0 & -\mathsf{U} \otimes (\mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1})) \\ 0 & 0 & I_{|\mathcal{G}|T} & -I_{|\mathcal{G}|T} \\ 0 & 0 & I_{|\mathcal{G}|T} & I_{|\mathcal{G}|T} \end{bmatrix}.$$

In the expression above, with some abuse of notation we have omitted for simplicity the dimension of the zero matrix blocks.

Proof. In order to show that M is positive semidefinite, we will show that M_i^j in Definition 5.4 is positive semidefinite, for $i, j \in \{1, 2\}$, and we use Lemma 5.6 to conclude the result. Note that M_1^1 and M_2^1 are trivially positive semidefinite. Likewise, notice that M_1^2 can be written as $M_1^2 = [\ I_T \quad I_T \quad \mathsf{U}\ ]^\top [\ I_T \quad I_T \quad \mathsf{U}\ ]$. It immediately implies that M_1^2 is positive semidefinite.

In order to show the positive semidefiniteness of M_2^2, we use the Schur complement test. This consists on checking the positive semidefiniteness on one of the Schur complements of the matrix, defined on a block partition of it [57]. Consider the partition of M_2^2 as follows:

$$M_2^2 = \left[\begin{array}{c|cc} \frac{2}{U_0^2}W & \cos\theta\mathbf{d}(\Delta V)^{-1} & \cos\theta(\mathbf{d}(\beta^{-1})) \\ \hline \cos\theta\mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2}\mathbf{d}(\Delta V)^{-1}W^{-1}\mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2}\mathbf{d}(\Delta V)^{-1}W^{-1}(\mathbf{d}(\beta^{-1})) \\ \cos\theta(\mathbf{d}(\beta^{-1})) & \frac{U_0^2}{2}(\mathbf{d}(\beta^{-1}))W^{-1}\mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2}(\mathbf{d}(\beta^{-1}))W^{-1}(\mathbf{d}(\beta^{-1})) \end{array}\right]. \quad (5.18)$$

We compute the Schur complement of $\frac{2}{U_0^2}W$ in (5.18). By definition, the Schur complement is:

$$\begin{bmatrix} \frac{U_0^2}{2}\mathbf{d}(\Delta V)^{-1}W^{-1}\mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2}\mathbf{d}(\Delta V)^{-1}W^{-1}\mathbf{d}(\beta^{-1}) \\ \frac{U_0^2}{2}\mathbf{d}(\beta^{-1})W^{-1}\mathbf{d}(\Delta V)^{-1} & \frac{U_0^2}{2}\mathbf{d}(\beta^{-1})W^{-1}\mathbf{d}(\beta^{-1}) \end{bmatrix}$$
$$-\frac{U_0^2}{2}\begin{bmatrix} \cos\theta\mathbf{d}(\Delta V)^{-1} \\ \cos\theta\mathbf{d}(\beta^{-1}) \end{bmatrix}W^{-1}\begin{bmatrix} \cos\theta\mathbf{d}(\Delta V)^{-1} & \cos\theta\mathbf{d}(\beta^{-1}) \end{bmatrix} =$$
$$(1-\cos^2\theta)\frac{U_0^2}{2}\begin{bmatrix} \mathbf{d}(\Delta V)^{-1}W^{-1}\mathbf{d}(\Delta V)^{-1} & \mathbf{d}(\Delta V)^{-1}W^{-1}\mathbf{d}(\beta^{-1}) \\ \mathbf{d}(\beta^{-1})W^{-1}\mathbf{d}(\Delta V)^{-1} & \mathbf{d}(\beta^{-1})W^{-1}\mathbf{d}(\beta^{-1}) \end{bmatrix}.$$

Since $1-\cos^2\theta$ is nonnegative, and the matrix above can be expressed as the product of the matrix $[\ \mathbf{d}(\Delta V)^{-1}W^{-1/2}\quad \mathbf{d}(\beta^{-1})W^{-1/2}\ ]^\top$ times its transpose, we conclude that M_2^2 is positive semidefinite. Since M_i^j are positive semidefinite, for $i, j \in \{1,2\}$, the result follows from Lemma 5.6.

Now, let us prove the second, third and fourth bullets.

First let us show that $\text{rank}(M) \geq 2|\mathcal{G}|T$. This follows by the construction of M. Consider a block partition of M in a 6×6 block matrix such that the block columns have $r \times |\mathcal{G}|T$ size. Further, denote the block columns of M as $(M)_j$, for $j \in \{1,\ldots,6\}$. Likewise, denote the block columns of E as $(E)_j$, $j \in \{1,\ldots,4\}$. The block columns

$(M)_1$ and $(M)_3$ can be written out as follows:

$$(M)_1 = \begin{bmatrix} I_T \otimes (\frac{2}{U_0^2} W) \\ -I_T \otimes (\frac{2}{U_0^2} W) \\ -I_T \otimes (\cos\theta \mathbf{d}(\Delta V)^{-1}) \\ I_T \otimes (\cos\theta \mathbf{d}(\Delta V)^{-1}) \\ -\mathsf{L} \otimes (\cos\theta \mathbf{d}(\beta^{-1})) \\ \mathsf{L} \otimes (\cos\theta \mathbf{d}(\beta^{-1})) \end{bmatrix}, \quad (M)_3 = \begin{bmatrix} -I_T \otimes (\cos\theta \mathbf{d}(\Delta V)^{-1}) \\ I_T \otimes (\cos\theta \mathbf{d}(\Delta V)^{-1}) \\ I_T \otimes (\frac{U_0^2}{2} \mathbf{d}(\Delta V)^{-1} W^{-1} \mathbf{d}(\Delta V)^{-1}) \\ -I_T \otimes (\frac{U_0^2}{2} \mathbf{d}(\Delta V)^{-1} W^{-1} \mathbf{d}(\Delta V)^{-1}) \\ \mathsf{L} \otimes (\cos\theta \frac{U_0^2}{2} \mathbf{d}(\beta^{-1}) W^{-1} \mathbf{d}(\Delta V)^{-1}) \\ -\mathsf{L} \otimes (\cos\theta \frac{U_0^2}{2} \mathbf{d}(\beta^{-1}) W^{-1} \mathbf{d}(\Delta V)^{-1}) \end{bmatrix}, \tag{5.19}$$

Since W^{-1} has rank $|\mathcal{G}|$, then, $I_T \otimes \frac{U_0^2}{2} W^{-1}$ has rank $|\mathcal{G}|T$. This means that $(M)_3$ has rank $|\mathcal{G}|T$. Next, notice that $(I_T \otimes (\frac{U_0^2}{2} \mathbf{d}(\Delta V)^{-1} W^{-1} \mathbf{d}(\Delta V)^{-1}))(I_T \otimes \frac{2}{U_0^2} \cos(\theta) \mathbf{d}(\Delta V) W) = I_T \otimes (\cos\theta \mathbf{d}(\Delta V)^{-1})$. Since $I_\tau \otimes (\frac{U_0^2}{2} \mathbf{d}(\Delta V)^{-1} W^{-1} \mathbf{d}(\Delta V)^{-1})$ is invertible from (5.19) we can conclude that in order for $(M)_1$ to be linearly dependent of $(M)_3$, $(I_T \otimes (\cos\theta \mathbf{d}(\Delta V)^{-1}))(I_T \otimes \frac{2}{U_0^2} \cos(\theta) \mathbf{d}(\Delta V) W)$ must be equal to $I_\tau \otimes (\frac{2}{U_0^2} W)$. However, it is equal to $I_T \otimes (\cos^2(\theta) \frac{2}{U_0^2} W)$, which means that for $\theta \neq 0$, there is no matrix X such that $(M)_3 X = (M)_1$. Finally, since W is invertible, the rank of $(M)_1$ is equal to $|\mathcal{G}|T$. Then, it follows that $\text{rank}(M) \geq 2|\mathcal{G}|T$.

Now we show that the dimension of $\text{null}(M)$ is at least $4|\mathcal{G}|T$. The reader can verify that the $(M)_j = -(M)_{j+1}$, for $j \in \{1, 3, 5\}$. Therefore it follows that $M(E)_j = 0$, $j \in \{1, \ldots, 3\}$, which means that $(E)_1$, $(E)_2$ and $(E)_3$ are formed by eigenvectors associated with zero eigenvalues. It can also be verified that $(M)_3(\mathsf{U} \otimes (\mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1}))) = -(M)_5$, and $-(M)_3(\mathsf{U} \otimes (\mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1}))) = (M)_6$. Hence, $M(E)_4 = 0$, which means that all columns of $(E)_4$ are eigenvectors of M associated with zero eigenvalues. It is also easy to verify by sparsity of E, that E has full column rank, which means that we have found $4|\mathcal{G}|T$ linearly independent eigenvectors of M with eigenvalue zero. Then, $\text{null}(M)$ has a dimension greater or equal than $4|\mathcal{G}|T$. This, along with the fact that $\text{rank}(M) = 6|\mathcal{G}|T$, ends the proof, since it implies that the sum of $\text{rank}(M)$ and the dimension of $\text{null}(M)$ is greater or equal than r. □

The following theorem establishes that the MICROGRID CONTROL ALGORITHM converges asymptotically to the optimal solution of the problem defined in (5.5), provided the parameter γ is small enough.

Theorem 5.1 (Algorithm convergence)**.** *Let assumptions 5.1, 5.2 on the input voltage magnitude and the transmission lines' impedance angle hold. Assume that the* MICROGRID CONTROL PROBLEM *in* (5.5) *is feasible. Then, for* $\gamma < \frac{2}{\rho(M)}$*, where* M *is described in Definition 5.4, the execution of the* MICROGRID CONTROL ALGORITHM *(Algorithm 3) by each node* $l \in \mathcal{G}$*, leads to* $q_{in}^k(t) \to q_{in}^\star(t)$*,* $v^k(t) \to v^\star(t)$*, for all* $t \in \tau$*, where* $(q_{in}^\star, v^\star)$ *is the unique optimizer of the* MICROGRID CONTROL PROBLEM*.*

Proof. By Lemma 5.4, we have that Algorithm 3 for each node $l \in \mathcal{G}$ is equivalent to the dynamics described by (5.10), (5.11), and (5.12). Therefore, the following analysis is performed directly on these expressions. Let $(q_{\text{in}}^\star, v^\star, \psi^\star)$ be a fixed point for the algorithm. Existence is guaranteed by assuming that the problem is feasible.

Now, let us show convergence to a fixed point. Let us define variables $y \triangleq q_{\text{in}} - q_{\text{in}}^\star$, $z \triangleq v - v^\star$, $\underline{\boldsymbol{\lambda}} \triangleq \underline{\lambda} - \underline{\lambda}^\star$, $\overline{\boldsymbol{\lambda}} \triangleq \overline{\lambda} - \overline{\lambda}^\star$, $\underline{\boldsymbol{\eta}} \triangleq \underline{\eta} - \underline{\eta}^\star$, $\overline{\boldsymbol{\eta}} \triangleq \overline{\eta} - \overline{\eta}^\star$, $\underline{\boldsymbol{\mu}} \triangleq \underline{\mu} - \underline{\mu}^\star$, $\overline{\boldsymbol{\mu}} \triangleq \overline{\mu} - \overline{\mu}^\star$. Further, define $\boldsymbol{\psi} \triangleq \psi - \psi^\star \in \mathbb{R}^r$. Notice that $\boldsymbol{\psi} \triangleq [\overline{\boldsymbol{\lambda}}^\top, \underline{\boldsymbol{\lambda}}^\top, \overline{\boldsymbol{\eta}}^\top, \underline{\boldsymbol{\eta}}^\top, \overline{\boldsymbol{\mu}}^\top, \underline{\boldsymbol{\mu}}^\top]^\top$. After some computations, we obtain that:

$$y^k(t) = -\left(\overline{\boldsymbol{\lambda}}^k(t) - \underline{\boldsymbol{\lambda}}^k(t)\right)\sin\theta, \tag{5.20}$$

$$\begin{aligned} z^k(t) = \cos\theta[\overline{\boldsymbol{\lambda}}^k(t) - \underline{\boldsymbol{\lambda}}^k(t)] - \frac{U_0^2}{2} W^{-1}\mathbf{d}(\Delta V)^{-1}[\overline{\boldsymbol{\eta}}^k(t) - \underline{\boldsymbol{\eta}}^k(t)] \\ - \frac{U_0^2}{2} W^{-1}\sum_{h=t}^{T}\mathbf{d}(\beta^{-1})[\overline{\boldsymbol{\mu}}^k(h) - \underline{\boldsymbol{\mu}}^k(h)]. \end{aligned} \tag{5.21}$$

Applying the proposed change of variables on the system described by (5.12), we obtain a dynamics that can be expressed as $\boldsymbol{\psi}^{k+1} = [f_1(\boldsymbol{\psi}^k, y^k, z^k, \psi^\star, q_{\text{in},G}^\star, v^\star)]_+ - [f_2(\psi^\star, q_{\text{in}}^\star, v^\star)]_+$, for linear maps $f_1 : \mathbb{R}^{2(r+2|\mathcal{G}|T)} \to \mathbb{R}^r$, $f_2 : \mathbb{R}^{2(r+2|\mathcal{G}|T)} \to \mathbb{R}^r$. Notice that y^k and z^k are simply linear functions of $\boldsymbol{\psi}^k$, which do not depend on past values of y, z, hence we can write the dynamics for $\boldsymbol{\psi}$ as $\boldsymbol{\psi}^{k+1} = [f_1(\boldsymbol{\psi}^k, g_1(\boldsymbol{\psi}^k), g_2(\boldsymbol{\psi}^k), \psi^\star, q_{\text{in}}^\star, v^\star)]_+ - [f_2(\psi^\star, q_{\text{in}}^\star, v^\star)]_+$, where $g_1 : \mathbb{R}^r \to \mathbb{R}^{|\mathcal{G}|T}$ is given by the right-hand side of (5.20), and $g_2 : \mathbb{R}^r \to \mathbb{R}^{|\mathcal{G}|T}$ is given by the right-hand side of (5.21). Then, consider $V(\boldsymbol{\psi}) \triangleq \|\boldsymbol{\psi}\|^2$ as a Lyapunov candidate function that can help us show that all the solutions of the dual decomposition algorithm converge to $\psi^\star$. Moreover, following the same analysis, we show that the

fixed point is unique, and by Lemma 5.5 we conclude that the optimizer of the problem is unique.

Notice that from the change of variables we introduced above, we have that:

$$\overline{\boldsymbol{\lambda}}^{k+1}(t) = \left[\overline{\lambda}^{k}(t) + \gamma(w_G^k(t) - W_{\max})\right]_+ - \left[\overline{\lambda}^{\star}(t) + \gamma(w_G^{\star}(t) - W_{\max})\right]_+,$$

for all $t \in \tau$. From the non-expansive property of the operator $[\cdot]_+$, i.e., $\|[a]_+ - [b]_+\| \leq \|a - b\|$, we obtain:

$$\|\overline{\boldsymbol{\lambda}}^{k+1}(t)\| \leq \left\| \overline{\boldsymbol{\lambda}}^{k}(t) + \gamma\left(w_G^k(t) - w_G^{\star}(t)\right) \right\|,$$

for all $t \in \tau$. Further, we replace $w_G^k(t)$ and $w_G^{\star}(t)$ according to the expression in (5.6) in order to obtain:

$$\|\overline{\boldsymbol{\lambda}}^{k+1}(t)\| \leq \left\| \overline{\boldsymbol{\lambda}}^{k}(t) + \gamma\left(\frac{2}{U_0^2}\Big(\sin\theta W(q_{\text{in}}^k(t) - q_{\text{in}}^{\star}(t)) - \cos\theta W(v^k(t) - v^{\star}(t))\Big)\right)\right\|.$$

Next, we replace $y^k(t) = q_{\text{in}}^k(t) - q_{\text{in}}^{\star}(t)$ and $z^k(t) = v^k(t) - v^{\star}(t)$, by the expression in (5.20), (5.21), derive the following equation:

$$\begin{aligned}\|\overline{\boldsymbol{\lambda}}^{k+1}(t)\| \leq \Bigg\| \overline{\boldsymbol{\lambda}}^{k}(t) + \gamma\Bigg(& - \frac{2}{U_0^2} W(\overline{\boldsymbol{\lambda}}^{k}(t) - \underline{\boldsymbol{\lambda}}^{k}(t)) \\ & + \cos\theta\left(\mathbf{d}(\Delta V)^{-1}(\overline{\boldsymbol{\eta}}^{k}(t) - \underline{\boldsymbol{\eta}}^{k}(t)) + \sum_{h=t}^{T} \mathbf{d}(\beta^{-1})(\overline{\boldsymbol{\mu}}^{k}(h) - \underline{\boldsymbol{\mu}}^{k}(h))\right)\Bigg)\Bigg\|, \end{aligned} \tag{5.22}$$

for all $t \in \tau$. This procedure can be repeated for $\underline{\boldsymbol{\lambda}}^{k+1}(t)$, $\overline{\boldsymbol{\eta}}^{k+1}(t)$,$\overline{\boldsymbol{\eta}}^{k+1}(t)$, $\overline{\boldsymbol{\mu}}^{k+1}(t)$ and $\underline{\boldsymbol{\mu}}^{k+1}(t)$, for all $t \in \tau$, obtaining:

$$\begin{aligned}\|\underline{\boldsymbol{\lambda}}^{k+1}(t)\| \leq \Bigg\| \underline{\boldsymbol{\lambda}}^{k}(t) - \gamma\Bigg(& - \frac{2}{U_0^2} W(\overline{\boldsymbol{\lambda}}^{k}(t) - \underline{\boldsymbol{\lambda}}^{k}(t)) \\ & + \cos\theta\left(\mathbf{d}(\Delta V)^{-1}(\overline{\boldsymbol{\eta}}^{k}(t) - \underline{\boldsymbol{\eta}}^{k}(t)) + \sum_{h=t}^{T} \mathbf{d}(\beta^{-1})(\overline{\boldsymbol{\mu}}^{k}(h) - \underline{\boldsymbol{\mu}}^{k}(h))\right)\Bigg)\Bigg\|, \end{aligned}$$

$$\|\overline{\boldsymbol{\eta}}^{k+1}(t)\| \leq \Bigg\| \overline{\boldsymbol{\eta}}^{k}(t) + \gamma \mathbf{d}(\Delta V)^{-1} \Bigg(\cos\theta (\overline{\boldsymbol{\lambda}}^{k}(t) - \underline{\boldsymbol{\lambda}}^{k}(t)) - \frac{U_0^2}{2} W^{-1} \left(\mathbf{d}(\Delta V)^{-1} (\overline{\boldsymbol{\eta}}^{k}(t) - \underline{\boldsymbol{\eta}}^{k}(t)) + \sum_{h=t}^{T} \mathbf{d}(\beta^{-1}) (\overline{\boldsymbol{\mu}}^{k}(h) - \underline{\boldsymbol{\mu}}^{k}(h)) \right) \Bigg) \Bigg\|, \quad (5.23)$$

$$\|\underline{\boldsymbol{\eta}}^{k+1}(t)\| \leq \Bigg\| \underline{\boldsymbol{\eta}}^{k}(t) - \gamma \mathbf{d}(\Delta V)^{-1} \Bigg(\cos\theta (\overline{\boldsymbol{\lambda}}^{k}(t) - \underline{\boldsymbol{\lambda}}^{k}(t)) - \frac{U_0^2}{2} W^{-1} \left(\mathbf{d}(\Delta V)^{-1} (\overline{\boldsymbol{\eta}}^{k}(t) - \underline{\boldsymbol{\eta}}^{k}(t)) + \sum_{h=t}^{T} \mathbf{d}(\beta^{-1}) (\overline{\boldsymbol{\mu}}^{k}(h) - \underline{\boldsymbol{\mu}}^{k}(h)) \right) \Bigg) \Bigg\|,$$

$$\|\overline{\boldsymbol{\mu}}^{k+1}(t)\| \leq \Bigg\| \overline{\boldsymbol{\mu}}^{k}(t) + \gamma \Bigg(\cos\theta \sum_{h=1}^{T} (\overline{\boldsymbol{\lambda}}^{k}(h) - \underline{\boldsymbol{\lambda}}^{k}(h)) - \frac{U_0^2}{2} \mathbf{d}(\beta^{-1}) W^{-1} \sum_{h=1}^{T} \left(\mathbf{d}(\Delta V)^{-1} (\overline{\boldsymbol{\eta}}^{k}(h) - \underline{\boldsymbol{\eta}}^{k}(h)) + \sum_{s=h}^{T} \mathbf{d}(\beta^{-1}) (\overline{\boldsymbol{\mu}}^{k}(s) - \underline{\boldsymbol{\mu}}^{k}(s)) \right) \Bigg) \Bigg\|, \quad (5.24)$$

$$\|\underline{\boldsymbol{\mu}}^{k+1}(t)\| \leq \Bigg\| \underline{\boldsymbol{\mu}}^{k}(t) - \gamma \Bigg(\cos\theta \sum_{h=1}^{T} (\overline{\boldsymbol{\lambda}}^{k}(h) - \underline{\boldsymbol{\lambda}}^{k}(h)) - \frac{U_0^2}{2} \mathbf{d}(\beta^{-1}) W^{-1} \sum_{h=1}^{T} \left(\mathbf{d}(\Delta V)^{-1} (\overline{\boldsymbol{\eta}}^{k}(h) - \underline{\boldsymbol{\eta}}^{k}(h)) + \sum_{s=h}^{T} \mathbf{d}(\beta^{-1}) (\overline{\boldsymbol{\mu}}^{k}(s) - \underline{\boldsymbol{\mu}}^{k}(s)) \right) \Bigg) \Bigg\|,$$

for all $t \in \tau$. From (5.22) through (5.24) we can write $\|\boldsymbol{\psi}^{k+1}\|^2 \leq \|(I_r - \gamma M)\boldsymbol{\psi}^k\|^2$, for M as defined in Lemma 5.7. It is straightforward to see that the eigenvalues of $I - \gamma M$ are related to the eigenvalues of M as $\lambda_i(I_r - \gamma M) = 1 - \gamma\lambda_i(M)$, where $\lambda_i(M)$ is an eigenvalue of M, with identical eigenvectors, for $i \in \{1, \ldots, r\}$. Then, by Lemma 5.7, $I_r - \gamma M$ has $r/2$ eigenvalues 1, with eigenvectors in null(M). Since M is positive semidefinite, the remaining $r/2$ eigenvalues lie in the interval $[1 - \gamma\rho(M), 1)$. Hence, with $0 < \gamma < \frac{2}{\rho(M)}$, the spectral radius $\rho(I_r - \gamma M) = 1$, but $\lambda_{\max,2} \triangle \max\{\lambda \in \text{spec}(I_r - \gamma M) \mid \lambda \neq 1\}$ is strictly less than one.

