Università degli studi di Genova
Dipartimento di Informatica, Bioingegneria,
Robotica ed Ingegneria dei Sistemi

# Competitive and Cooperative Approaches to the Balancing Market in Distribution Grids

by

Giulio Ferro

Theses Series **DIBRIS-TH-2020-**

DIBRIS, Università di Genova
Via Opera Pia, 13 16145 Genova, Italy http://www.dibris.unige.it/

**Università degli Studi di Genova**

**Dipartimento di Informatica, Bioingegneria,**

**Robotica ed Ingegneria dei Sistemi**

**Ph.D. Thesis in Computer Science and Systems Engineering**
**Systems Engineering Curriculum**

# Competitive and Cooperative Approaches to the Balancing Market in Distribution Grids

by

Giulio Ferro

May, 2020

**Dottorato di Ricerca in Informatica ed Ingegneria dei Sistemi**
**Indirizzo Informatica**
**Dipartimento di Informatica, Bioingegneria, Robotica ed Ingegneria dei Sistemi**
**Università degli Studi di Genova**

DIBRIS, Univ. di Genova
Via Opera Pia, 13
I-16145 Genova, Italy
`http://www.dibris.unige.it/`

**PhD Thesis in Computer Science and Systems Engineering**
**Systems Engineering Curriculum**
(S.S.D. ING INF/04)

Submitted by Giulio Ferro DIBRIS, Univ. di Genova

Date of submission: 09/03/2020

Title: Competitive and Cooperative Approaches to the Balancing Market in Distribution Grids

Advisor: Prof. Michela Robba.
Dipartimento di Informatica, Bioingegneria, Robotica ed Ingegneria dei Sistemi
Università di Genova
Co-advisor: Prof. Riccardo Minciardi.
Dipartimento di Informatica, Bioingegneria, Robotica ed Ingegneria dei Sistemi
Università di Genova
Co-advisor: Dr. Anuradha M. Annaswamy
Active-Adaptive Control (AAC) Laboratory Massachusetts Institute of Technology (MIT)

Ext. Reviewers:

Prof. Mariagrazia Dotoli. Dipartimento Ingegneria Elettrica e dell'Informazione Politecnico di Bari

Prof. Luigi Gliemo. Dipartimento di Ingegneria Università degli studi del Sannio

Prof. Riccardo Scattolini. Dipartimento di Elettronica, Informazione e Bioingegneria Politecnico di Milano

# Acknowledgements

These thanks are for all those who have contributed to the realization of my thesis and my Ph.D. path.

A huge thank you goes to my tutor, Prof Michela Robba, who tolerated and supported me throughout my entire Ph.D. path, even though on some occasions I was not the best Ph.D. student ever (especially when speaking of calm and peacefulness). Thank you for helping me grow professionally and teaching me how to be a good researcher (and person).

My gratitude goes to the two co tutors of this thesis, Prof Riccardo Minciardi and Dr. Anuradha Annaswamy from MIT, for their valuable advice and teachings and for having introduced me to new scientific fields, thus widening my horizons.

A special thanks also go to prof Mansueto Rossi, who wanted to accompany me also in this Ph.D. adventure, having already been my supervisor for the master thesis, always continuing to be a mentor for me in the field of power systems.

I would like to thank all those with whom I have had the chance to collaborate in these three years on different thematic areas from those of the thesis: Prof Roberto Sacile, Prof Massimo Paolucci, and Prof Federico Delfino.

I would like to express my gratitude to everyone who took part in my American adventure, starting with my office colleagues 3-441: Rabab, David, Yohan, and Bella; friends from MIT VISTA and MITaly associations. A big thank you to my parents who despite everything have accepted my life choices and supported me through this path.

Finally, I wish to thank all my friends, with whom I also shared my experiences, and who have lived with me most of my academic life.

# Abstract

The electrical grid has been changing in the last decade due to the presence, at the distribution level, of renewables, distributed generation, storage systems, microgrids, and electric vehicles. The introduction of new legislation and actors in the smart grid's system opens new challenges for the activities of companies, and the development of new energy management systems, models, and methods. In order to face this revolution, new market structures are being defined as well as new technologies and optimization and control algorithms for the management of distributed resources and the coordination of local users to contribute to active power reserve and ancillary services. One of the main problems for an electricity market operator that also owns the distribution grid is to avoid congestions and maximize the quality of the service provided. The thesis concerns the development and application of new methods for the optimization of network systems (with multi-decision makers) with particular attention to the case of power distribution networks

This Ph.D. thesis aims to address the current lack of properly defined market structures for the determination of balancing services in distribution networks. As a first study, to be able to handle the power flow equation in a computationally better way, a new convex relaxation has been proposed. Thereafter, two opposite types of market structure have been developed: competitive and cooperative. The first structure presents a two-tier mechanism where the market operator is in a predominant position compared to other market players. Vice versa in the cooperative mechanism (solved through distributed optimization techniques ) all actors are on the same level and work together for social welfare. The main methodological novelties of the proposed work are to solve complex problems with formally correct and computationally efficient techniques.

# Table of Contents

# Chapter 1

# Introduction and motivation

The increment of renewables resources use, distributed generation and storage, demand response programs, and electric vehicles charging stations has led to the necessity of changing the structure of the electrical grid into a more flexible one. In particularly distributed power production systems based on renewables, due to uncertainties associated with their presence can create some problems to the distribution grid, such as power quality losses in the power network, voltage unbalances, and undesired peaks. Usually, traditional controllable generators are called to compensate these fluctuations, operating at different working points from the optimal ones. This function for the traditional generators cause efficiency losses (increasing in operational and maintenance costs) and a decrease of regulation margins for the distribution grid.