Recall that any vector $\boldsymbol{\psi}^k \in \mathbb{R}^r$ can be written as unique linear combination $\boldsymbol{\psi}^k = \boldsymbol{\psi}^k_{\text{row}} + \boldsymbol{\psi}^k_0$, where $\boldsymbol{\psi}^k_{\text{row}} \in \text{row}(M)$, $\boldsymbol{\psi}^k_0 \in \text{null}(M)$ and $\boldsymbol{\psi}^k_{\text{row}} \cdot \boldsymbol{\psi}^k_0 = 0$. Therefore, we obtain $\|\boldsymbol{\psi}^{k+1}\|^2 \leq \|(I_r - \gamma M)(\boldsymbol{\psi}^k_{\text{row}} + \boldsymbol{\psi}^k_0)\|^2$. Since M is symmetric, null(M) and row(M) are invariant under the operator M, hence $(I_r - \gamma M)\boldsymbol{\psi}^k_{\text{row}}$ lies in row(M), and $(I_r - \gamma M)\boldsymbol{\psi}^k_0 = \boldsymbol{\psi}^k_0 \in \text{null}(M)$. By orthogonality of $(I_r - \gamma M)\boldsymbol{\psi}^k_{\text{row}}$ and $\boldsymbol{\psi}^k_0$, it holds

that $\|(I_r - \gamma M)(\boldsymbol{\psi}^k_{\text{row}} + \boldsymbol{\psi}^k_0)\|^2 = \|(I_r - \gamma M)\boldsymbol{\psi}^k_{\text{row}}\|^2 + \|\boldsymbol{\psi}^k_0\|^2$. With this result, we have that $\|\boldsymbol{\psi}^{k+1}\|^2 - \|\boldsymbol{\psi}^k\|^2 \leq \|(I_r - \gamma M)\boldsymbol{\psi}^k_{\text{row}}\|^2 + \|\boldsymbol{\psi}^k_0\|^2 - \|\boldsymbol{\psi}^k\|^2 = \|(I_r - \gamma M)\boldsymbol{\psi}^k_{\text{row}}\|^2 - \|\boldsymbol{\psi}^k_{\text{row}}\|^2$. It is well known that the bound $\|(I_r - \gamma M)\boldsymbol{\psi}^k_{\text{row}}\|^2 \leq \lambda^2_{\max,2}\|\boldsymbol{\psi}^k_{\text{row}}\|^2$ holds [58]. Then, we have that $V(\boldsymbol{\psi}^{k+1}) - V(\boldsymbol{\psi}^k) = \|\boldsymbol{\psi}^{k+1}\|^2 - \|\boldsymbol{\psi}^k\|^2 \leq -(1 - \lambda^2_{\max,2})\|\boldsymbol{\psi}^k_{\text{row}}\|^2$. Using LaSalle's invariance principle, we have that any solution of the MICROGRID CONTROL ALGORITHM converges to the largest invariant set contained in $\text{null}(M) \cap \{x \in \mathbb{R}^r \mid \|x\| \leq \|\boldsymbol{\psi}^0_0\|\}$.

Next, we show that for any $\psi = \psi^\star + \boldsymbol{\psi}$ such that $\boldsymbol{\psi} \in \text{null}(M)$, it holds that $q_{\text{in}} = q^\star_{\text{in}}$, and $v = v^\star$. Notice that if $\boldsymbol{\psi}^k \in \text{null}(M)$, it can be written as $\boldsymbol{\psi}^k = E\boldsymbol{\kappa}^k$, $\boldsymbol{\kappa}^k \in \mathbb{R}^d$, where $d = 4|\mathcal{G}|T$, and E is defined in Lemma 5.7. Let us partition the vector $\boldsymbol{\kappa}^k$ according to the block partition of E, as $\boldsymbol{\kappa}^k = [(\boldsymbol{\kappa}^k_1)^\top, (\boldsymbol{\kappa}^k_2)^\top, (\boldsymbol{\kappa}^k_3)^\top, (\boldsymbol{\kappa}^k_4)^\top]^\top$, where we concisely denote $\boldsymbol{\kappa}^k_1 = [\boldsymbol{\kappa}^k_1(1)^\top, \ldots, \boldsymbol{\kappa}^k_1(T)^\top]^\top$, $\boldsymbol{\kappa}^k_2 = [\boldsymbol{\kappa}^k_2(1)^\top, \ldots, \boldsymbol{\kappa}^k_2(T)^\top]^\top$, $\boldsymbol{\kappa}^k_3 = [\boldsymbol{\kappa}^k_3(1)^\top, \ldots, \boldsymbol{\kappa}^k_3(T)^\top]^\top$, and $\boldsymbol{\kappa}^k_4 = [\boldsymbol{\kappa}^k_4(1)^\top, \ldots, \boldsymbol{\kappa}^k_4(T)^\top]^\top$, with $\boldsymbol{\kappa}^k_1(t), \boldsymbol{\kappa}^k_2(t), \boldsymbol{\kappa}^k_3(t), \boldsymbol{\kappa}^k_4(t) \in \mathbb{R}^{|\mathcal{G}|}$, for all $t \in \tau$.

Given the structure of E, it is easy to verify that $\overline{\boldsymbol{\lambda}}^k(t) = \underline{\boldsymbol{\lambda}}^k(t) = \boldsymbol{\kappa}^k_1(t)$, $\overline{\boldsymbol{\eta}}^k(t) = \boldsymbol{\kappa}^k_2(t) + \sum_{h=t}^T \mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1})\boldsymbol{\kappa}^k_4(h)$, $\underline{\boldsymbol{\eta}}^k(t) = \boldsymbol{\kappa}^k_2(t) - \sum_{h=t}^T \mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1})\boldsymbol{\kappa}^k_4(h)$, $\overline{\boldsymbol{\mu}}^k(t) = \boldsymbol{\kappa}^k_3(t) - \boldsymbol{\kappa}^k_4(t)$ and $\overline{\boldsymbol{\mu}}^k(t) = \boldsymbol{\kappa}^k_3(t) + \boldsymbol{\kappa}^k_4(t)$, for all $t \in \tau$. We plug these values in (5.10) and (5.11) to obtain:

$$\begin{aligned} q^k_{\text{in}}(t) &= q_{\text{load},G}(t) + W^{-1}Fq_{\text{load},L}(t) - \sin\theta[\overline{\lambda}^\star(t) + \boldsymbol{\kappa}^k_1(t) - (\underline{\lambda}^\star(t) + \boldsymbol{\kappa}^k_1(t))] \\ &= q^\star_{\text{in}}(t), \end{aligned}$$

$$\begin{aligned} v^k(t) =& - p_{\text{load},G}(t) - p_{\text{in},G}(t) - W^{-1}Fp_{\text{load},L}(t) \\ &+ \cos\theta(\overline{\lambda}^\star(t) + \boldsymbol{\kappa}^k_1(t) - (\underline{\lambda}^\star(t) + \boldsymbol{\kappa}^k_1(t))) \\ &- \frac{U_0^2}{2}W^{-1}\Bigg(\mathbf{d}(\Delta V)^{-1}\Big(\overline{\eta}^\star(t) + \boldsymbol{\kappa}^k_2(t) + \sum_{h=t}^T \mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1})\boldsymbol{\kappa}^k_4(h)\Big) \\ &- \mathbf{d}(\Delta V)^{-1}\Big(\underline{\eta}^\star(t) + \boldsymbol{\kappa}^k_2(t) - \sum_{h=t}^T \mathbf{d}(\Delta V)\mathbf{d}(\beta^{-1})\boldsymbol{\kappa}^k_4(h)\Big) \\ &+ \sum_{h=t}^T \mathbf{d}(\beta^{-1})\left(\overline{\mu}^\star(h) + \boldsymbol{\kappa}^k_3(t) + \boldsymbol{\kappa}^k_4(t) - (\underline{\mu}^\star(h) + \boldsymbol{\kappa}^k_3(t) - \boldsymbol{\kappa}^k_4(t))\right) + \delta c(t)\mathbf{1}\Bigg) \\ =& v^\star(t), \end{aligned}$$

for all $t \in \tau$. Then, since $\psi^k - \psi^\star \to \text{null}(M)$ as $k \to +\infty$, it follows that $v^k \to v^\star$, $q_{\text{in}}^k \to q_{\text{in}}^\star$ as $k \to +\infty$. □

5.3.2 Study of the rate of convergence

Although the dual decomposition algorithm has been shown to converge to the optimal solution of the MICROGRID CONTROL PROBLEM, simulations show that the convergence rate of the algorithm is prohibitively slow.

Consider the implementation of the MICROGRID CONTROL ALGORITHM on a single-phase approximation of the IEEE 37 standard model, with the same location of generators as in [6]. In addition, we add 4 nodes that have only storage capacity and no generation. The complete list of the simulation description and parameters, as well as the commented simulation code we have used, can be found at
`http://fausto.dynamic.ucsd.edu/andres/project_reactive.html`.

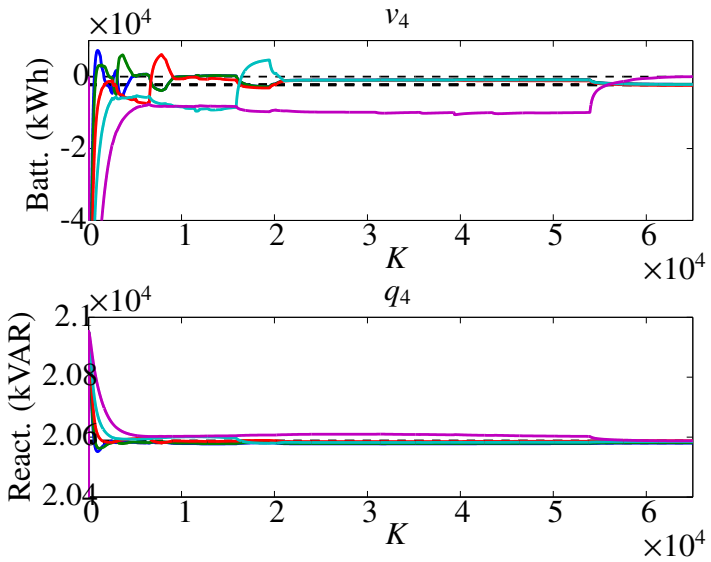

Figure 5.1: a) Evolution of v_4, b) evolution of q_4, as a function of the iteration number K, respectively, for each $t \in \tau$. Dashed lines show the optimal values.

Figure 5.1 shows the evolution of the decision variables for the whole time horizon, for node 4. Dashed lines represent the optimal values for the decision variables (presenting some overlap). It can be seen that the algorithm leads the decision variables to their optimizers.

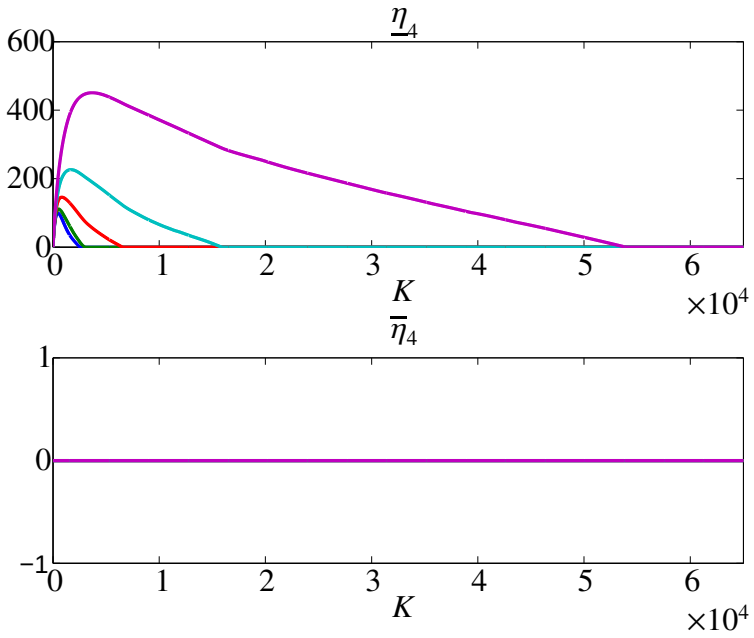

Figure 5.2: Evolution of η_4, as a function of the iteration number K, respectively, for each $t \in \tau$.

Notice that there is a remarkable difference between the amount of iterations that it takes the variable q_{in} to converge and those for the variable v. There are two main observed reasons for this difference: the first one is that the amount of local constraints related to the variable v is very large as compared to those affecting q_{in}. The second reason is that the geometry of the feasible set given by the local constraints on v activates some multipliers that in their optimal state should be zero. In particular, we observe a very fast and large growth in the $\underline{\eta}$ multipliers. Once the satisfaction of the $\underline{\mu}$ constraints guarantee the satisfaction of the $\underline{\eta}$ constraints, the components of $\underline{\eta}$ start to decrease. However, the nonlinear dynamics for the multipliers do not allow for a fast decrease rate. It can be seen in Figures 5.2, 5.3 that the $\underline{\eta}$, $\underline{\mu}$ multipliers grow very fast, however, the multipliers $\underline{\eta}$ eventually start a very slow decrease to end up reaching the optimal value 0. This slow decrease leads to very large number of iterations for convergence, which may affect the possibility of using the algorithm in applications with short discretization steps. Since the parameter γ is near the limit for stability, the speed of convergence cannot be significantly improved. We have tested the speed of execution of several

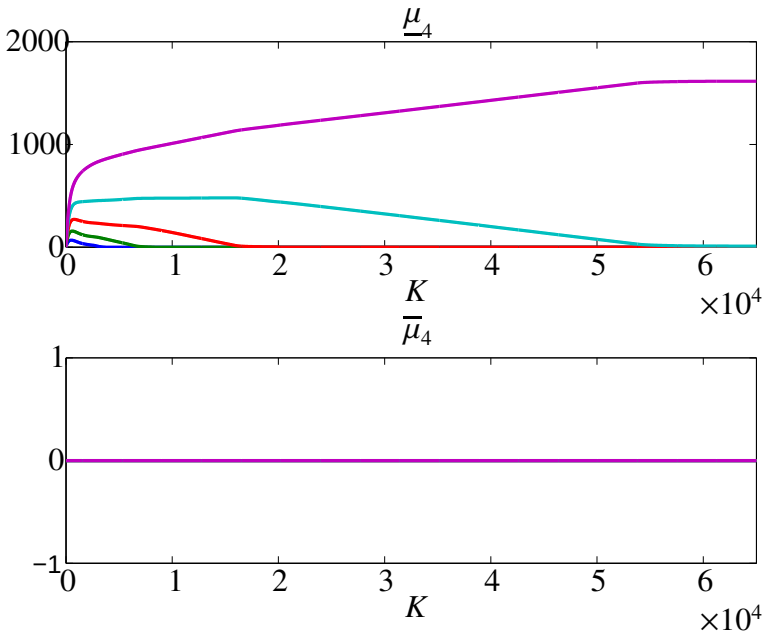

Figure 5.3: Evolution of μ_4 as a function of the iteration number K, respectively, for each $t \in \tau$.

iterations of the algorithm in a computer with processor 2.8GHz Intel Core i7, and 8GB 1333 MHz RAM, and on average, each iteration takes 0.0364 seconds. Of course, the computations for all nodes have been performed in a sequential manner, while they are meant to be performed in a parallel way. Nevertheless, the simulation for 60×10^3 iterations that is shown in the case study required more than 36 minutes to run. This in turn means that if the simulation was carried out in a parallel way, it would have taken at least 4 minutes.

In practice, it implies that: i) a large amount of communication between neighbors will be required to achieve a reasonable solution, and ii) the time it takes to compute an optimal solution will not allow to perform the computation periodically in order to account for the rapid changes in solar/wind generation.

5.4 The MICROGRID OPTIMIZATION ALGORITHM

In the previous Section we introduced a Dual Decomposition Algorithm for the computation of control of storage and reactive power in a microgrid. Unfortunately, it turned out to be too slow if we want to perform a periodic recomputation of the control, given the rapid variability of renewable generation.

In order to circumvent this limitation, we propose a combination of the dual decomposition algorithm to handle the coupling constraints, along with a component-wise minimization and local projection to handle local constraints.

Let us divide the constraints of our problem into two groups: *global constraints*, which are the ones related to node voltages, in (5.7) and *local constraints*, which are the ones related to the storage and reactive power generation, in (5.5b), (5.5c), and (5.5d). Let us define sets $\mathcal{U}_\ell \triangleq \{v_\ell \mid \text{Constraints (5.5c), (5.5d) hold}\}$ and also $\mathcal{Q}_\ell \triangleq \{q_{\text{in},\ell} \mid \text{Constraint (5.5b) holds}\}$ for each $\ell \in \mathcal{G}$. Further, define $\mathcal{U} \triangleq \bigtimes_{\ell\in\mathcal{G}} \mathcal{U}_\ell$ and $\mathcal{Q} \triangleq \bigtimes_{\ell\in\mathcal{G}} \mathcal{Q}_\ell$.

With a slight abuse of notation, we redefine $\mathcal{L}$ as a Lagrangian function for the problem in (5.5), taking into account only the global constraints:

$$\mathcal{L}(q_{\text{in}}, v, \lambda) \triangleq J(q_{\text{in}}, v) + \sum_{t=1}^{T} \underline{\lambda}^T(t)(W_{\min} - w_G(t)) + \sum_{t=1}^{T} \overline{\lambda}^T(t)(w_G(t) - W_{\max}), \tag{5.25}$$

where $\underline{\lambda}(t), \overline{\lambda}(t) \geq 0$ for all $t \in \tau$, are Lagrange multipliers associated to the constraints (5.5a).

Define the function $(q_{\text{in}}, v) \mapsto \mathcal{L}_\lambda(q_{\text{in}}, v)$ as $\mathcal{L}_\lambda(q_{\text{in},,}v) \triangleq \mathcal{L}(q_{\text{in}}, v, \lambda)$. Given that $\mathcal{L}_\lambda$ is a quadratic form of (q_{in}, v) and the matrix W is positive definite, $\mathcal{L}_\lambda$ has a unique minimizer that can be found directly by solving $\nabla_{q_{\text{in}},v}\mathcal{L}_\lambda(q_{\text{in}}, v) = 0$. Note that, by definition, minimizing $\mathcal{L}_\lambda(q_{\text{in}}, v)|_{\lambda=\lambda^k}$ is the primal step of the dual decomposition approach.

The dual decomposition iteration given the Lagrangian in (5.25) will converge to an optimal solution of the problem without local constraints. A direct projection of such solution onto $\mathcal{U} \times \mathcal{Q}$ will not be optimal unless the quadratic cost function has spherical level sets. Hence, we can not use this on our problem directly, as it will not lead to the optimal value of v, q_{in} in general. Instead of a direct projection, we propose that each node performs an approximate step in the direction of the minimizer of $\mathcal{L}$,

parameterized by the value of the decision variables that other nodes have.

To explain this idea, let us write v as (v_l, v_{-l}), with $v_l = [v_l(1), \ldots, v_l(T)]^\top$, for each $l \in \mathcal{G}$, and v_{-l} defined as the concatenation of all v_h, for $h \in \mathcal{G} \setminus \{l\}$. Likewise, let us write q_{in} as $(q_{\text{in},l}, q_{\text{in},-l})$, for each $l \in \mathcal{G}$.

Then, for the procedure that we propose, each node $l \in \mathcal{G}$, computes:

$$\overline{v}_l^{k+1} = (1-\eta)v_l^k + \eta \operatorname{argmin}_{v_l} \mathcal{L}(q_{\text{in}}^k, v_l, v_{-l}^k, \lambda^k), \tag{5.26}$$

$$\overline{q}_{\text{in},l}^{k+1} = (1-\eta)q_{\text{in},l}^k + \eta \operatorname{argmin}_{q_{\text{in},l}} \mathcal{L}(q_{\text{in},l}, q_{\text{in},-l}^k, v^k, \lambda^k), \tag{5.27}$$

where $\eta \in (0, 1)$,

In order to compute the expressions above, we remove from the Lagrangian the summands that do not depend on v_l for $\operatorname{argmin}_{v_l} \mathcal{L}(q_{\text{in}}^k, v_l, v_{-l}^k, \lambda^k)$, and the ones that do not depend on $q_{\text{in},l}$ for $\operatorname{argmin}_{q_{\text{in},l}} \mathcal{L}(q_{\text{in},l}, q_{\text{in},-l}^k, v^k, \lambda^k)$. This leads to modified expressions for the Lagrangian as follows:

$$\tilde{\mathcal{L}}(v, \lambda) = (v^\top (W \otimes I_T)v + 2\xi_v^\top v) + \delta(\mathbf{1} \otimes c)^\top v + (\overline{\lambda} - \underline{\lambda})^\top \Big(- \frac{2}{U_0^2} \cos\theta W \otimes I_T \Big) v,$$

where $\xi_v \in \mathbb{R}^{|\mathcal{G}|T}$ is defined as $\xi_v = [\xi_{v,1}^\top, \ldots, \xi_{v,|\mathcal{G}|}^\top]^\top$, and:

$$\tilde{\mathcal{L}}(q_{\text{in}}, \lambda) = (q_{\text{in}}^\top (W \otimes I_T) q_{\text{in}} + 2\xi_q^\top q_{\text{in}}) + (\overline{\lambda} - \underline{\lambda})^\top \Big(\frac{2}{U_0^2} \sin\theta W \otimes I_T \Big) q_{\text{in}},$$

where $\xi_q \in \mathbb{R}^{|\mathcal{G}|T}$ is defined as $\xi_q = [\xi_{q,1}^\top, \ldots, \xi_{q,|\mathcal{G}|}^\top]^\top$. Note that $\operatorname{argmin}_{v_l} \mathcal{L}(q_{\text{in}}^k, v_l, v_{-l}^k, \lambda^k)$ is the solution of $\nabla_{v_l} \tilde{\mathcal{L}}(v_l, v_{-l}^k, \lambda^k) = 0$. Likewise, $\operatorname{argmin}_{q_{\text{in},l}} \mathcal{L}(q_{\text{in},l}, q_{\text{in},-l}^k, v^k, \lambda^k)$ is the solution of $\nabla_{q_{\text{in},l}} \tilde{\mathcal{L}}(q_{\text{in},l}, q_{\text{in},-l}^k, \lambda^k) = 0$, where:

$$\begin{aligned} \nabla_{v_l} \tilde{\mathcal{L}}(v_l, v_{-l}^k, \lambda^k) = & 2W_{ll} I_T v_l + 2(W_l \otimes I_T v^k - W_{ll} I_T v_l^k) + \frac{2}{U_0^2} (\xi_{v,l} - \cos\theta W_l \otimes I_T (\overline{\lambda}^k - \underline{\lambda}^k)) \\ & + \delta c, \end{aligned}$$

and:

$$\nabla_{q_{\text{in},l}}\tilde{\mathcal{L}}(q_{\text{in},l}, q^k_{\text{in},-l}, \lambda^k) = 2W_{ll}I_T q_{\text{in},l} + 2(W_l \otimes I_T q^k_{\text{in}} - W_{ll}I_T q^k_{\text{in},l}) + \frac{2}{U_0^2}(\xi_{q,l} + \sin\theta W_l \otimes I_T(\overline{\lambda}^k - \underline{\lambda}^k)).$$

It is also important to notice that $\nabla_{v_l}\tilde{\mathcal{L}}(v_l, v^k_{-l}, \lambda^k) = 0$ and $\nabla_{q_{\text{in},l}}\tilde{\mathcal{L}}(q_{\text{in},l}, q^k_{\text{in},-l}, \lambda^k) = 0$ for all $l \in \mathcal{G}$ can be written in compact form as:

$$\mathbf{d}(W) \otimes I_T v + ((W - \mathbf{d}(W)) \otimes I_T)v^k + \xi_v - \frac{1}{U_0^2}\cos\theta W \otimes I_T(\overline{\lambda}^k - \underline{\lambda}^k) + \frac{1}{2}\delta(\mathbf{1} \otimes c) = 0, \tag{5.28}$$

and

$$\mathbf{d}(W) \otimes I_T q_{\text{in}} + ((W - \mathbf{d}(W)) \otimes I_T)q^k_{\text{in}} + \xi_q + \frac{1}{U_0^2}\sin\theta W \otimes I_T(\overline{\lambda}^k - \underline{\lambda}^k) = 0, \tag{5.29}$$

respectively. From (5.26), (5.28), we obtain:

$$\overline{v}^{k+1} = (1-\eta)v^k + \eta(\mathbf{d}(W)^{-1} \otimes I_T)\Big(- \xi_v - ((W - \mathbf{d}(W)) \otimes I_T)v^k + \frac{1}{U_0^2}\cos\theta W \otimes I_T(\overline{\lambda}^k - \underline{\lambda}^k) - \frac{1}{2}\delta(\mathbf{1} \otimes c)\Big), \tag{5.30}$$

and from (5.27), (5.29), we obtain:

$$\overline{q}^{k+1}_{\text{in}} = (1-\eta)q^k_{\text{in}} + \eta(\mathbf{d}(W)^{-1} \otimes I_T)\Big(- \xi_q - ((W - \mathbf{d}(W)) \otimes I_T)q^k_{\text{in}} - \frac{1}{U_0^2}\sin\theta W \otimes I_T(\overline{\lambda}^k - \underline{\lambda}^k))\Big). \tag{5.31}$$

Finally, we perform a projection in the feasible set of v and q_{in}, i.e., $v^{k+1} = \text{Proj}_{\mathcal{U}}\{\overline{v}^{k+1}\}$ and $q^{k+1}_{\text{in}} = \text{Proj}_{\mathcal{Q}}\{\overline{q}^{k+1}_{\text{in}}\}$. By construction, $\mathcal{U} = \bigtimes_{l\in\mathcal{G}}\mathcal{U}_l$ and $\mathcal{Q} = \bigtimes_{l\in\mathcal{G}}\mathcal{Q}_l$. Then, the

operation can be carried out in a distributed way:

$$v_l^{k+1} = \text{Proj}_{\mathcal{U}_l}\{\overline{v}_l^{k+1}\}, \tag{5.32a}$$

$$q_{\text{in},l}^{k+1} = \text{Proj}_{\mathcal{Q}_l}\{\overline{q}_{\text{in},l}^{k+1}\}, \qquad \forall l \in \mathcal{G}. \tag{5.32b}$$

The gradient ascent for the dual variables is given by:

$$\begin{aligned} \overline{\lambda}^{k+1}(t) &= \left[\overline{\lambda}^{k}(t) + \eta\gamma(w_G^k(t) - W_{\max})\right]_+, \\ \underline{\lambda}^{k+1}(t) &= \left[\underline{\lambda}^{k}(t) + \eta\gamma(W_{\min} - w_G^k(t))\right]_+, \end{aligned} \tag{5.33}$$

for all $t \in \tau$, with $\gamma > 0$.

In order to analyze this algorithm, hereinafter referred to as MICROGRID OPTIMIZATION ALGORITHM, we present a general quadratic problem with a structure that includes that of the MICROGRID CONTROL PROBLEM.

Our problem of interest is a particular case of the following PARTIALLY SEPARABLE QUADRATIC PROBLEM, expressed as:

$$\begin{aligned} &\min_z \; z^\top A z + b^\top z \\ &\text{subject to} \\ &Mz \preceq h \\ &z \in Z \end{aligned} \tag{5.34}$$

for $z \in \mathbb{R}^n$, $h \in \mathbb{R}^{\kappa_1}$, $A \in \mathbb{R}^{n\times n}$ symmetric positive definite, $M \in \mathbb{R}^{\kappa_1\times n}$, and a convex polytope Z, which can be expressed as the set of points z such that $Gz \preceq r$, for $r \in \mathbb{R}^{\kappa_2}$ and a matrix G. From now on, let us assume that the problem is feasible. Therefore, since all the inequality constraints are affine, the PARTIALLY SEPARABLE QUADRATIC PROBLEM satisfies the Slater's condition.

Let the components of z be assigned to a set of agents $I \triangleq \{1, \ldots, N\}$. In particular, consider that $\mathcal{I}_i \triangleq \{i_p\}_{p=1}^{n_i}$ be the indices of those components of z associated to agent i, for each $i \in \{1, \ldots, N\}$. where $\sum_{i=1}^N n_i = n$. Let z_i encompass all the components of z with indices in $\mathcal{I}_i$. In addition, let Z be also separable according to the set of agents

I, i.e., $Z \triangleq \bigtimes_{i=1}^{N} Z_i$, where $Z_i \subset \mathbb{R}^{n_i}$ is a convex polytope, for all $i \in I$.

Assumption 5.3. *(Structure of the* PARTIALLY SEPARABLE QUADRATIC PROBLEM*): Let us write the vector z_i as $z_i = [z_{i,1}^\top, \ldots, z_{i,m_i}^\top]^\top$ where each $z_{i,\ell}$ is a vector with components of z_i, $\ell \in \{1, \ldots, m_i\}$, for all $i \in I$. Let us assume:*

- *The matrix A has a structure such that the quadratic form $z^\top A z$ satisfies:*

$$z^\top A z = \sum_{\ell=1}^{m_i} a_{i,\ell} \|z_{i,\ell}\|^2 + 2 z_i^\top A_{i,12} z_{-i} + z_{-i}^\top A_{i,22} z_{-i},$$

  where $a_{i,\ell}$ is a real scalar, $A_{i,12}$ is composed by all entries A_{kl} such that $k \in \mathcal{I}_i$ and $l \notin \mathcal{I}_i$ and $A_{i,22}$ is composed by those entries A_{kl} such that $k, l \notin \mathcal{I}_i$, for all $i \in \{1, \ldots, N\}$. Recall that z can be written as $z = (z_i, z_{-i})$ for all $i \in I$, where z_{-i} are the components of z that are not associated to agent $i \in I$.

- *The set Z_i is separable in terms of $z_{i,\ell}$, for $\ell \in \{1, \ldots, m_i\}$, i.e., $Z_i = \bigtimes_{\ell=1}^{m_i} Z_{i,\ell}$, and $z_{i,\ell} \in Z_{i,\ell}$.*

$\diamond$

The first part of the assumption implies that there are no cross terms between any two entries of z associated to the same agent $i \in I$. As an example, any quadratic form $z^\top \mathsf{C} z$ with C symmetric, positive definite matrix $\mathsf{C} = C \otimes I_T$, where $C \in \mathbb{R}^{n \times n}$, with $z = [z_1^\top, \ldots, z_N^\top]^\top$, and $z_i \in \mathbb{R}^T$ for all $i \in \{1, \ldots, N\}$, satisfies the first statement of the assumption.

The Lagrangian function associated to the PARTIALLY SEPARABLE QUADRATIC PROBLEM is:

$$\mathcal{L}(z, \psi, \theta) = z^\top A z + b^\top z + \psi^\top (M z - h) + \theta^\top (G z - r),$$

where $\psi \geq 0$ and $\theta \geq 0$ are dual variables for the problem. Let $\psi^\star, \theta^\star$ be an optimizer of the dual of the PARTIALLY SEPARABLE QUADRATIC PROBLEM:

$$\max_{\psi, \theta \geq 0} \min_{z \in \mathbb{R}^n} \mathcal{L}(z, \psi, \theta).$$

Then, since the PARTIALLY SEPARABLE QUADRATIC PROBLEM is convex and satisfies the Slater's condition, it can be recast as:

$$\begin{aligned} &\min_{z} f(z, \psi^\star) \\ &\text{subject to} \\ &z \in Z, \end{aligned} \tag{5.35}$$

where:

$$f(z, \psi) \triangleq z^\top A z + b^\top z + \psi^\top (Mz - h). \tag{5.36}$$

Next, we write a general form of the dynamics, which we will call the CONSTRAINT SEPARATION DYNAMICS:

$$\psi^{k+1} = [\psi^k + \eta\gamma(Mz^k - h)]_+, \tag{5.37a}$$

$$z_i^{k+1} = \text{Proj}_{Z_i}[(1-\eta)z_i^k + \eta \,\text{argmin}_{z_i \in \mathbb{R}^{n_i}} f((z_i, z_{-i}^k), \psi^k)], \qquad \forall i \in I. \tag{5.37b}$$

Given the definition of f in (5.36), we can write this dynamics in a compact form as follows:

$$\psi^{k+1} = [\psi^k + \eta\gamma(Mz^k - h)]_+, \tag{5.38a}$$

$$z^{k+1} = \text{Proj}_Z[z^k - \eta \mathbf{d}(A)^{-1}(Az^k + \frac{1}{2}b + \frac{1}{2}M^\top \psi^k)]. \tag{5.38b}$$

Lemma 5.8. *For $z^\star = \text{argmin}_{y \in Z} f(y, \psi^\star)$, with f defined in Equation* (5.36)*, where $Z = \bigtimes_i Z_i$, it holds that $z_i^\star = \text{argmin}_{y_i \in Z_i} f(y_i, z_{-i}^\star, \psi^\star)$, for all $i \in \{1, \ldots, N\}$.*

Proof. Let us proceed by contradiction. Assume that for some $i \in \{1, \ldots, N\}$, $z_i^\star \neq \text{argmin}_{y_i \in Z_i} f(y_i, z_{-i}^\star, \psi^\star)$. Then, there exists $y_i \in Z_i$ such that $f(y_i, z_{-i}^\star, \psi^\star) < f(z_i^\star, z_{-i}^\star, \psi^\star)$, which contradicts the fact that $x^\star$ is an optimizer of $f(\cdot, \psi^\star)$ in Z. □

Proposition 5.1. *If Assumption 5.3 which describes a structure for the* PARTIALLY SEPARABLE

QUADRATIC PROBLEM *holds, then:*

$$\begin{aligned}&\operatorname{argmin}_{y_i \in Z_i} (y_i, z_{-i})^\top A(y_i, z_{-i}) + b^\top (y_i, z_{-i}) + \psi^\top (M(y_i, z_{-i}) - h) \\ &\quad = \operatorname{Proj}_{Z_i}[\operatorname{argmin}_{y_i \in \mathbb{R}^{n_i}} ((y_i, z_{-i})^\top A(y_i, z_{-i}) + b^\top (y_i, z_{-i}) + \psi^\top (M(y_i, z_{-i}) - h))].\end{aligned}$$

Proof. From Assumption 5.3, we have that:

$$z^\top A z + b^\top z + \psi^\top (Mz - h) = \sum_{\ell=1}^{m_i} a_\ell \|z_{i,\ell}\|^2 + \varphi^\top z_i + s - \psi^\top h,$$

where s is a bilinear function of z_{-i} and φ is a linear function of z_{-i} and ψ. Then, the statement of the proposition can be formulated as:

$$\operatorname{argmin}_{y_i \in Z_i} \sum_{\ell=1}^{m_i} a_\ell \|y_{i,\ell}\|^2 + \varphi^\top y_i + s - \psi^\top h = \operatorname{Proj}_{Z_i}[\operatorname{argmin}_{y_i \in \mathbb{R}^{n_i}} \sum_{\ell=1}^{m_i} a_\ell \|y_{i,\ell}\|^2 + \varphi^\top y_i + s]. \tag{5.39}$$

Let us write out $\varphi = [\varphi_1, \ldots, \varphi_{m_i}]^\top$, such that $\varphi^\top y_i = \sum_{\ell=1}^{m_i} \varphi_\ell^\top y_{i,\ell}$. By structure of Z_i, which is component-wise decomposable, (5.39) is equivalent to saying that:

$$\operatorname{argmin}_{y_{i,\ell} \in Z_{i,\ell}} a_\ell \|y_{i,\ell}\|^2 + \varphi_\ell^\top y_{i,\ell} = \operatorname{Proj}_{Z_{i,\ell}}[\operatorname{argmin}_{y_{i,\ell}} a_\ell \|y_{i,\ell}\|^2 + \varphi_\ell^\top y_{i,\ell}],$$

for all $\ell \in \{1, \ldots, m_i\}$. In order to show this equivalence, we have that:

$$\begin{aligned}&\operatorname{Proj}_{Z_{i,\ell}}[\operatorname{argmin}_{y_{i,\ell}} a_\ell \|y_{i,\ell}\|^2 + \varphi_\ell^\top y_{i,\ell}] \\ &\quad = \operatorname{argmin}_{q \in Z_{i,\ell}} \left(-\frac{1}{2a_\ell}\varphi_\ell - q\right)^\top \left(-\frac{1}{2a_\ell}\varphi_\ell - q\right) \\ &\quad = \operatorname{argmin}_{q \in Z_{i,\ell}} \left(\frac{\varphi_\ell^\top \varphi_\ell}{4a_\ell} + \frac{1}{a_\ell}\varphi_\ell^\top q + q^\top q\right) \\ &\quad = \operatorname{argmin}_{q \in Z_{i,\ell}} \varphi_\ell^\top q + a_\ell q^\top q.\end{aligned}$$

The first equality comes from the definition of the projection operator and the fact that $\operatorname{argmin}_{y_{i,\ell}} [a_\ell \|y_{i,\ell}\|^2 + \varphi_\ell^\top y_{i,\ell}] = -\frac{1}{2a_\ell}\varphi_\ell$. The last equality simply comes from removing terms that do not depend on the decision variable, and use a scaling of the cost function

by a positive constant, completing the proof. □

Lemma 5.9. *Let $Y \subset \mathbb{R}^n$ be a convex closed polytope. Consider a point $y \notin Y$. Let $\tilde{y}$ be the projection of y in Y, i.e., $\tilde{y} = \text{Proj}_Y[y]$. Define $q = (1-\eta)y + \eta\tilde{y}$, for $\eta \in [0,1]$. Then, $\tilde{y} = \text{Proj}_Y[q]$.*

Proof. Assume that $\tilde{q} = \text{Proj}_Y[q] \neq \tilde{y}$. By definition of q as a convex combination of y and $\tilde{y}$, it holds that $\|y - \tilde{y}\| = \|y - q\| + \|q - \tilde{y}\|$. Then, $\|y - \tilde{y}\| \geq \|y - q\| + \|q - \tilde{q}\|$. By the triangular inequality, if follows that $\|y - q\| + \|q - \tilde{q}\| \geq \|y - \tilde{q}\|$, then $\|y - \tilde{y}\| \geq \|y - \tilde{q}\|$, which contradicts the fact that $\tilde{y} = \text{argmin}_{\phi \in Y} \|\phi - y\|$, which follows from the definition of the projection operator. □

Remark 5.1. *We would like to notice that if we apply the primal-dual subgradient method introduced in [59] to the problem* (5.35)*, we obtain a dynamics that is similar to the* CONSTRAINT SEPARATION DYNAMICS*. However, the method in [59] is intended to optimize general convex cost functions. The proposed algorithm therein is only guaranteed to converge to a neighborhood of the optimizer, whose size depends on the step size chosen for the dynamics. The result in [59, Proposition 5.1] presents an upper bound for the cost at each iteration in terms of the optimal cost, the step size, and the iteration number. Unlike the aforementioned paper, we present convergence result for the* CONSTRAINT SEPARATION DYNAMICS*, and characterize the step size required for convergence to the optimizer of the problem* (5.35)*. Moreover, in Assumption 5.3 we characterize the type of quadratic cost functions that allow us to perform a distributed computation of this dynamics.* ◇

Corollary 5.1. *(Convergence of the* MICROGRID OPTIMIZATION ALGORITHM*): The* MICROGRID OPTIMIZATION ALGORITHM *converges to the optimal solution of the* MICROGRID CONTROL PROBLEM*.*

Proof. It follows directly by defining $z = [v^\top, q_{\text{in}}^\top]^\top$.

$$A = \begin{bmatrix} W \otimes I_T & 0 \\ 0 & W \otimes I_T \end{bmatrix}, \quad b = \begin{bmatrix} 2\xi_v + \delta \mathbf{1} \otimes c \\ 2\xi_q \end{bmatrix}$$

$\psi = [\overline{\lambda}^\top, \underline{\lambda}^\top]^\top$, and $Z = \Xi$. It is easy to see that the problem fulfills Assumption 5.3, where $z_i = [v_i^\top, q_{\text{in},i}^\top]^\top$, for all $i \in \mathcal{G}$. □

5.4.1 Handling the non-distributed terms

Still, we can see that $\overline{v}^{k+1}$ and $\overline{q}_{\text{in}}^{k+1}$ must be computed in terms of the matrix W respectively, which is by no means sparse in the communication network. Then, $\overline{v}^{k+1}$ and $\overline{q}_{\text{in}}^{k+1}$ cannot be computed in a distributed way using the expressions (5.30) and (5.31).

To go around this problem, we define μ_k^v, μ_k^q, ζ_k^v, and ζ_k^q as follows:

$$\begin{aligned}
\mu_k^v =& \frac{1}{\eta}(\mathbf{d}(W) \otimes I_T)(\overline{v}^{k+1} - v^k) + \frac{1}{2}\delta(\mathbf{1} \otimes c),\\
\zeta_k^v =& - v^k - (W^{-1} \otimes I_T)\xi_v + \frac{1}{U_0^2}\cos\theta(\overline{\lambda}^k - \underline{\lambda}^k),\\
\mu_k^q =& \frac{1}{\eta}(\mathbf{d}(W) \otimes I_T)(\overline{q}_{\text{in}}^{k+1} - q_{\text{in}}^k),\\
\zeta_k^q =& - q_{\text{in}}^k - (W^{-1} \otimes I_T)\xi_v - \frac{1}{U_0^2}\sin\theta(\overline{\lambda}^k - \underline{\lambda}^k).
\end{aligned} \tag{5.40}$$

Then, we can recast the equalities in (5.30) and (5.31) as:

$$\begin{aligned}
\mu_k^v &= (W \otimes I_T)\zeta_k^v,\\
\mu_k^q &= (W \otimes I_T)\zeta_k^q.
\end{aligned}$$

This follows from reorganizing terms in (5.30) and (5.31). Since μ_k^v, and μ_k^q are not known, and from [60], W^{-1} is sparse in the sense of the communication network, we can use a fast execution of the Jacobi overrelaxation method on the linear systems of equations $(W \otimes I_T)^{-1}\mu_k^v = \zeta_k^v$ and $(W \otimes I_T)^{-1}\mu_k^q = \zeta_k^q$ to compute an approximation of μ_k^v and μ_k^q. For further information on the Jacobi overrelaxation, see [61]. Let $\tilde{\mu}_k^v$, $\tilde{\mu}_k^q$ be ϵ-approximations of μ_k^v and μ_k^q. Then, we compute $\overline{v}^{k+1}$ and $\overline{q}_{\text{in}}^{k+1}$ as:

$$\begin{aligned}
\overline{v}^{k+1} =& v^k + \eta\mathbf{d}(W)^{-1} \otimes I_T\Big[\tilde{\mu}_k^v - \frac{1}{2}\delta(\mathbf{1} \otimes c)\Big],\\
\overline{q}_{\text{in}}^{k+1} =& q_{\text{in}}^k + \eta\mathbf{d}(W)^{-1} \otimes I_T\tilde{\mu}_k^q,
\end{aligned}$$

where we neglect the approximation error.

The following result gives us a valid distributed approximation of ξ_q and ξ_v.

Lemma 5.10 (Distributed computation)**.** *The vectors ξ_q and ξ_v are such that:*

$$(W^{-1} \otimes I_T)\xi_q = -\operatorname{Im}\left(e^{-j\theta}\begin{bmatrix} 0 & \mathbf{d}(\hat{u}_G^k) \end{bmatrix}(\mathsf{G} \otimes I_T)\begin{bmatrix} u_0^k \mathbf{1} \\ u_G^k \end{bmatrix}\right) - q_{in}^k + o\left(\frac{1}{U_0^2}\right), \tag{5.41}$$

$$(W^{-1} \otimes I_T)\xi_v = -\operatorname{Re}\left(e^{-j\theta}\begin{bmatrix} 0 & \mathbf{d}(\hat{u}_G^k) \end{bmatrix}(\mathsf{G} \otimes I_T)\begin{bmatrix} u_0^k \mathbf{1} \\ u_G^k \end{bmatrix}\right) - v^k + o\left(\frac{1}{U_0^2}\right), \tag{5.42}$$

where $\mathbf{1} \in \mathbb{R}^T$.

Proof. First, note that from Lemma 5.3, we have that G can be written as:

$$\mathsf{G} = \begin{bmatrix} (\mathsf{G})_{11} & (\mathsf{G})_{12} \\ (\mathsf{G})_{12}^\top & W^{-1} \end{bmatrix},$$

Second, we replace this expression above and the expression for the voltage in (5.3) into the right-hand side of (5.41), (5.42). While replacing the expression for voltage, we use the properties of the matrix X in Lemma 5.1. Finally, we rearrange terms to obtain the expressions in (5.4). □

Using this result in the equalities for ζ_k^v and ζ_k^q in (5.40), results in a completely distributed optimization algorithm. However, it presents a new problem: information on future voltages is required for computing $q_{\text{in}}(t)$, $v(t)$ is required. In Algorithm 4 we summarize the described procedure for one iteration of the control computation.

5.5 Voltage prediction

In order to compute $q_{\text{in}}(t)$, $v(t)$, for $t \in \tau$, the two algorithms we have introduced in this chapter require voltage information for all $t \in \tau$. However, this information is not available for two reasons: i) voltages $u(t)$ depend on the power injections at all the nodes of the microgrid at time $t \in \tau$, i.e., they depend on future values of the decision variables, and ii) although for time $t = 1$, it is theoretically possible to inject the power given by decision variables $v^k(1)$ and $q_{\text{in}}^k(1)$ into the system, these variables are only

Algorithm 4 Execution for each node $l \in \mathcal{G}$

Set $u_l^{k-1}(t)$, $v_l^{k-1}(t)$ (also $q_l^{k-1}(t)$ and $\overline{\lambda}_l^{k-1}(t)$, $\underline{\lambda}_l^{k-1}(t)$, if $l \in \mathcal{G}$) for all $t \in \tau$
for $t = \{1, \ldots, T\}$ **do**
$\overline{\lambda}_l^k(t) = \left[\overline{\lambda}_l^{k-1}(t) + \eta\gamma(w_l^{k-1}(t) - W_{\max})\right]_+$
$\underline{\lambda}_l^k(t) = \left[\underline{\lambda}_l^{k-1}(t) + \eta\gamma(W_{\min} - w_l^{k-1}(t))\right]_+$
end for
Initialize $\tilde{\mu}_l^v(t)$, $\tilde{\mu}_l^q(t) = 0$
Gather $u_h^{k-1}(t)$ for all $h \in \mathcal{N}_G(l)$
for $\ell \leq \ell_{\max}$ **do**
Gather $\tilde{\mu}_h^v(t)$, $\tilde{\mu}_h^q(t)$ for all $h \in \mathcal{N}_G(l)$, $t \in \tau$
for $t = \{1, \ldots, T\}$ **do**

$$\tilde{\mu}_l^v(t) = (1-\alpha)\tilde{\mu}_l^v(t) - \frac{\alpha}{\mathsf{G}_{ll}}\Bigg(\sum_{h\in\mathcal{N}_G(l)\setminus\{0,l\}} \mathsf{G}_{lh}\tilde{\mu}_h^v(t) - \sum_{h\in\mathcal{N}_G(l)} \mathsf{G}_{lh}\Big(\|u_l^{k-1}(t)\|_{\mathbb{C}}\|u_h^{k-1}(t)\|_{\mathbb{C}} \cos(\angle u_h^{k-1}(t) - \angle u_l^{k-1}(t) - \theta)\Big) + \frac{1}{U_0^2}\cos(\theta)(\overline{\lambda}_l^k(t) - \underline{\lambda}_l^k(t))\Bigg)$$

$$\tilde{\mu}_l^q(t) = (1-\alpha)\tilde{\mu}_l^q(t) - \frac{\alpha}{\mathsf{G}_{ll}}\Bigg(\sum_{h\in\mathcal{N}_G(l)\setminus\{0,l\}} \mathsf{G}_{lh}\tilde{\mu}_h^q(t) - \sum_{h\in\mathcal{N}_G(l)} \mathsf{G}_{lh}\Big(\|u_l^{k-1}(t)\|_{\mathbb{C}}\|u_h^{k-1}(t)\|_{\mathbb{C}} \sin(\angle u_h^{k-1}(t) - \angle u_l^{k-1}(t) - \theta)\Big) + \frac{1}{U_0^2}\sin(\theta)(\overline{\lambda}_l^k(t) - \underline{\lambda}_l^k(t))\Bigg)$$

end for
end for
for $t = \{1, \ldots, T\}$ **do**
$\overline{v}_l^k(t) = v_l^{k-1}(t) + \eta(\tilde{\mu}_l^v(t) - \frac{1}{2}\delta c(t))/W_{ll}$
$q_l^k(t) = q_l^{k-1}(t) + \eta\tilde{\mu}_l^q(t)/W_{ll}$
end for
$v_l^k = \mathrm{Proj}_{\mathcal{U}_l}\{\overline{v}_l^k\}$
$q_l^k = \mathrm{Proj}_{\mathcal{Q}_l}\{\overline{v}_l^k\}$

asymptotically feasible. Then, it may be either harmful or impossible to inject such power values into the actual microgrid. Thus, it is necessary to use a model to predict values of $u(t)$, for $t \in \tau$. Recall that a microgrid is modeled by the nonlinear memoryless system of equations formulated in (5.1). Finding the solution for voltages $u(t)$ given $P(t)$, $Q(t)$ and u_0 is not only a computationally expensive procedure, but it is also not distributed. We formulate two alternatives to address this problem.

5.5.1 A multilayer control approach

A first possibility is to define an additional layer for the control, which contains a model of the microgrid, and is such that it can exchange information with all nodes in $\mathcal{G}$. Thus, the control structure has two layers: the upper layer with a 'super agent' that knows the microgrid model, and the lower layer, with the node controllers. At each iteration, nodes in $\mathcal{G}$ will provide the super agent in the upper layer the value of q_{in}^k and v^k. The agent uses these values to compute u^k by solving the power flow equations in (5.1). Then, the agent provides the node controllers with u^k according to their sparsity in the sense of the network topology of Definition 5.1. Then, the dual decomposition algorithm is executed again. By means of this approach, part of the computations are carried out in a centralized way. Even though part of the computations are parallelized, i.e., the control computations, the centralized computations performed by the super agent destroy the scalability and robustness properties that justify the use of a distributed algorithm.

5.5.2 Distributed approximation

An alternative is the VOLTAGE PREDICTION ALGORITHM: a novel idea which is based on the voltage expression given in (5.3). It consists of executing several sub-iterations at each iteration k of the control computation, in order to approximate the voltage u^k for the computation of q_{in}^{k+1} and v^{k+1}. Figure 5.4 shows the interaction between the MICROGRID CONTROL ALGORITHM and the VOLTAGE PREDICTION ALGORITHM. Let us assume that all the loads in nodes that belong to $\mathcal{M}$ are constant, for all $t \in \tau$, i.e., $s_L(t) = s_L$, for all $t \in \tau$. Let us consider $u_0 \triangleq [U_0 e^{j\phi}, u_{0,G}^\top, u_{0,L}^\top]^\top \in \mathbb{R}^N$ and $s_0 \triangleq [S_0, s_{0,G}^\top, s_L^\top]^\top \in \mathbb{R}^N$, with

$s_{0,L} = p_L + jq_L$, such that the pair u_0, s_0 solves the power flow equations that model the microgrid. Assume that each node $l \in \mathcal{G}$ knows $u_{0,l}$ and $s_{0,l}$, where $u_{0,l}$ and $s_{0,l}$ denote the l^{th} components of u_0 and s_0 respectively, which means that the node l requires entirely local information.

Using the expression in (5.3), we have that:

$$\Delta u_t = e^{j\phi}\left(\frac{e^{j\theta}}{U_0} W \Delta \hat{s}_t\right),$$

where $\Delta u_t \triangleq u_G(t) - u_{G,0}$, and $\Delta \hat{s}_t \triangleq \hat{s}_G(t) - \hat{s}_{G,0}$, for all $t \in \{1, \ldots, T\}$. Then, we obtain:

$$\begin{aligned}
\operatorname{Re}(\Delta u_t) &= W\left(\frac{\cos(\theta+\phi)}{U_0}\operatorname{Re}(\Delta \hat{s}_t) - \frac{\sin(\theta+\phi)}{U_0}\operatorname{Im}(\Delta \hat{s}_t)\right),\\
\operatorname{Im}(\Delta u_t) &= W\left(\frac{\sin(\theta+\phi)}{U_0}\operatorname{Re}(\Delta \hat{s}_t) + \frac{\cos(\theta+\phi)}{U_0}\operatorname{Im}(\Delta \hat{s}_t)\right).
\end{aligned}$$

Notice that the l^{th} component of $\Delta \hat{s}_t$ is known to node $l \in \mathcal{G}$ for all $t \in \tau$. Since W is invertible, it is possible to solve Δu_t, using the Jacobi overrelaxation algorithm, as follows:

$$u_{e,t}^{\ell+1} = (1-h)u_{e,t}^{\ell} - h\mathbf{d}(W^e)^{-1}\left((W^e - \mathbf{d}(W^e))u_{e,t}^{\ell} - s_{e,t}^{\ell}\right), \tag{5.43}$$

where $W^e \triangleq I_2 \otimes W^{-1}$ and:

$$\begin{aligned}
u_{e,t} &= \begin{bmatrix} \operatorname{Re}(\Delta u_t) \\ \operatorname{Im}(\Delta u_t) \end{bmatrix},\\
s_{e,t} &= \begin{bmatrix} \frac{\cos(\theta+\phi)}{U_0}\operatorname{Re}(\Delta \hat{s}_t) - \frac{\sin(\theta+\phi)}{U_0}\operatorname{Im}(\Delta \hat{s}_t) \\ \frac{\sin(\theta+\phi)}{U_0}\operatorname{Re}(\Delta \hat{s}_t) + \frac{\cos(\theta+\phi)}{U_0}\operatorname{Im}(\Delta \hat{s}_t) \end{bmatrix}.
\end{aligned}$$

By construction of W, it also holds that the diagonal elements of W^e are strictly positive. Then the matrix $\mathbf{d}(W^e)^{-1}$ is well defined. Given that W^{-1} is symmetric and positive definite, W^e is also symmetric and positive definite. Then, it holds that if $h < 2/|\mathcal{G}|$, the Jacobi overrelaxation converges from any initial condition and also presents a linear rate of convergence. Since W^{-1} is distributed in the sense of the communication network in Definition 5.1, the computation of u can be made using only local information.

Algorithm 5 summarizes the prediction procedure for each node $l \in \mathcal{G}$. Notice that a

Algorithm 5 The VOLTAGE PREDICTION ALGORITHM. Execution for node $l \in \mathcal{G}$

Get $v_l^k(t)$, $q_{\text{in},l}^k(t)$ (if $l \in \mathcal{G}$), for all $t \in \tau$.
for $t = \{1, \ldots, T\}$ **do**
  $\delta_{t,Re}^0(l) = 0$
  $\delta_{t,Im}^0(l) = 0$
  $s_{l,t} = p_{\text{in},l}(t) - p_{\text{load},l}(t) - v_l^k(t) + j(q_{\text{in},l}^k(t) - q_{\text{load},l}(t))$
  $b_{t,Re} = \frac{\cos(\theta+\phi)}{U_0} \operatorname{Re}(s_{l,t} - s_{l,0}) - \frac{\sin(\theta+\phi)}{U_0} \operatorname{Im}(s_{l,t} - s_{l,0})$
  $b_{t,Im} = \frac{\sin(\theta+\phi)}{U_0} \operatorname{Re}(s_{l,t} - s_{l,0}) + \frac{\cos(\theta+\phi)}{U_0} \operatorname{Im}(s_{l,t} - s_{l,0})$

end for
for $\ell \in \{1, \ldots, \ell_{\max} - 1\}$ **do**
  for $t = \{1, \ldots, T\}$ **do**
    $\delta_{t,Re}^{\ell+1}(l) = (1-h)\delta_{t,Re}^{\ell}(l) - \frac{h}{\mathsf{G}_{S,ll}} \left(\sum_{r \in \mathcal{N}_S(l) \setminus \{0,l\}} \mathsf{G}_{S,lr} \delta_{t,Re}^{\ell}(r) - b_{t,Re} \right)$
    $\delta_{t,Im}^{\ell+1}(l) = (1-h)\delta_{t,Im}^{\ell}(l) - \frac{h}{\mathsf{G}_{S,ll}} \left(\sum_{r \in \mathcal{N}_S(l) \setminus \{0,l\}} \mathsf{G}_{S,lr} \delta_{t,Im}^{\ell}(r) - b_{t,Im} \right)$
  end for
end for
for $t = \{1, \ldots, T\}$ **do**
  $u_l^k(t) \approx \delta_{t,Re}^{\ell_{\max}}(l) + j\delta_{t,Im}^{\ell_{\max}}(l) + u_{l,0}$

end for

fast execution of this algorithm (until some error tolerance is reached) at the end of each iteration of the dual decomposition algorithm can give an approximation of $u_G(t)$, $t \in \tau$, for the next iteration.

Let us recall that the load and the active power generation at all nodes are system parameters that come from forecasting processes. In the particular case of load forecasting, persistence models are widely used. A persistence model assumes that the best forecast for a variable in the near future is the current value of that variable. It leads to an estimate of a future step for loads in nodes $l \in \mathcal{M}$ computed as follows:

$$s_L(t+1) = s_L(t).$$

A propagation of this estimation over the whole time horizon τ leads to a load estimation in which $s_L(t) = s_L$ for all $t \in \tau$, which is the setting under which our distributed prediction model works. Thus, even though the assumption that the load of nodes in

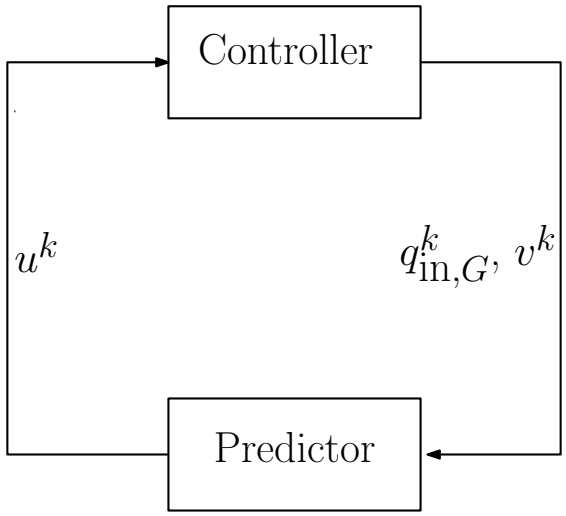

Figure 5.4: Control computation using voltage prediction.

$\mathcal{M}$ does not vary with time seems to be restrictive, it is not introducing any additional assumptions than those that are already posed in real applications.

5.6 Simulation results and discussion

In order to show the algorithm performance, we use single-phase versions of the IEEE13, IEEE34, IEEE37, and IEEE123 feeders. All the data on transmission line impedances, loads, power generation and storage capacities, as well as further details of the simulation, and simulation files can be found in the project website: `http://fausto.dynamic.ucsd.edu/andres/project_reactive_fast`. For each testbed we implement the MICROGRID OPTIMIZATION ALGORITHM and also the dual decomposition algorithm introduced in [60]. To obtain the voltage prediction required for the algorithm execution, we use in both cases the distributed prediction algorithm in [60], which works at a fast time scale to compute an approximation of the voltage given the values of the decision variables. The prediction algorithm has been tuned to use the parameter that maximizes the speed of convergence. Both the dual decomposition algorithm and the MICROGRID OPTIMIZATION ALGORITHM have been tuned to a step size that maximizes the speed of convergence. In Table 5.1 we show the parameters for each algorithm involved in the simulation: MO stands for the MICROGRID OPTIMIZATION ALGORITHM,

Table 5.1: Parameters for each testbed

	MO			DD	Prediction	
Testbed	η	α	$\ell_{\max}$	γ	h	FE
IEEE13	0.95	0.271	30	0.0737	0.995	100
IEEE34	0.95	0.211	30	0.0052	0.995	2500
IEEE37	0.95	0.211	30	0.0728	0.995	100
IEEE123	0.95	0.095	30	0.0356	0.995	1000

Table 5.2: Practical Convergence

	MO		DD	
Testbed	OL	Total	OL	Total
IEEE13	50	1.5×10^5	1.8×10^5	1.8×10^7
IEEE34	500	37.5×10^6	3.6×10^6	9×10^9
IEEE37	50	1.5×10^5	6×10^4	6×10^6
IEEE123	150	4.5×10^6	2×10^5	2×10^8

DD stands for the dual decomposition algorithm in [60] and prediction is the voltage prediction algorithm. For the MICROGRID OPTIMIZATION ALGORITHM, η is the step size as it has been introduced in this document, α is the step size of the fast execution of the Jacobi overrelaxation intended to estimate $\tilde{\mu}_k^v$ and $\tilde{\mu}_k^q$, and $\ell_{\max}$ is the number of iterations that we run for such estimation. For DD, γ is a dual ascent step size as described in [60]. For the prediction algorithm, h corresponds to the Jacobi overrelaxation step size and FE stands for the number of iterations that are executed for the voltage prediction. Notice that the parameters η, α, γ and h have been chosen in a way that each of the algorithms exhibits near the fastest possible convergence for the case study. The parameters FE and $\ell_{\max}$ have been chosen so that each approximation that is being made via the Jacobi overrelaxation, reaches its actual value within 3 significant figures.

It is interesting to notice that, in the case of the feeder IEEE34, the convergence of the MICROGRID OPTIMIZATION ALGORITHM, the convergence of the dual decomposition algorithm, and also that of the voltage prediction algorithm, are significantly slower that those of the other feeders we simulate. This is due to the largest electrical distance between two actuated nodes in the network.

In Table 5.2, we compare the speed of convergence of the MICROGRID OPTIMIZATION ALGORITHM with that of the dual decomposition algorithm. For the MICROGRID OPTIMIZATION ALGORITHM, the OL (outer loop) column indicates how many times the procedure de-

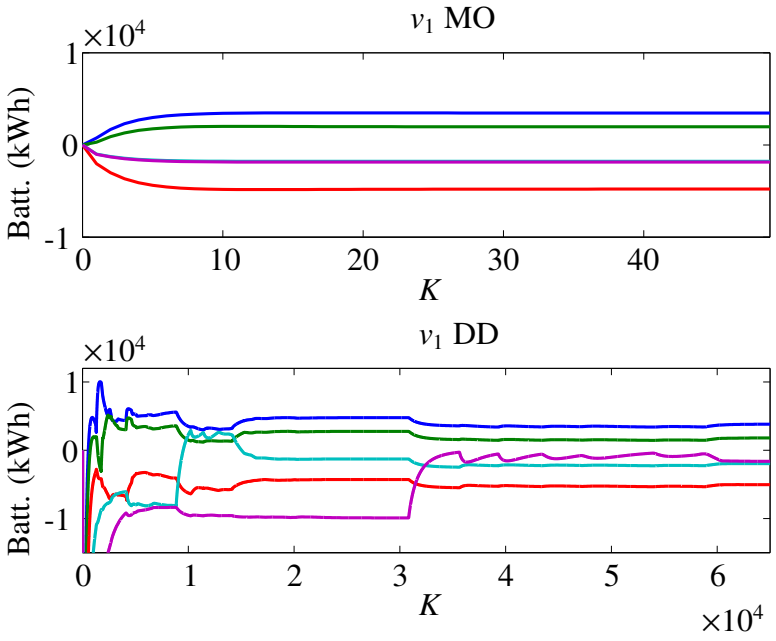

Figure 5.5: Top: Evolution of $v_1(t)$ for all $t \in \tau$, using the MICROGRID OPTIMIZATION ALGORITHM. Bottom: Evolution of $v_1(t)$ for all $t \in \tau$ using the dual decomposition algorithm.

scribed in Algorithm 4 is repeated to reach practical convergence. This number has been chosen in a conservative way, i.e., practical convergence is reached in a lesser number of iterations. The column called Total shows the amount of communication rounds executed to reach practical convergence, this is, the number of outer loop executions, multiplied by the number of iterations used for the voltage prediction, which corresponds to the column FE in Table 5.1, and also multiplied by the amount of iterations carried out for the approximation of μ_k^v, μ_k^q, which can be found in the column $\ell_{\max}$ in Table 5.1.

For the dual decomposition algorithm, we show in the OL column the amount of iterations of Algorithm 1 in [60] necessary to reach practical convergence. In the column labeled Total, we multiply this number by the number of iterations taken by the voltage prediction algorithm. We can see that the amount of iterations needed for practical convergence in the MICROGRID OPTIMIZATION ALGORITHM is at least two orders of magnitude less than the dual decomposition algorithm in each of the simulation cases. It is important to highlight that if it were not for the need to use the Jacobi overrelaxation

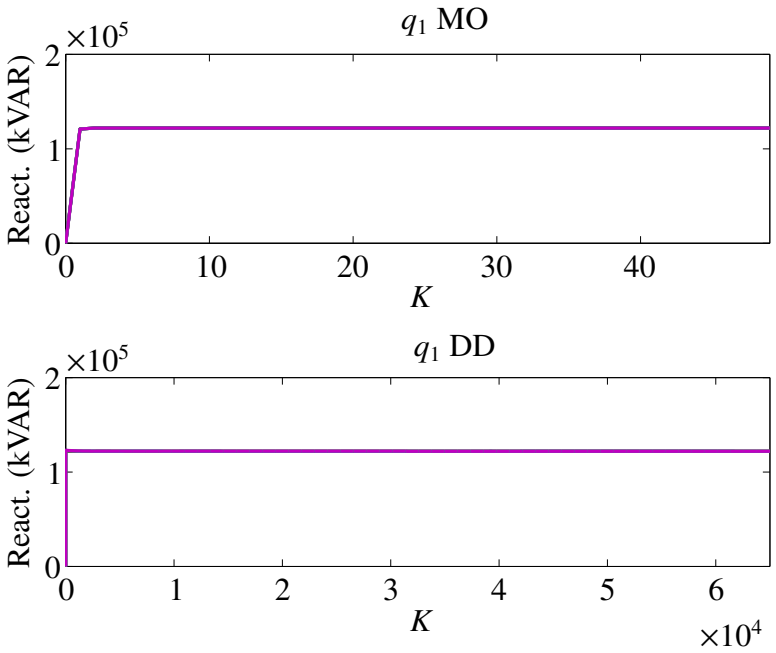

Figure 5.6: Top: Evolution of $q_1(t)$ for all $t \in \tau$, using the MICROGRID OPTIMIZATION ALGORITHM. Bottom: Evolution of $q_1(t)$ for all $t \in \tau$ using the dual decomposition algorithm.

in the computation of the primal step of the algorithm, the speed of convergence would be much greater.

Figures 5.5 through 5.9 show the trajectory of all the state variables related to node 1 in the IEEE37 feeder for both algorithms.

5.7 Summary

This chapter presents two distributed algorithm for the computation of predictive control sequences of reactive power and storage in microgrids. The algorithms use forecasted parameters such as the electricity cost in time and the solar-power generation, in order to compute the reactive power that must be injected by generators and the charging/charging rate that storage devices must follow in order to minimize electricity cost and also transmission losses. Both designs are based on a previous convexification approach which relaxes the power flow equations onto a linear relation between power and voltage in a microgrid, under some assumptions on impedances in the transmission lines

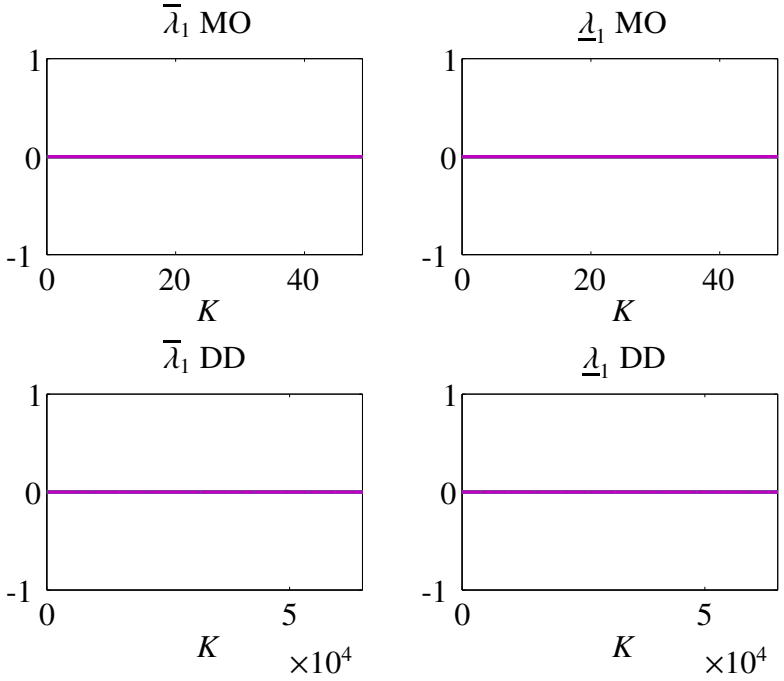

Figure 5.7: Top plots: $\underline{\lambda}_1(t)$, $\overline{\lambda}_1(t)$ for all $t \in \tau$, for the MICROGRID OPTIMIZATION ALGORITHM. Bottom plots: $\underline{\lambda}_1(t)$, $\overline{\lambda}_1(t)$ for all $t \in \tau$, for the dual decomposition algorithm.

and the magnitude of the input voltage. New constraints on voltage regulation, and operational constraints on the storage systems are considered. The first algorithm, a dual-decomposition-based dynamics is thoroughly analyzed, concluding that the dynamics globally converge to the unique optimizer of the problem. Given the slow rate of convergence of this algorithm, we propose the MICROGRID OPTIMIZATION ALGORITHM. The algorithm manages the global constraints using a dual decomposition-like approach while all the local constraints are taken care of using primal projections. This allows us to obtain a much faster convergence than that of the first algorithm. We show convergence of the MICROGRID OPTIMIZATION ALGORITHM towards the optimizer of the MICROGRID CONTROL MICROGRID CONTROL PROBLEM, by analyzing the CONSTRAINT SEPARATION DYNAMICS. This dynamics encompasses the structure of the MICROGRID OPTIMIZATION ALGORITHM, and all its solutions converge to an optimizer of the PARTIALLY SEPARABLE QUADRATIC PROBLEM, which has the same structure as the MICROGRID CONTROL MICROGRID CONTROL PROBLEM. Further, we present a distributed implementation of the algorithm, which is based on an approximation carried out using the Jacobi overrelaxation in a fast-time scale.

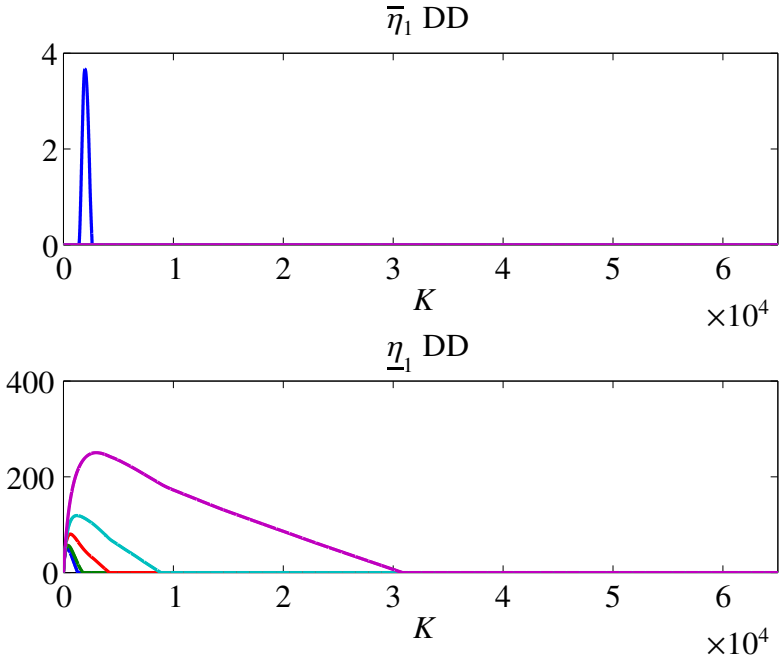

Figure 5.8: Evolution of Lagrange multipliers for the dual decomposition algorithm, top: $\overline{\eta}_1(t)$ for all $t \in \tau$, bottom: $\underline{\eta}_1(t)$ for all $t \in \tau$.

Since the two algorithms require predicted voltage valued, we discuss two prediction ways, including a novel distributed prediction model.

In order to show that the MICROGRID OPTIMIZATION ALGORITHM is faster than the first algorithm, we implement them in four different testbeds that are microgrids with the structure to the IEEE13, IEEE34, IEEE37 and IEEE123 test feeders, in a single-phase setting. All trials show that the MICROGRID OPTIMIZATION ALGORITHM requires $1/100$ of iterations needed by the first algorithm to reach practical convergence.

Acknowledgments

This paper contains material that has been published in the following works:

- A. Cortés and S. Martínez, "A Projection-based Dual Algorithm for Fast Computation of Control in Microgrids," in the proceedings of the SIAM Conference on Control and its Applications (2015), 109-114.

- A. Cortés and S. Martínez, "Distributed Control of Reactive Power and Storage

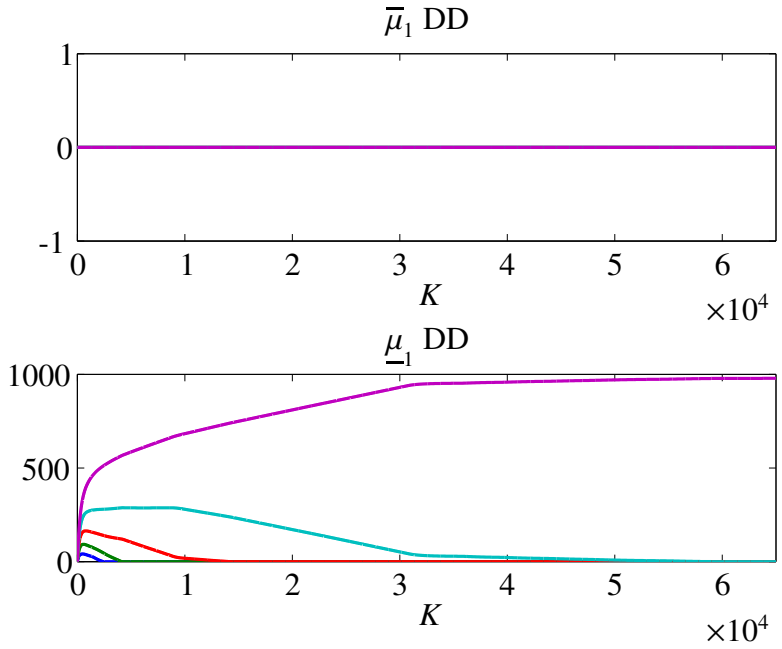

Figure 5.9: Evolution of Lagrange multipliers for the dual decomposition algorithm, top: $\underline{\mu}_1(t)$ for all $t \in \tau$, bottom: $\overline{\mu}_1(t)$ for all $t \in \tau$.

in Microgrids," in the proceedings of the 21st International Symposium on Mathematical Theory of Networks and Systems (2014)

- A. Cortés and S. Martínez, "A Projection-Based Dual Algorithm for Fast Computation of Control in Microgrids," submitted to the SIAM Journal on Control and Optimization (2014).

- A. Cortés and S. Martínez, "On Distributed Reactive Power and Storage Control on Microgrids," submitted to the International Journal of Robust and Nonlinear Control (2014), revised 2015.

Chapter 6

Self-triggered best response dynamics for continuous games

The last years have witnessed an intense research activity in the development of novel distributed algorithms for multi-agent systems with performance guarantees. A particular effort has been devoted to the study of game-theoretic approaches that can model and regulate selfish agent interactions. By means of these, the multi-agent coordination objective is formulated in terms of Nash Equilibria (NE), which correspond to the natural emergent behavior arising from the interaction of selfish players. Due to their modularity, game dynamics can easily be implemented by agents relying on local information, leading to a robust performance. Even though the resulting emerging behavior may not be optimal, it is generally expected that the behavior is as close as possible to that of the benchmark given by a centralized design. However, finding algorithms to reach a NE is not always an easy task, mainly due to the fact that in general, the NE is very difficult to compute, and even some games do not have any.

In particular, the best-response dynamics can be thought of as a natural dynamics describing an interaction in which each player is able to compute its own best action against other players' action profile. Then, the player's action evolves continuously towards its best-response set. Convergence of the best-response dynamics to the set of NE has been studied for games under well defined conditions. In [62, 63], convergence is proven for finite zero-sum games with bilinear payoff functions and potential games with continuously differentiable potential functions. In [64], the authors consider

best-response dynamics for two-player zero-sum games, relaxing the previous differentiability assumptions to concave and convex payoff functions. Convergence to the saddle point set is proven, since this set corresponds to the NE set of the game. The current effort in [65] extends the above result to a two-player zero-sum continuous game with quasi-concave and quasi-convex continuous payoff functions. Although these works consider non-differentiable payoff functions, their approach is oriented to two-players zero-sum games.

The idea of restricting communication efforts to time instants at which it is absolutely necessary to have current information leads to the self-triggered and event-triggered concepts; e.g. see [66, 67] and references therein. This idea has been recently extended to the context of multi-agent systems and distributed optimization with applications to cooperative control [68, 69]. The present work contributes further to this area by studying a complementary game-theoretic setting with applications to coverage control and deployment similarly to [70, 71]. A main difference with gradient-based methods is given by the need of estimating the evolution of best-response sets as actions of other agents change.

More precisely, we start by analyzing the convergence properties of the so-called continuous-time best-response dynamics for a continuous-action-space game by means of the invariance theory for set-valued dynamical systems. We show that all the solutions of the best-response dynamics converge to the NE set of the potential game for component-wise pseudoconcave, component-wise quasiconcave potential functions. Next, we introduce a novel self-triggered best-response dynamics relying on Lipschitz payoff functions. Then, we prove how this strategy still ensures convergence to the set of NE while decreasing communication efforts. An application example can be found in the extended manuscript [72]. The paper [73] presents an alternative self-triggered best-response dynamics for 1D action spaces where the self-triggering condition comes from the assumption that the best-response sets are Lipschitz set-valued maps.

Notations. In what follows, $\text{sign} : \mathbb{R} \to \mathbb{R}$ is defined as $\text{sign}(x) = 1$ if $x > 0$, $\text{sign}(x) = -1$ if $x < 0$, and $\text{sign}(x) = 0$ if $x = 0$. Let S be a subset of $\mathbb{R}^n$, then $\text{co}(S)$ denotes the convex hull of S, $\overline{S}$ denotes the closure of S, and $\rho(S)$ denotes the diameter of S, $\rho(S) = \sup_{x,y \in S} \|x - y\|$. If f is a map, $\text{dom}(f)$ represents the domain of f. The

open ball with radius r centered at x is denoted as $B_r(x)$. Given $f : \mathbb{R}^n \to \mathbb{R}$, we define a level set of f as $f^{-1}(r) = \{x \in \text{dom}(f) \mid f(x) = r\}$. For $A, B \subset \mathbb{R}^n$, we denote $A \setminus B = \{x \in A \mid x \notin B\}$. Let A be a subset of $\mathbb{R}^n$. Then, int(A) represents the interior of A, and bnd(A) represents the boundary of A.

6.1 Game theoretical notions

In this section, we first introduce some basic definitions from Game Theory [74] and an adaptation from [71] to deal with constrained motion coordination problems.

Definition 6.1. *A continuous-action-space game is a* 3*-tuple* $\Gamma = (I, X, u)$*, such that (i)* $I = \{1, \ldots, N\}$ *is the set of N players, (ii)* $X = \prod_{i=1}^{N} X_i \subset \mathbb{R}^d$ *is the action space of the game, with* $X_i \subset \mathbb{R}^{n_i}$*,* $i \in I$*,* $d = \sum_i n_i$*, a compact and convex set representing the action space of the* i^{th} *player, and (iii)* $u : X \to \mathbb{R}^N$ *is a function whose component* $u_i : X \to \mathbb{R}$ *defines the payoff of the* i^{th} *player,* $i \in I$*.*

Let $x_i \in X_i \subset \mathbb{R}^{n_i}$ be the action for the i^{th} player and $x \in X$ be the action profile for all players, such that $x = (x_1, \ldots, x_N)^\top$. In the sequel, we will use the notation $x = (x_i, x_{-i})$, where $x_{-i} \in X_{-i} = \prod_{j \in I, j \neq i} X_j$, for all $i \in I$, are the actions of all players except that of the i^{th} player.

A repeated, continuous-time, game associated with Γ, $\mathcal{R}(\Gamma)$, is a game in which, at each time $t \in \mathbb{R}_{\geq 0}$, each agent $i \in I$ modifies $x_i(t) \in X_i$ simultaneously while receiving $u_i(x(t))$. This is in contrast to repeated, discrete-time games, which follow a discrete-time schedule.

In the context of (vehicle) motion coordination, agents' actions can be identified with system states, and thus it makes sense that these change in continuous time according to some vehicle dynamics. In particular, the way in which player i modifies $x_i(t)$ can be constrained by a (state-dependent) set W. Let $W_i(x_i, x_{-i}) \subset X_i$ be a constraint subset associated with $x \in X$, $W(x) = \Pi_{i \in I} W_i(x) \subset X$, and $W = \cup\{(x, W(x)) \mid x \in X\} \subseteq X \times X$. We will refer to W as a *fiber bundle* over X. The introduction of W leads to the notion of constrained repeated game associated with Γ and W, $\mathcal{R}_W(\Gamma)$, and the following concept.

Definition 6.2. *Let* $\Gamma = (I, X, u)$ *be a continuous-space game and W a fiber bundle over*

X. A constrained Nash Equilibrium (NE) for Γ with respect to W is an action profile $(x_i^\star, x_{-i}^\star) \in X$ such that $u_i(x_i, x_{-i}^\star) \leq u_i(x_i^\star, x_{-i}^\star)$, for all $x_i \in W_i(x_i^\star, x_{-i}^\star)$ and all $i \in I$.

In the extended version [72], we use W to represent collision-avoidance type of constraints, or restricted reachable sets, thus it will be additionally assumed that $x \in W(x)$. Finally, we recall the following notion.

Definition 6.3 ([75])**.** *Consider a game $\Gamma = (I, X, u)$. Let us assume that there exists a function $\Phi : X \to \mathbb{R}$ such that $\operatorname{sign}\left(u_i(x_i, x_{-i}) - u_i(x_i', x_{-i})\right) = \operatorname{sign}\left(\Phi(x_i, x_{-i}) - \Phi(x_i', x_{-i})\right)$, for $x_i, x_i' \in X_i$, $x_{-i} \in X_{-i}$, for all $i \in I$. Then, the game is called an ordinal potential game.*

6.2 Continuous-time best-response dynamics

Here, we introduce some basic facts about continuous-time best-response dynamics [76] and show their convergence to the set of equilibria under some general conditions. Definitions of regularity, upper-semicontinuity, generalized gradient and set-valued Lie derivative can be found in [72]. The generalized LaSalle's invariance principle for differential inclusions can be found in [77].

Let $\Gamma = (I, X, u)$ be a continuous-space game, let W be a continuous fiber bundle over X, such that $W_i(x)$ convex and compact for all $x \in X$, $i \in I$, and consider the game $\mathcal{R}_W(\Gamma)$.

Definition 6.4. *The best-response dynamics for $\mathcal{R}_W(\Gamma)$ is defined by the differential inclusion $F : X \rightrightarrows \mathbb{R}^d$, $F_i(x) = \mathrm{BR}_i(x_{-i}) - x_i = \operatorname{argmax}_{y \in W_i(x)} u_i(y, x_{-i}) - x_i$, for all $i \in I$. That is,*

$$\dot{x}_i \in F_i(x) := \mathrm{BR}_i(x_{-i}) - x_i, \quad i \in I. \tag{6.1}$$

We denote $F(x) = \mathrm{BR}(x) - x$ for conciseness. Existence of solutions for differential inclusions is guaranteed for F, nonempty, upper semicontinuous and taking compact and convex values. Let us assume that payoff functions u_i are continuous maps on X. By compactness of X, u_i reaches its maximum value on X, and the set of maximizers is compact. Then, F_i is nonempty and takes compact values. Further, let us assume

that u_i is quasiconcave on W. Then, the set of maximizers of u_i and F_i are convex for each $x \in W$[1]. By continuity of u_i on X, and continuity of W we can apply directly the maximum theorem [78] to conclude that F_i is upper semicontinuous for each $i \in I$. Alternatively, in potential games, one can exchange the continuity assumption on the u_i by continuity on the Φ. Since F_i is nonempty, compact, convex and upper semicontinuous at every $x \in X$, and each $i \in I$, there exists a solution to (6.1) for every initial condition. These solutions are absolutely continuous functions, $\varphi : [0, +\infty) \to X$, such that $\dot{\varphi}_i(t) \in \mathrm{BR}_i(\varphi_{-i}(t)) - \varphi_i(t)$, for almost every $t \in [0, +\infty)$, and for all $i \in I$; see [77]. The equilibria set of system (6.1) is

$$X^\star = \{x \in X \mid x_i \in \mathrm{BR}_i(x_{-i}),\ \forall i \in \{1, \dots, N\}\}. \tag{6.2}$$

This set corresponds exactly to the set of constrained Nash equilibria for Γ with respect to W. The existence of a Nash equilibrium is guaranteed by the existence of a continuous potential function that reaches its supremum over a compact action space.

The above theorem will be used to analyze the best-response dynamics associated with a potential game. The potential function should satisfy the following property.

Definition 6.5. *Let $Y \subseteq \mathbb{R}^d$ be a convex set. A function $\Phi : Y \to \mathbb{R}$ is said to be component-wise pseudoconcave if for every $i \in I$, and every $w = (w_i, w_{-i}), y = (y_i, y_{-i}) \in Y$, and $s \in (0, 1)$, with $y_{-i} = w_{-i}$, it holds that if $\Phi(w) > \Phi(y)$, then $\Phi(sw_i + (1-s)y_i, w_{-i}) \geq \Phi(y_i, w_{-i}) + (1 - s)sb(w_i, y_i)$ where $b(w_i, y_i)$ is a positive function. The function Φ is component-wise quasiconcave if $\Phi(sw_i + (1 - s)y_i, w_{-i}) \geq \min\{\Phi(y_i, w_{-i}), \Phi(w_i, w_{-i})\}$, for every $i \in I$, and every $w = (w_i, w_{-i}), y = (y_i, y_{-i}) \in Y$. If Φ is pseudoconcave (resp. quasiconcave), $-\Phi$ is pseudoconvex (resp. quasiconvex).*

Theorem 6.1. *Let $\Gamma = (I, X, u)$ be an ordinal potential game with potential function Φ. Let W be a continuous fiber bundle over X such that $W_i(x)$ is compact and convex for all $x \in X$, $i \in I$. Assume that Φ is component-wise quasiconcave, component-wise pseudoconcave with b_i continuous over each $W(x)$, $x \in X$, Lipschitz, and regular over X. Let $F : X \rightrightarrows \mathbb{R}^d$ be the best-response dynamics for $\mathcal{R}_W(\Gamma)$. Then, all solutions of the system $\dot{x} \in F(x)$ converge to the set $X^\star$ of constrained Nash equilibria defined in* (6.2).

[1]The same result holds for a potential game with a component-wise quasiconcave potential function (Definition 6.5).

Proof. Consider $\Psi = -\Phi$. Since Φ is component-wise pseudoconcave, then Ψ is component-wise pseudoconvex. We will see that Ψ is a Lyapunov function for our set-valued map; that is, it holds that $\max \mathcal{L}_F \Psi(x) \leq 0$, for all $x \in X$.

Let x be a point in X. Any $v \in F(x)$ has the form $v = x^\star - x$, with $x^\star \in \mathrm{BR}(x)$. Define $\Omega_\Psi \subset X$ as the zero-measure set for which Ψ is non-differentiable. Consider a $\zeta \in \partial\Psi(x)$ of the form $\zeta = \lim_k \nabla\Psi(y^k)$, with $y^k \to x$, $y^k \notin \Omega_\Psi$. If $x^\star = x$, then it trivially holds that $v^T\zeta = 0$. Suppose that $x^\star \neq x$. Since BR is nonempty and upper-semicontinuous for all $x \in X$, it holds that there exists a sequence $x^{k,\star} \to x^\star$ such that $x^{k,\star} \in \mathrm{BR}(y^k)$, for all k. Thus, we have $v^T\zeta = (x^\star - x)^T \lim_k \nabla\Psi(y^k) = \lim_k (x^{k,\star} - y^k)^T \nabla\Psi(y^k)$.

Let us define $\nabla_i \Psi \in \mathbb{R}^{n_i}$ as the partial derivative of Ψ with respect to the action of the i^{th} player. Since Ψ is differentiable at y^k, the term $(x^{k,\star} - y^k)^T \nabla\Psi(y^k)$ is the directional derivative of Ψ at y^k along the direction $x^{k,\star} - y^k$. In particular,

$$
\begin{aligned}
v^T\zeta &= \lim_{k\to\infty} (x^{k,\star} - y^k)^T \nabla\Psi(y^k) \\
&= \lim_{k\to\infty} \sum_{i\in I} (x_i^{k,\star} - y_i^k) \nabla_i \Psi(y^k) \\
&= \lim_{k\to\infty} \sum_{i\in I} (x_i^{k,\star} - y_i^k, 0_{-i})^T \nabla\Psi(y^k) \\
&= \lim_{k\to\infty} \sum_{i\in I} \lim_{h\to 0} \frac{\Psi(y_i^k + h(x_i^{k,\star} - y_i^k), y_{-i}^k) - \Psi(y^k)}{h},
\end{aligned}
$$

where in the last equality we have used the limit definition of directional derivative.

Notice that since $x^{k,\star} \in \mathrm{BR}(y^k)$, then it holds that $\Psi(y^k) \geq \Psi(x_i^{k,\star}, y_{-i}^k)$ for any $i \in I$. Moreover, since $x \neq x^\star$, we have that there is a $k_1 < \infty$ for which $y^k \neq x^{k,\star}$ for all $k > k_1$. Next, assume that $x \notin \mathrm{BR}(x)$, then there is an $i \in I$ such that $x_i \notin \mathrm{BR}_i(x)$. By continuity of Ψ, the set $\mathrm{BR}(x)$ is closed, therefore for each $x_i \in W_i(x) \setminus \mathrm{BR}_i(x)$ there exists ε such that $B_\varepsilon(x) \cap W_i(x) \subset W_i(x) \setminus \mathrm{BR}_i(x)$. Therefore, there is $k_2 < \infty$ such that $y_i^k \notin \mathrm{BR}_i(x^{k,\star})$ for all $k > k_2$. Thus, when we study the behavior as $k \to \infty$, we will consider only sequences y^k such that $y^k \neq x^{k,\star}$ and $y^k \notin \mathrm{BR}(y^k)$.

Using the fact that $y_i^k \notin \mathrm{BR}_i(y^k)$ and by component-wise pseudoconvexity of Ψ, it holds that since $\Psi(x_i^{k,\star}, y_{-i}^k) < \Psi(y^k)$, then $\Psi(y_i^k + h(x_i^{k,\star} - y_i^k), y_{-i}^k) \leq \Psi(y_i, y_{-i}^k) + (1 - h)hb_i(x_i^{k,\star}, y_i^k)$, for any $h \in (0, 1)$, and each $i \in I$. From here, $\Psi(y_i^k + h(x_i^{k,\star} - y_{-i}^k), y_{-i}^k) -$

$\Psi(y^k) \leq (1-h)hb_i(x_i^{k,\star}, y_i^k)$, which implies that $\lim_{h\to 0} \frac{\Psi(y_i^k + h(x_i^{k,\star} - y_i^k), y_{-i}^k) - \Psi(y^k)}{h} \leq b_i(x_i^{k,\star}, y_i^k)$. Now, for each $j \in I$ such that $y_j^k \in \mathrm{BR}_j(y^k)$, we have that $\Psi(y^k) = \Psi(x_j^{k,\star}, y_{-j}^k)$. It means that y_j^k and $x_j^\star$ are minimizers of $\Psi(\cdot, y_{-j}^k)$. By component-wise quasiconvexity, the set of minimizers of $\Psi(\cdot, y_{-j}^k)$ is convex, then $\Psi(y_j^k + h(x_j^{k,\star} - y_j^k), y_{-j}^k) = \Psi(y^k) = \Psi(x^{k,\star})$, therefore we can conclude that $\lim_{h\to 0} \frac{\Psi(y_i^k + h(x_i^{k,\star} - y_i^k), y_{-i}^k) - \Psi(y^k)}{h} = 0$. Then, it follows by using the continuity of b_i that

$$
\begin{aligned}
v^T\zeta &= \lim_{k\to\infty} (x^{k,\star} - y^k)\nabla\Psi(y_k) \\
&\leq \lim_{k\to\infty} \sum_{\substack{i\in I \\ y_i^k \notin \mathrm{BR}_i(y^k)}} b_i(x_i^{k,\star}, y_i^k) = \sum_{\substack{i\in I \\ x_i \notin \mathrm{BR}_i(x)}} b_i(x_i^\star, x_i). \qquad (6.3)
\end{aligned}
$$

Now, consider the case when $x \in \mathrm{BR}(x)$. In this case, if $x \in \mathrm{int}(\mathrm{BR}(x))$, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subset \mathrm{BR}(x)$. Note that for $h \to 0$, $(x_i + h(x_i^\star - x_i), x_{-i}) \in B_\varepsilon(x) \subset \mathrm{BR}(x)$. Then, $\Psi(x_i + h(x_i^\star - x_i), x_{-i}) = \Psi(x_i, x_{-i})$, and $\lim_{h\to 0} \frac{\Psi(x_i + h(x_i^\star - x_i), x_{-i}) - \Psi(x)}{h} = 0$, for each $i \in I$. It implies that $v^T\zeta = 0$. If $x \in \mathrm{bnd}(\mathrm{BR}(x))$, there are: i) sequences $y^k \to x$ such that for every k, $y^k \notin \mathrm{BR}(y^k)$, ii) sequences such that $y^k \in \mathrm{BR}(y^k)$ for all k, and iii) sequences such that there is a subsequence $\{y^{k,l}\}_l \subset \mathrm{BR}(y^k)$ and a subsequence $\{y^{k,l}\}_l \subset W(x) \setminus \mathrm{BR}(y^k)$. In case i) we follow the same analysis as that for $x \notin \mathrm{BR}(x)$, then $v^T\zeta < 0$, in case ii) the analysis is analogous to that for $x \in \mathrm{int}(\mathrm{BR}(x))$, to show that $v^T\zeta = 0$. In the third case, if x is a point of non-differentiability, the gradient of Ψ at y^k does not converge, then we do not need to consider these sequences. If Ψ is differentiable at x, then, as x_i is a minimizer of $\Psi(\cdot, x_{-i})$, for all $i \in I$, we have that $v^T\zeta = 0$. Hence, $v^T\zeta \leq 0$ for all $x \in \mathrm{bnd}(\mathrm{BR}(x))$. Then, we can conclude that for all $x \in X$, it holds that $v^T\zeta \leq 0$ for all sequences $y^k \to x$, such that $y^k \notin \Omega_\Psi$ and $\lim_k \nabla\Psi(y^k) = \zeta$.

Now consider any $\zeta \in \partial\Psi(x)$. By the definition of *generalized gradient* (see [77]), there exist $\alpha_1, \ldots, \alpha_l$, with $0 \leq \alpha_s \leq 1$, and $\alpha_1 + \cdots + \alpha_l = 1$, and sequences $\{y^{k_1}\}, \ldots, \{y^{k_l}\}$ converging to x such that $\zeta = \alpha_1 \lim_{k_1} \nabla\Psi(y^{k_1}) + \cdots + \alpha_l \lim_{k_l} \nabla\Psi(y^{k_l})$. Then it follows that $v^T\zeta = \alpha_1 v^T\zeta_1 + \cdots + \alpha_l v^T\zeta_l$. Using the previous analysis for each ζ_s, it follows that $v^T\zeta \leq 0$. From here we conclude that $\max \mathcal{L}_F\Psi(x) \leq 0$ for all $x \in X$. From the generalized LaSalle invariance principle [77], we have that all solutions will converge to the largest invariant set contained in $X \cap \overline{\{x \in \mathbb{R}^d \mid 0 \in \mathcal{L}_F\Psi\}}$. In the following, we prove

that the largest invariant set is contained in $X^\star$.

Suppose that $x \notin \mathrm{BR}(x)$, and x belongs to the invariant set. Take a $x^\star \in \mathrm{BR}(x)$, define $v = x^\star - x$, and take a $\zeta \in \partial\Psi(x)$ such that $\zeta = \lim_k \nabla\Psi(y^k)$, with $y^k \to x$, when $k \to +\infty$. From (6.3), we have that $v^T\zeta \leq \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(x_i^{k,\star}, y_i^k) < 0$, where, the second inequality follows from the fact that there is a $j \in I$ such that $x_j \notin \mathrm{BR}_j(x_{-j})$. Taking the maximum over $\mathrm{BR}(x)$, we have that

$$\begin{aligned}\max_{x^\star\in\mathrm{BR}(x)} (x^\star - x)^T\zeta &= \max_v v^T\zeta \\ &\leq \max_{x^\star\in\mathrm{BR}(x)} \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(x_i^\star, x_i) \\ &= \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(\bar{x}_i^\star, x_i).\end{aligned}$$

That is, the continuous function $\sum_{i\in I} b_i(\bar{x}_i^\star, x_i)$ achieves its maximum over the compact $\mathrm{BR}(x)$ at some $\bar{x}^\star \in \mathrm{BR}(x)$. Note that the inequality holds for all ζ of the form considered. Since $x \notin \mathrm{BR}(x)$, $\bar{x}^\star \neq x$, then $\Psi(\bar{x}_i^\star, x_{-i}) < \Psi(x)$ for some $i \in I$, hence we have that $\sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(\bar{x}_i^\star, x_i) < 0$.

Now consider any ζ that is a convex combination of $\zeta_s = \lim_s \nabla\Psi(y^{k_s})$. From the above considerations, we have that

$$\begin{aligned}v^T\zeta &= \alpha_1 v^T\zeta_1 + \cdots + \alpha_l v^T\zeta_l \\ &\leq \alpha_1 \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(\bar{x}_i^\star, x_i) + \cdots + \alpha_l \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(\bar{x}_i^\star, x_i) \\ &= \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(\bar{x}_i^\star, x_i) < 0,\end{aligned}$$

for all v and ζ. From here we conclude that $0 \notin \mathcal{L}_F\Psi(x)$, if $x \notin \mathrm{BR}(x)$. Thus, x does not belong to $X \cap \overline{\{x \in \mathbb{R}^d \mid 0 \in \mathcal{L}_F\Psi\}}$, which is a contradiction. □

6.3 Self-triggered Communications in Best-Response Dynamics

In this section, we present a sufficient self-triggered communication law as in [69] to lower the frequency at which neighbors' information needs to be updated while still guaranteeing convergence to the set of NE. It is worth highlighting that the amount of neighbors from which information needs to be updated for each player, depends uniquely on the sparsity of the game. That is, our law does not deal with generating a distributed execution of a non-distributed game, but it rather has to do with reducing the time-between-updates. All proofs for the results introduced in this section can be found in [72].

Let $\{t_k^i\}_{k=0}^{\infty} \subseteq \mathbb{R}_{>0}$, such that $t_k^i < t_{k+1}^i$, be the time sequence at which player i updates information about other players, for each $i \in I$. Assume that the i^{th} player has obtained up-to-date information of agent $j \in I \setminus \{i\}$ at some time t_k^i. In what follows, we aim to estimate the largest possible time $t_{k+1}^i > t_k^i$ that an agent i can wait for in order to update information about neighbors while guaranteeing convergence to the set of NE of the game. To do this, we assume that each player has available up-to-date information about its own state at every time $t > t_0^i$. The i^{th} player's action is driven by

$$\dot{x}_i(t) \in \begin{cases} \text{BR}_i(x_{-i}(t_k^i)) - x_i(t), & \text{if } x_i(t_k^i) \notin \text{BR}_i(x_{-i}(t_k^i)), \\ \{0\}, & \text{otherwise,} \end{cases} \tag{6.4}$$

for time $t \in (t_k^i, t_{k+1}^i]$. See Remark 6.1 about the introduction of zero when $x_i(t_k^i) \in \text{BR}_i(x_{-i}(t_k^i))$.

In the sequel, let us assume that each agent payoff function is Lipschitz continuous with Lipschitz constant $L_i > 0$; that is, $|u_i(x^1) - u_i(x^2)| \leq L_i \|x^1 - x^2\|$, for any $x^1, x^2 \in X$. Let us assume that player i knows L_i. This will help us to compute a self-triggering condition which makes each agent update information whenever its payoff is no longer increasing. First, let us find an upper bound on uncertainty about other player's action with respect to time.

At time t_k^i, player i knows other players' actions, and thus can compute pre-

cisely its best-response set, as well as the value of $u_i(x_i^\star(t_k^i), x_{-i}(t_k^i))$, where $x_i^\star(t_k^i) \in \mathrm{BR}_i(x_{-i}(t_k^i))$. Let $j \in I$ be an arbitrary agent $j \neq i$. Let l be such that $t_{l+1}^j > t_k^i \geq t_l^j$ for the given k. Notice that since $\mathrm{BR}_j(x_{-j}(t_l^j))$ is compact, there exists a point $x_j^{\text{fast}} \in \mathrm{BR}_j(x_{-j}(t_l^j))$ such that $x_j^{\text{fast}} \in \operatorname{argmax}_{y \in \mathrm{BR}_j(x_{-j}(t_l^j))} \|y - x_j(t_k^i)\|$. Then, the magnitude of $\dot{x}_j(t)$ defined in (6.4) is maximized by x_j^{fast}, for all time $t \in (t_k^i, t_{l+1}^j]$ (i.e., $\dot{x}_j(t) = x_j^{\text{fast}} - x_j(t)$ has maximum norm). Assume that $x_j(t_k^i) \notin \mathrm{BR}_j(t_l^j)$. Thus, a fastest solution of (6.4) for $t \in (t_k^i, t_{l+1}^j]$, is $x_j(t) = x_j^{\text{fast}} - (x_j^{\text{fast}} - x_j(t_k^i))e^{-(t-t_k^i)}$. This implies that the distance $\|x_j(t) - x_j(t_k^i)\|$ is upper bounded by $\|x_j^{\text{fast}} - x_j(t_k^i)\| \left(1 - e^{-(t-t_k^i)}\right)$, for $t \in (t_k^i, t_{l+1}^j]$. However, the i^{th} player does not know the j^{th} player's best-response set, then, the only option is to compute the worst possible case with the available information. Assume that all agents know the action space X. Then, the i^{th} agent can find a point $x_j^{\text{far}}(t_k^i) \in X_j$, which maximizes the distance from the last known position of j. That is, $x_j^{\text{far}}(t_k^i) \in \operatorname{argmax}_{y \in X_j} \|x_j(t_k^i) - y\|$. Then, $\|x_j^{\text{fast}} - x_j(t_k^i)\| \left(1 - e^{-(t-t_k^i)}\right) \leq \|x_j^{\text{far}} - x_j(t_k^i)\| \left(1 - e^{-(t-t_k^i)}\right)$, holds for every $t \in (t_k^i, t_{l+1}^j]$. Since the left-hand side of this inequality is an upper bound for the movement of j, for $t \in (t_k^i, t_{l+1}^j]$, we have that

$$\|x_j(t) - x_j(t_k^i)\| \leq \|x_j^{\text{far}}(t_k^i) - x_j(t_k^i)\| \left(1 - e^{-(t-t_k^i)}\right). \tag{6.5}$$

At this point, we can neglect the assumption of $x_j(t_k^i) \notin \mathrm{BR}_j(t_l^j)$, since, if $x_j(t_k^i) \in \mathrm{BR}_j(t_l^j)$, then the solution is trivially $x_j(t) = x_j(t_k^i)$, for $t \in (t_k^i, t_{l+1}^j]$. Therefore, the upper bound in (6.5) holds.

Lemma 6.1. *Inequality* (6.5) *holds for all* $t_k^i < t$.

Next, we can find an upper bound for $\|x_{-i}(t) - x_{-i}(t_k^i)\|$ as

$$\sum_{j \in I \setminus \{i\}} \|x_j(t_k^i) - x_j^{\text{far}}(t_k^i)\| \left(1 - e^{-(t-t_k^i)}\right) \geq \|x_{-i}(t) - x_{-i}(t_k^i)\|. \tag{6.6}$$

This upper bound only depends on information available to player i up to time t_k^i. We will use it next to determine the time-update instant t_{k+1}^i.

Next, we use the Lipschitz property of u_i, to obtain

$$u_i(x_i^\star(t_k^i), x_{-i}(t)) \geq u_i(x_i^\star(t_k^i), x_{-i}(t_k^i)) - L_i \|x_{-i}(t) - x_{-i}(t_k^i)\|, \tag{6.7}$$

and similarly,

$$u_i(x_i(t), x_{-i}(t_k^i)) \leq u_i(x_i(t_k^i), x_{-i}(t_k^i)) + L_i \|x_{-i}(t) - x_{-i}(t_k^i)\|. \tag{6.8}$$

The following lemma states a combination of the bounds in u_i in equations (6.7), (6.8), and the bound on $\|x_{-i}(t) - x_{-i}(t_k^i)\|$ from Lemma 6.1, to formulate the self-triggering update condition for each player.

Lemma 6.2. *Let $\Gamma = (I, X, u)$ be an ordinal potential game with potential function Φ, fulfilling all properties defined in Theorem 6.1. Assume that u is Lipschitz over X. Let W be a continuous fiber bundle over X such that $W_i(x)$ is compact and convex for all $x \in X$, $i \in I$. Let us consider the self-triggered best-response dynamics as defined in equation* (6.4). *Let $\varepsilon > 0$, and suppose $t_k^i > t_0$ is the last time instant when agent i updated information about other agents. Consider any $x_i^\star(t_k^i) \in \mathrm{BR}_i(x_{-i}(t_k^i))$. Let t_{wait}^i be a positive constant. If $t_{k+1}^i > t_k^i$ is such that either*

$$\begin{aligned} &u_i(x_i^\star(t_k^i), x_{-i}(t_k^i)) \\ &- 2L_i \sum_{j \in I \setminus \{i\}} \|x_j(t_k^i) - x_j^{\mathrm{far}}(t_k^i)\| \left(1 - e^{-(t_{k+1}^i - t_k^i)}\right) \\ &\qquad = u_i(x_i(t_{k+1}^i), x_{-i}(t_k^i)) + \varepsilon, \end{aligned} \tag{6.9}$$

provided $x_i(t_k^i) \notin \mathrm{BR}_i(x_{-i}(t_k^i))$, or $t_{k+1}^i = t_k^i + t_{\mathrm{wait}}^i$, if $x_i(t_k^i) \in \mathrm{BR}_i(x_{-i}(t_k^i))$, then it holds that $\Phi(x_i(t), x_{-i}(t)) < \Phi(x_i^\star(t_k^i), x_{-i}(t))$ for all $t \in (t_k^i, t_{k+1}^i]$, such that $x_i(t_k^i) \notin \mathrm{BR}_i(x_{-i}(t_k^i))$, and such that $\Phi(x_i(t), x_{-i}(t)) = \Phi(x_i(t_k^i), x_{-i}(t))$, for all $t \in (t_k^i, t_{k+1}^i]$, if $x_i(t_k^i) \in \mathrm{BR}_i(x_{-i}(t_k^i))$.

It also holds that $t_{k+1}^i - t_k^i \leq \max\left\{\log\left(\frac{2L_i N \max_{j \in I} \rho(X_j)}{2L_i N \max_{j \in I} \rho(X_j) - L_i \rho(X_i) - \varepsilon}\right), t_{\mathrm{wait}}^i\right\}$. This upper bound follows from (6.9), but we omit its computation for brevity. It will be important to establish precompactness of solutions in the analysis in Section 6.4.

Remark 6.1. *Here, we analyze the behavior of player i if its dynamics was given by*

$$\dot{x}_i \in \mathrm{BR}_i(x_{-i}(t_k^i)) - x_i(t), \tag{6.10}$$

for $t \in (t_k^i, t_{k+1}^i]$. If for some time in $(t_k^i, t_{k+1}^i]$, player i is not in its best-response set, the self-triggering time-update policy of Lemma 6.2 guarantees that the payoff at time

t is at worst less than the last known best payoff by some $\varepsilon > 0$*, provided* $x_i(t_k^i) \notin \mathrm{BR}_i(x_{-i}(t_k^i))$*. Notice that* t_{k+1}^i *is the maximum time when this property holds. Then, at* t_{k+1}^i *information is updated by player i and uncertainty becomes zero again, leading to a new best-response set. This will produce a larger or equal payoff than the current action's payoff. In particular, if the payoff value is the same, then* $x_i(t_{k+1}^i) \in \mathrm{BR}_i(x_{-i}(t_{k+1}^i))$*.*

Suppose a player has reached its best-response set and follows the dynamics given by (6.10)*. Once in* $\mathrm{BR}_i(x_{-i}(t_k^i))$*, the motion of player i can evolve arbitrarily in the set. In the meantime, the evolution of* $\mathrm{BR}_i(x_{-i}(t))$ *can lead to a situation where* $\mathrm{BR}_i(x_{-i}(t)) \neq \mathrm{BR}_i(x_{-i}(t_k^i))$*, while* $x_i(t) \in \mathrm{BR}_i(x_{-i}(t))$*. In this case, moving toward a point* $y \in \mathrm{BR}_i(x_{-i}(t_k^i)) \setminus \mathrm{BR}_i(x_{-i}(t))$ *will clearly produce a lower payoff. Thus, the set of velocities that an agent can take needs to be restricted, and it makes necessary to estimate how the best-response set will evolve. Alternatively, one can leverage the fact that* $x_i(t_k^i) \in \mathrm{BR}_i(x_{-i}(t_k^i))$*, to prescribe the agent velocity to be zero. This motivates the definition of a self-triggered best-response dynamics as in* (6.4)*, and not as in* (6.10)*. By means of this, one can guarantee* $\Phi(x_i(t_k^i), x_{-i}(t)) = \Phi(x_i(t), x_{-i}(t))$ *if and only if* $x_i(t) \in \mathrm{BR}_i(x_{-i}(t_k^i))$*.* ⋄

Remark 6.2. *The self-triggered best-response dynamics in* (6.4) *may lead to a zeno-behavior in some examples. That is, as agents approach their best-response sets, they may require information updates more and more often, creating an accumulation point in the time-update sequence. This is a typical trait of general event and self-triggered dynamics. In general, the only way to guarantee a lower bound on the time between updates by this approach is to force it, for example by taking* $\max\{t_i^k + \Delta t_{\min}, t_i^{k+1}\}$*, where* $\Delta t_{\min}$ *is a small positive number. Introducing this constant is an acceptable trade-off: on the one hand, the nature of the self-triggered approach is still preserved as much as possible, i.e., if possible, communications will be reduced by being triggered at times larger than* $\Delta t_{\min}$ *until being close to converge, leading to a type of practical convergence. On the other hand, the zeno-behavior is forced to disappear.* ⋄

6.4 System analysis via invariance theory

In order to formally analyze the self-triggered best-response dynamics, we over-approximate it by means of a larger hybrid system whose solutions include those of interest. To do this, first we associate each agent with a data structure $P^i = (x_i, x^i_{-i}, t_i) \in X \times \mathbb{R}_{\geq 0}$, where $x^i_{-i} = (x^i_j)_{j \in I} \in X_{-i}$ represents the information that agent i maintains on all other agents $j \neq i$, i.e., $x^i_{-i}(t) = x_{-i}(t^i_k)$ for $t \in (t^i_k, t^i_{k+1}]$, and $t_i = t - t^i_k$ for $t \in (t^i_k, t^i_{k+1}]$. Then, $P = (P^1, \ldots, P^N) \in (X \times \mathbb{R}_{\geq 0})^N = O$ is an extended state that includes the data structure P^i for each agent. Finally, let us define the projection $\pi : O \to \prod_{i \in I} \mathbb{R}^{n_i}$ as $\pi(P) = (\pi_i(P)) = (x_1, \ldots, x_N)$. Using this new notation, we can write the self-triggering condition of Lemma 6.2 as $\Delta_i(P) \leq 0$, where

$$\begin{aligned} \Delta_i(P) =& u_i(x^\star_i, x^i_{-i}) \\ & - 2L_i \sum_{j \in I \setminus \{i\}} \|x^i_j - x^{\text{far}}_j\| \left(1 - e^{-t_i}\right) \\ & - u_i(x_i, x^i_{-i}) - \varepsilon, \end{aligned} \tag{6.11}$$

with $\|x^i_j - x^{\text{far}}_j\| = \max_{y \in X_j} \|x^i_j - y\|$, and $x^\star_i \in \text{BR}_i(x^i_{-i})$. We now define a hybrid system on $O = (\mathbb{R}^d \times \mathbb{R})^N$ as follows. First, let $C \subset O$ be the set $C = \cap_{i \in I} C_i = \cap_{i \in I} (\{P \in O \mid \Delta_i(P) \geq 0,\ x_i \in \overline{W_i(x_i, x^i_{-i}) \setminus \text{BR}_i(x^i_{-i})}\} \cup \{P \in O \mid t_i \leq t^i_{\text{wait}} \text{ and } x_i \in \text{BR}_i(x^i_{-i})\})$. Secondly, we let $D = \cup_{i \in I} (\{P \in O \mid \Delta_i(P) \leq 0 \text{ and } x_i \in \overline{W_i(x_i, x^i_{-i}) \setminus \text{BR}_i(x^i_{-i})}\} \cup \{P \in O \mid t_i \geq t^i_{\text{wait}} \text{ and } x_i \in \text{BR}_i(x^i_{-i})\})$. Define the flow map $F : O \rightrightarrows O$ as $F(P) = \Pi_{i \in I} F_i(P)$, with $F_i(P) = \{(x^\star_i - \pi_i(P), 0, 1) \mid x^\star_i \in \text{BR}_i(x^i_{-i})\}$, for all $i \in I$. Define the jump map $G : O \rightrightarrows O$ so that $Y \in G(P)$ if and only if $Y^i \in \{P, (x_i, x_{-i}, 0)\}$, for each $i \in I$. Finally, define the hybrid system $\mathcal{H} = (F, G, C, D)$ as

$$\mathcal{H} : \begin{cases} \dot{P} \in F(P), & \text{if } P \in C, \\ P^+ \in G(P), & \text{if } P \in D. \end{cases}$$

Solutions for this system are given by functions $\phi : E \to O$, such that for each $j \in \mathbb{N}$ it holds that $t \mapsto \phi(t, j)$ is locally absolutely continuous on the interval $I^j = \{t \in \mathbb{R}_{\geq 0} \mid (t, j) \in E\}$, where E is a hybrid domain; see [79] for the definition of this concept. Let $S_{\mathcal{H}}$ be the set of all solutions of $\mathcal{H}$. By definition of $\mathcal{H}$, for each $P \in D$, it holds

that $P \in G(P)$. It means that the hybrid system overapproximation generates solutions that remain at the same fixed point P via infinite switching. However, note that these are not solutions of the self-triggered best-response dynamics. Additionally, the set $S_{\mathcal{H}}$ contains trajectories that allow motion inside $\mathrm{BR}_i(x^i_{-i})$ when x_i has reached $\mathrm{BR}_i(x^i_{-i})$, see Remark 6.1. Given that Δ_i is a continuous function of P, the sets C and D are closed sets in O. Under the assumption that the u_i are Lipschitz over X, and Φ is component-wise quasiconcave, one can see that F has compact, convex values, it is also locally bounded, and outer semicontinuous in C. The map G is outer semicontinuous by construction.

Let $\Psi = -\Phi$, and consider its extension $\tilde{\Psi} : O \to \mathbb{R}$ defined as $\tilde{\Psi}(P) = \Psi(\pi(P)) = \Psi(x_1, \ldots, x_N)$. In this way, $\tilde{\Psi}$ is a continuous function on O, and a locally Lipschitz function on a neighborhood of C. We now focus on the trajectories of $\mathcal{H}$ whose velocities take values in a subset of the differential inclusion. In other words, we define $\overline{F} : O \rightrightarrows O$ as $\overline{F} = \Pi_{i\in I}\overline{F}_i(P)$, where

$$\overline{F}_i(P) = \begin{cases} (0,0,1), & \text{if } x_i \notin \mathrm{int}(\mathrm{BR}_i(x^i_{-i})), \\ (\mathrm{BR}_i(x^i_{-i}) - x_i, 0, 1), & \text{otherwise.} \end{cases}$$

We have that $\overline{F}(P) \subseteq F(P)$ for all $P \in O$. Note that $\overline{F}$ selects the velocities according to the self-triggered dynamics.

Lemma 6.3. *For all $P \in C$ it holds that* $\max \mathcal{L}_{\overline{F}}\tilde{\Psi}(P) \leq 0$*, and for all $P \in D$, it holds that* $\max_{P^+\in G(P)} \tilde{\Psi}(P^+) - \tilde{\Psi}(P) \leq 0$*. Moreover, if $P \in C$ is such that for some $i \in I$, $x_i \notin \mathrm{BR}_i(x^i_{-i})$, then* $\max \mathcal{L}_{\overline{F}}\tilde{\Psi}(P) < 0$*.*

Proof. The condition $\max_{P^+\in G(P)} \tilde{\Psi}(P^+) - \tilde{\Psi}(P) \leq 0$ holds trivially, as $\tilde{\Psi}(P^+) = \tilde{\Psi}(P)$ for any $P \in O$ and $P^+ \in G(P)$. In order to verify the first condition, we follow along the lines of the proof of Theorem 6.1.

Consider $P \in C$. Any $V \in \overline{F}(P)$ can be written as $V = (V^1, \ldots, V^N)$, where each component V^i has the form $V^i = (x_i^\star - x_i, 0, 1)$, for some $x_i^\star \in \mathrm{BR}_i(x^i_{-i})$ if $x_i \notin \mathrm{BR}_i(x^i_{-i})$, or $V^i = (0,0,1)$ if $x_i \in \mathrm{BR}_i(x^i_{-i})$. Let us write $V = P^\star - P$, for an appropriate $P^\star$. Consider any $\zeta \in \partial\tilde{\Psi}(P)$ such that $\zeta = \lim_k \nabla\tilde{\Psi}(Y^k)$, with $Y^k \to P$, and $Y^k \notin \Omega_{\tilde{\Psi}}$. For convenience, let us recall that $\Omega_{\tilde{\Psi}}$ is the set of points at which $\tilde{\Psi}$ is non-differentiable. Since $\tilde{\Psi}$ is independent of the components x^i_{-i}, t_i, note that $\zeta^i = (\pi_i(\zeta), 0, 0)$. If $V^i = (0,0,1)$ for

all i, then it holds trivially that $\zeta^T V = 0$. Suppose then that $V^i \neq 0$ for some $i \in I$. Since $Y^k \to P$ we can write $V^T\zeta = \lim_{k\to\infty}(P^\star - Y^k)\nabla\tilde{\Psi}(Y^k)$. Denote and $y^k = \pi(Y^k)$. Using component-wise pseudoconvexity and component-wise quasiconvexity of Ψ, the computations in the proof of Theorem 6.1 can be repeated until we reach that $V^T\zeta \leq \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(x_i^\star, x_i)$, for all $\zeta \in \partial\tilde{\Psi}(P)$.

Now, for $P \in C$, either the condition $\Delta_i(P) \geq 0$ holds for some i or $x_i \in \mathrm{BR}_i(x^i_{-i})$, for all $i \in I$. Then, by Lemma 6.2 it is true that $\Psi(x_i^\star, x_{-i}) - \Psi(x) \leq 0$, for all $i \in I$. Thus, we have that $V^T\zeta \leq 0$. The result can be extended for any $\zeta \in \partial\tilde{\Psi}(x)$ similarly to Theorem 6.1. Therefore, $\max_{\overline{F}} \mathcal{L}_{\overline{F}}\tilde{\Psi}(P) = \max_{V,\zeta} V^T\zeta \leq 0$.

To prove the second part of this lemma, note that $\max_{\overline{F}} \mathcal{L}_{\overline{F}}\tilde{\Psi}(P) = \max_{V,\zeta} V^T\zeta$, and $\max_{V,\zeta} V^T\zeta \leq \max_{x^\star\in\{y \mid y_i\in \mathrm{BR}_i(x^i_{-i})\}} \sum_{\substack{i\in I \\ x_i\notin \mathrm{BR}_i(x)}} b_i(x_i^\star, x_i)$. Since $b_i(x_i^\star, x_i)$ is continuous for all $i \in I$, and $\mathrm{BR}_i(x^i_{-i})$ is a compact set, the right-hand side of the above inequality achieves its maximum at some $\bar{x}_i^\star \in \mathrm{BR}_i(x^i_{-i})$ for all $i \in I$. Then, if for some $i \in I$, $x_i \notin \mathrm{BR}_i(x^i_{-i})$, since $P \in C$, it must be that $\Delta_i(P) \geq 0$. By Lemma 6.2, this implies $b_i(x_i^\star, x_i) < 0$ for all $x_i^\star \in \mathrm{BR}_i(x^i_{-i})$ and, in particular, $b_i(\bar{x}_i^\star, x_i) < 0$. Thus, the strict inequality $\max \mathcal{L}_{\overline{F}}\tilde{\Psi}(P) < 0$ follows.

□

Theorem 6.2. *Let $\Gamma = (I, X, u)$ be an ordinal potential game with potential function Φ, fulfilling all properties defined in Theorem 6.1. Assume that u is a Lipschitz continuous function over X. Let W be a continuous fiber bundle over X such that $W_i(x)$ is compact and convex for all $x \in X$, $i \in I$. Let* (6.4) *be the self-triggered best-response dynamics for $\mathcal{R}_W(\Gamma)$. Then, all precompact solutions of the self-triggered best-response dynamics converge to the set $X^\star$ of constrained Nash Equilibria.*

Proof. Consider a precompact solution ϕ of the self-triggered best response dynamics which, in particular, is a precompact solution of $\mathcal{H}$. Then, the ω-limit set $\Omega(\phi)$ is nonempty, compact, and weakly invariant [79]. Since $\tilde{\Psi}$ satisfies Lemma 6.3, then $\Omega(\phi) \subseteq \tilde{\Psi}^{-1}(r)$ for some r. First, the conditions in Lemma 6.3 imply that $\tilde{\Psi} \circ \phi$ is non-increasing and bounded below. Let r satisfy that $\lim_{t\to\infty, j\to\infty} \tilde{\Psi}(\phi(t, j)) = r$. Take any $P \in \Omega(\phi)$. By definition, $\lim_{l\to\infty} \phi(t_l, j_l) = P$, with $(t_l, j_l) \in E$. Since $\tilde{\Psi}$ is continuous, then it holds that $\lim_{k\to\infty} \tilde{\Psi}(\phi(t_k, j_k)) = \tilde{\Psi}(P) = r$.

Take a $P \in \Omega(\phi)$. Since $\Omega(\phi)$ is weakly forward invariant, there exists a solution

to the self-triggered best response dynamics such that $P = \phi(t_0, j_0)$ for some (t_0, j_0). Suppose that there is an i such that for $P_i = (x_i, x^i_{-i}, t_i)$, we have $x_i \notin \mathrm{BR}_i(x^i_{-i})$. Then, by Lemma 6.3 we have that $\max \mathcal{L}_{\overline{F}} \tilde{\Psi}(P) < 0$. If $P \in C$, and the solution flows, then it must be that $\tilde{\Psi}(\phi(t_0^+, j_0)) < r$, which contradicts $P \in \tilde{\Psi}^{-1}(r)$. Thus, it must be that $P \in \overline{D \setminus C}$. This implies that there exists an i such that either $\Delta_i(P) \leq 0$ or $t^i_{\mathrm{wait}} \leq t_i$. However, after the jump, $\overline{P} = \phi(t_0, j_0 + 1) \in C$. Then, either we have that $\pi_i(\overline{P}) \in \mathrm{BR}_i(\pi(\overline{P}))$, for some $i \in I$, in which case the conclusion follows, or else it will continue flowing afterwards, which leads to a contradiction again. □

6.5 Summary

In this chapter, we characterize the convergence properties of the continuous time best-response dynamics for a continuous-action-space potential game, with N players and n_i-dimensional action space for each player i. We show under general conditions that all solutions of the best-response dynamics of a potential game will converge to the set of Nash equilibria set of the game. With the aim of making the best-response dynamics more practical, a self-triggered communication strategy is proposed to reduce communications among agents while still guaranteeing convergence to the desired configurations. The self-triggered best response dynamics is modeled as a hybrid system, and convergence analysis is made using analysis tools in [79].

As a future line of study, the effects of delays in the self-triggered best response dynamics can be analyzed.

Acknowledgments

This paper contains material that has been published in the following works:

- A. Cortés and S. Martínez, "Self-triggered Best Response Dynamics for Continuous Games", *IEEE Transactions on Automatic Control*, vol 30 (4), 2015, pp 1115-1120.

- A. Cortés and S. Martínez, "Self-triggered Best Response Dynamics for Mobile

Sensors Deployment", *Proceedings of the American Control Conference*, 2015, pp 6370-6375.

Chapter 7

Conclusions

In this thesis, we present multi-agent algorithms for coordination of resources in smart grids. In general, we propose various algorithms to coordinate storage, generation, and demand resources in a power grid using distributed computation and decentralized decision making.

We address three different energy management problems, each with a specific algorithm, which works under particular communication assumptions.

First, we address the problem of multi-agent coordination of Plug-in Electric Vehicles (PEVs) under two different paradigms, i) vehicle-1-grid (V1G), in which vehicles only act as flexible loads that can charge whenever it is more convenient for the grid/user performance, and ii) vehicle-2-grid (V2G), in which vehicles behave like batteries that can inject power back into the grid if needed. Both approaches use non-PEV load and usage forecast information for the decision making process. For the V1G problem, we introduce a load-balancing algorithm that uses information on the aggregate energy prices to drive PEV load to an optimal setting. The decision making process for each PEV is carried out by an agent that controls such PEV. The V2G problem introduces a hierarchical framework that computes an optimal V2G charging/discharging strategy, which also satisfies line capacity constraints in the distribution side of the grid.

Following, we introduce a hierarchical approach for indirect control of deferrable loads in a demand response event. This framework takes into account the fact that many loads only admit on/off control, and provides an algorithm that computes a sub-optimal solution that satisfies maximum power constraints. The algorithm does not require users

to share private information, but only an estimate of its demand. The solution for the on/off loads is computed in a greedy-like way, by computing a threshold to decide on the loads that must be turned off.

Finally, we present a framework that uses optimal AC power flow to decide on the optimal usage of storage and generation resources in a microgrid. This computation is distributed, and requires communication between neighboring nodes of a microgrid, according to certain network topology. This network topology is closely related to the microgrid topology.

The algorithms that we propose throughout this thesis are provably correct, and their convergence properties are established via theoretical results. In addition, we present simulations that complement the theoretical results. In particular, for the problem of control of resources in microgrids, we present a simulation study to compare the two proposed algorithms under different scenarios. We have found that the first approach was prohibitively slow, which encouraged us to develop the second approach, which presented an improvement on the speed of convergence of more than two orders of magnitude.

As a future direction, we propose to evaluate the practical implementation of the proposed algorithms in real life scenarios. The impact of communications required for the algorithms to compute a solution on the operational costs of a power systems must be studied. Additionally, the trade-off between benefits for end-users and benefits for utility companies and system operators in the obtained solutions is an open problem that should be addressed.

Bibliography

[1] A. Domínguez-García, C. N. Hadjicostis, Distributed algorithms for control of demand response and distributed energy resources, in: IEEE Int. Conf. on Decision and Control, Florida, USA, 2011, pp. 27–32.

[2] A. Domínguez-García, S. Cady, C. Hadjicostis, Decentralized optimal dispatch of distributed energy resources, in: IEEE Int. Conf. on Decision and Control, 2012, pp. 3688–3693.

[3] Z. Zhang, X. Ying, M. Chow, Decentralizing the economic dispatch problem using a two-level incremental cost consensus algorithm in a smart grid environment, in: North American Power Symposium, Boston, MA, 2011, pp. 1–7, electronic Proceedings.

[4] A. Cherukuri, J. Cortés, Distributed generator coordination for initialization and anytime optimization in economic dispatch, IEEE Transactions on Control of Network SystemsConditionally accepted. Available at `http://carmenere.ucsd.edu/jorge`.

[5] A. Cherukuri, S. Martínez, J. Cortés, Distributed, anytime optimization in power-generator networks for economic dispatch, in: 2014 American Control Conference, Portland, OR, 2014.

[6] S. Bolognani, R. Carli, G. Cavraro, S. Zampieri, A distributed feedback control strategy for optimal reactive power flow with voltage constraints, arXiv preprint arXiv:1303.7173.

[7] S. Bolognani, R. Carli, G. Cavraro, S. Zampieri, Distributed reactive power feedback control for voltage regulation and loss minimization, IEEE Transactions on Automatic Control 60 (4) (2015) 966–981.

[8] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, Foundations and Trends in Machine Learning 3 (1) (2010) 1–122.

[9] E. Wei, A. Ozdaglar, On the O(1/k) convergence of asynchronous distributed alternating direction method of multipliers, Mathematical ProgrammingSubmitted,Preprint available at http://arxiv.org/abs/1307.8254.

[10] M. Zhongjing, D. S. Callaway, I. A. Hiskens, Decentralized charging control of large populations of plug-in electric vehicles, IEEE Transactions on Control Systems Technology 21 (1) (2011) 67–78.

[11] L. Gan, U. Topcu, S. H. Low, Optimal decentralized protocol for electric vehicle charging, IEEE Transactions on Power Systems 28 (2) (2013) 940–951.

[12] A. Cortés, S. Martínez, Optimal plug-in electric vehicle charging with schedule constraints, in: Allerton Conf. on Communications, Control and Computing, 2013, pp. 262–266.

[13] S. Shafiei, R. Izadi-Zamanabadi, H. Rasmussen, J. Stoustrup, A decentralized control method for direct smart grid control of refrigeration systems, in: IEEE Int. Conf. on Decision and Control, IEEE, 2013, pp. 6934–6939.

[14] R. Pedersen, J. Schwensen, B. Biegel, J. Stoustrup, T. Green, Aggregation and control of supermarket refrigeration systems in a smart grid, in: IFAC World Congress, 2014, pp. 9942–9949.

[15] K. Mets, T. Verschueren, W. Haerick, C. Develder, F. D. Turck, Optimizing smart energy control strategies for plug-in hybrid electric vehicle charging, in: IEEE/IFIP Network Operations and Management Symposium (NOMS) Workshops, 2010, pp. 293–299.

[16] A. S. Masoum, S. Deilami, P. S. Moses, M. A. S. Masoum, A. Abu-Siada, Smart load management of plug-in electric vehicles in distribution and residential networks with charging stations for peak shaving and loss minimisation considering voltage regulation, IET Generation, Transmission & Distribution 5 (8) (2011) 877–888.

[17] M. Caramanis, J. M. Foster, Management of electric vehicle charging to mitigate renewable generation intermittency and distribution network congestion, in: IEEE Int. Conf. on Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009, 2009, pp. 4717–4722.

[18] O. Ardakanian, C. Rosenberg, S. Keshav, Distributed control of electric vehicle charging, in: Proceedings of the fourth international conference on Future energy systems, 2013, pp. 101–112.

[19] B. Gharesifard, T. Basar, A. D. Dominguez-Garcia, Price-based distributed control for networked plug-in electric vehicles, in: American Control Conference, 2013, pp. 5093–5098.

[20] C. Ahn, C. Li, H. Peng, Optimal decentralized charging control algorithm for electrified vehicles connected to smart grid, Journal of Power Sources 196 (23) (2011) 10369–10379.

[21] W. Ma, V. Gupta, U. Topcu, On distributed charging control of electric vehicle with power network capacity constraints, in: American Control Conference, 2014, pp. 4306–4311.

[22] Y. Ru, J. Kleissl, S. Martínez, Storage size determination for grid-connected PV systems, IEEE Transactions on Sustainable Energy 4 (1) (2013) 68–81.

[23] A. Arsie, C. Ebenbauer, Refining LaSalle's invariance principle, in: American Control Conference, 2009, pp. 108–112.

[24] J. P. Aubin, H. Frankowska, Set-valued analysis, Birkhäuser, 1990.

[25] J. A. P. Lopes, F. J. Soares, P. M. R. Almeida, Integration of electric vehicles in the electric power system, Proceedings of the IEEE 99 (1) (2011) 168–183.

[26] T. Sousa, Z. Vale, J. P. Carvalho, T. Pinto, H. Morais, A hybrid simulated annealing approach to handle energy resource management considering an intensive use of electric vehicles, Energy 67 (2014) 81–96.

[27] A. T. Al-Awami, E. Sortomme, Coordinating vehicle-to-grid services with energy trading, IEEE Transactions on Smart Grid 3 (1) (2012) 453–462.

[28] C. Wu, H. Mohsenian-Rad, J. Huang, Vehicle-to-aggregator interaction game, IEEE Transactions on Smart Grid 3 (1) (2012) 434–442.

[29] D. P. Bertsekas, Necessary and sufficient conditions for a penalty method to be exact, Mathematical Programming 9 (1) (1975) 87–99.

[30] S. Boyd, L. Vandenberghe, Convex Optimization, Cambridge University Press, 2004.

[31] D. P. Palomar, Y. C. Eldar, Convex optimization in signal processing and communications, Cambridge University Press, 2010.

[32] B. Biegel, P. Andersen, T. Pedersen, K. Nielsen, J. Stoustrup, L. Hansen, Electricity market optimization of heat pump portfolio, in: IEEE Conf. on Control Applications, 2013, pp. 294–301.

[33] F. Tahersima, P. Andersen, P. P. Madsen, Economic energy distribution and consumption in a microgrid part1: Cell level controller, in: IEEE Conf. on Control Applications, 2013, pp. 308–313.

[34] M. Avci, M. Erkoc, A. Rahmani, S. Asfour, Model predictive HVAC load control in buildings using real-time electricity pricing, Energy and Buildings 60 (2013) 199–209.

[35] A. Agnetis, G. Dellino, P. Detti, G. Innocenti, G. de Pascale, A. Vicino, Appliance operation scheduling for electricity consumption optimization, in: IEEE Int. Conf. on Decision and Control, 2011, pp. 5899–5904.

[36] L. Gkatzikis, T. Salonidis, N. Hegde, L. Massoulié, Electricity markets meet the home through demand response, in: IEEE Int. Conf. on Decision and Control, 2012, pp. 5846–5851.

[37] B. Biegel, P. Andersen, T. Pedersen, K. Nielsen, J. Stoustrup, L. Hansen, Smart grid dispatch strategy for on/off demand-side devices, in: European Control Conference, 2013, pp. 2541–2548.

[38] E. Kutanoglu, S. Wu, On combinatorial auction and Lagrangean relaxation for distributed resource scheduling, IIE transactions 31 (9) (1999) 813–826.

[39] D. P. Bertsekas, The auction algorithm: a distributed relaxation method for the assignment problem, Annals of Operations Research 14 (1988) 105–123.

[40] K. Hirayama, A distributed solution protocol that computes an upper bound for the generalized mutual assignment problem, in: Proceedings of the 7th International Workshop on Distributed Constraint Reasoning, 2006, pp. 102–116.

[41] Q. Wang, C. Zhang, Y. Ding, G. Xydis, J. Wang, J. Østergaard, Review of real-time electricity markets for integrating distributed energy resources and demand response, Applied Energy 138 (2015) 695–706.

[42] Y. Ma, G. Anderson, F. Borrelli, A distributed predictive control approach to building temperature regulation, in: American Control Conference, IEEE, 2011, pp. 2089–2094.

[43] N. Chatzipanagiotis, D. Dentcheva, M. Zavlanos, An augmented Lagrangian method for distributed optimization, Mathematical Programming (2013) 1–30.

[44] E. Camacho, T. Samad, M. Garcia-Sanz, I. Hiskens, Control for renewable energy and smart grids, The Impact of Control Technology, Control Systems Society (2011) 69–88.

[45] A. Cortés, S. Martínez, Decentralized management of demand response events with on/off loads, preprint available at http://fausto.dynamic.ucsd.edu/andres.

[46] J. Carpentier, Contribution to the economic dispatch problem, Bulletin de la Societe Francoise des Electriciens 3 (8) (1962) 431–447.

[47] R. A. Jabr, Radial distribution load flow using conic programming, IEEE Transactions on Power Systems 21 (3) (2006) 1458–1459.

[48] J. Lavaei, S. H. Low, Zero duality gap in optimal power flow problem, IEEE Transactions on Power Systems 27 (1) (2012) 92–107.

[49] J. Lavaei, S. H. Low, Convexification of optimal power flow problem, in: Allerton Conf. on Communications, Control and Computing, 2010, pp. 223–232.

[50] J. Lavaei, Zero duality gap for classical opf problem convexifies fundamental nonlinear power problems, in: American Control Conference, 2011, pp. 4566–4573.

[51] K. M. Chandy, S. H. Low, U. Topcu, H. Xu, A simple optimal power flow model with energy storage, in: IEEE Int. Conf. on Decision and Control, 2010, pp. 1051–1057.

[52] D. Gayme, U. Topcu, Optimal power flow with large-scale storage integration, IEEE Transactions on Power Systems 28 (2) (2013) 709–717.

[53] M. E. Baran, F. F. Wu, Optimal sizing of capacitors placed on a radial distribution system, IEEE Transactions on Power Delivery 4 (1) (1989) 735–743.

[54] K. R. C. Mamandur, R. D. Chenoweth, Optimal control of reactive power flow for improvements in voltage profiles and for real power loss minimization, IEEE Transactions on Power Apparatus and Systems (7) (1981) 3185–3194.

[55] S. Bolognani, S. Zampieri, A distributed control strategy for reactive power compensation in smart microgrids, IEEE Transactions on Automatic Control 58 (11) (2013) 2818–2833.

[56] S. Liu, Matrix results on the Khatri-Rao and Tracy-Singh products, Linear Algebra and its Applications 289 (1) (1999) 267–277.

[57] F. Zhang, The Schur complement and its applications, Vol. 4, Springer, 2005.

[58] F. Zhang, Matrix Theory. Basic Results and Techniques, 2nd Edition, Universitext, Springer, 2011.

[59] A. Nedić, A. Ozdaglar, Subgradient methods for saddle-point problems, Journal of Optimization Theory & Applications 142 (1) (2009) 205–228.

[60] A. Cortés, S. Martínez, Distributed control of reactive power and storage in microgrids, in: Mathematical Theory of Networks and Systems, Groningen, The Netherlands, 2014, pp. 563–570.

[61] F. E. Udwadia, Some convergence results related to the JOR iterative method for symmetric, positive-definite matrices, Applied Mathematics and Computation 47 (1) (1992) 37–45.

[62] C. Harris, On the rate of convergence of continuous-time fictitious play, Games and Economic Behavior 22 (2) (1998) 238–259.

[63] M. Benaïm, J. Hofbauer, S. Sorin, Stochastic approximations and differential inclusions, SIAM Journal on Control and Optimization 44 (1) (2005) 328–348.

[64] J. Hofbauer, S. Sorin, Best response dynamics for continuous zero-sum games, Discrete and Continuous Dynamical Systems Series B 6 (1) (2006) 215.

[65] E. N. Barron, R. Goebel, R. R. Jensen, Best response dynamics for continuous games, Proceedings of the American Mathematical Society 138 (3) (2010) 1069–1083.

[66] A. Anta, P. Tabuada, To sample or not to sample: Self-triggered control for nonlinear systems, IEEE Transactions on Automatic Control 55 (9) (2010) 2030–2042.

[67] E. Garcia, P. J. Antsaklis, Model-based event-triggered control for systems with quantization and time-varying network delays, IEEE Transactions on Automatic Control 58 (2) (2013) 422–434.

[68] P. Wan, M. Lemmon, Event-triggered distributed optimization in sensor networks, in: Symposium on Information Processing of Sensor Networks, San Francisco, CA, 2009, pp. 49–60.

[69] C. Nowzari, J. Cortés, Self-triggered coordination of robotic networks for optimal deployment, Automatica 48 (6) (2012) 1077–1087.

[70] J. R. Marden, G. Arslan, J. S. Shamma, Cooperative control and potential games, IEEE Transactions on Systems, Man, & Cybernetics. Part B: Cybernetics 39 (6) (2009) 1393–1407.

[71] M. Zhu, S. Martínez, Distributed coverage games for energy-aware mobile sensor networks, SIAM Journal on Control and Optimization 51 (1) (2013) 1–27.

[72] A. Cortés, S. Martínez, Self-triggered best response dynamics for continuous games, preprint available at `http://fausto.dynamic.ucsd.edu/andres`.

[73] A. Cortés, S. Martínez, Self-triggered, best-response dynamics for sensor coverage, in: American Control Conference, Washington, DC, 2013, pp. 6385–6390.

[74] A. B. MacKenzie, L. A. DaSilva, Game theory for wireless engineers, Synthesis Lectures on Communications 1 (1) (2006) 1–86.

[75] D. Monderer, L. S. Shapley, Potential games, Games and Economic Behavior 14 (1996) 124–143.

[76] J. Hofbauer, K. Sigmund, Evolutionary game dynamics, Bulletin of the American Mathematical Society 40 (4) (2003) 479.

[77] A. Bacciotti, F. Ceragioli, Stability and stabilization of discontinuous systems and nonsmooth lyapunov functions., ESAIM: Control, Optimisation & Calculus of Variations 4 (1999) 361–376.

[78] E. A. Ok, Real analysis with economic applications, Princeton University Press, 2011.

[79] R. Goebel, R. G. Sanfelice, A. Teel, Hybrid dynamical systems, IEEE Control Systems Magazine 29 (2) (2009) 28–93.