Moreover, due to the increase of generation from renewables, large fossil-fueled production plants are no more installed or reduced in rated power; this affects the capability of the grid to respond to emergencies. The characteristics of this new smart grid require an increase of the power reserve to face the sudden request of active/reactive injection/absorption from a Distribution System Operator (DSO) to compensate for example a sudden drop in the production from a photovoltaic plant.

Furthermore, the introduction of new legislation and actors (such as aggregators, microgrids, and active buildings) in the smart grid's system opens new challenges for the activities of companies, and for the development of new energy management systems, models and methods. As an example, the figure of an aggregator that manages different local consumers and/or producers to reduce power demand or to increase production makes arise different new possible optimization problems.

In this paradigm, there are different new challenges both from applications, modeling, and methodological viewpoints. First of all, new formulations of control and optimization problems are necessary as a consequence of the development of new technologies and decision frameworks. As a matter of fact, in addition to classic problems such as power flow management, new decision problems based on the cooperation or competition of multi decision-makers should be defined and solved. The definition of new electricity market mechanisms is a clear example. Besides, since most of the optimization problems in power systems need to include the model of the electrical grid, i.e the so-called power flow equations (which define a non-linear and non-convex set), the most important mathematical challenge is to solve these optimization problems for large scale networks in computationally acceptable timescales.

The last main challenge is related to how to solve the optimization problems in a fast and reliable way considering uncertainties, privacy features, new technologies, and multiple decision-makers.

The motivations that give rise to this Ph.D. work, come from the recent proposals at the regulatory level to introduce power market mechanisms at the distribution level. Allowing the participation of distributed resources and the demand response (DR) to the market, both in a single and aggregate manner, to optimally dispatch energy also at the distribution level. As this is a completely new topic and without a real legislative framework, the literature is full of works proposing different market mechanisms for distribution networks. These studies mainly formalize decision-making problems based on economic considerations, based on the supply and demand curves of market participants, partially integrating electricity grid models, sometimes in a simplified way. Differently from the current literature, this network wants to analyze a market session related to the quantification of the balancing reserve provided by active market users, i.e. the Balancing Market (BM). This market typology takes place after the result of the main energy market where the resources are dispatched. The network operator, the DSO, through technical and economic considerations, introduces a BM session to quantify the energy reserve aiming to improve the quality of the overall network service. The focus of this Ph.D. thesis is on the development of new models and methods for the power distribution network which is experiencing a significant change of paradigm, i.e. from a passive network to an active one, giving rise to new and never faced problems. In particular, this thesis will focus mainly on multilevel and distributed architectures for the BM in distribution power grids.

The preliminary work undertaken with the study of the BM begins with the modeling of the various market participants and formulating performance indicators related to the technologies present in the entire system (generators and distribution network).

The first important task was to model the distribution grid in a convex way to develop optimization algorithms that guarantee the overall optimality of the solution obtained. Since the power flow equations are nonlinear and define a nonconvex space, it was necessary to formulate a convex relaxation of the above equations ensuring a small distance from the global optimal solution obtained through a nonlinear solver. In contrast to the current approaches in the literature, which can formulate convex relaxations only for certain types of networks, the approach proposed in the thesis manages to describe any type of network, keeping the whole error limited in acceptable bounds.

The second part of the thesis is devoted to the proposal and development of two possible BM structures for distribution networks both based on the resolution of optimization problems: the first competitive and the second cooperative. The first structure wants to solve the multi decision-maker problem of the BM in a centralized and competitive way each agent does not cooperate with the others being completely selfish from the decision point of view. The problem has been solved through the use of multilevel programming. Finally, the second market structure wants to solve the problem of BM optimization through distributed optimization. As a matter of fact, the consensus technique is applied to be able to solve the problem cooperatively, i.e. each decision-maker cooperates with his neighbors to reach the global optimum.

## 1.1 Main contributions

The main contribution of this Ph.D. thesis can be summarized as:

- The statement of a new convex relaxation of the OPF problem based on McCormick envelopes. This new model can afford any type of distribution network with an unbalanced structure and mesh topology, without binding working hypothesis.
- The development of an explicit procedure that states optimal upper and lower bounds for injected current increasing the quality of the proposed relaxation.
- The statement a new multi decision-maker architecture to describe BMs for a distribution network, taking into account a new convex relaxation of the electrical grid constraints.
- The development of a bi-level BM structure based on multilevel programming, in which every low-level market participant shares its KKT conditions to the upper-level problem to guarantee the optimality of the problem's solution.
- The development of a distributed algorithm, based on the Proximal Atomic Coordination (PAC) method, with a clear articulation of conditions for proof of convergence.
- The definition of a cooperative market structure for BM based on distributed optimization

## 1.2 Thesis' organization

This thesis is the collection of several research works. The chapters are based on lecture notes, journal or conference papers, which are either published or currently under review that is reported at the beginning of each Chapter. This thesis consists of eight Chapters, namely:

- Chapter II presents and discusses the models for the balancing market participants, describing the mathematical model, the main issues related to this formulation and some possible approaches
- Chapter III aims to face one of the main problems in power systems, presenting a new convex relaxation for the Optimal power flow problem for unbalanced and mesh distribution grids
- Chapter IV shows the first possible balancing market approach, namely the competitive market structure, based on multilevel programming is proposed and tested over a different testbed

- Chapter V shows the second possible balancing market approach, i.e. an innovative cooperative balancing market structure based on a new distributed optimization algorithm, the PAC algorithm, developed during the research stage at the Active Adaptive Control (AAC) Lab at Massachusetts Institute of Technology (MIT).
- Chapter VI includes some concluding remarks and indicates possible future developments.
- Appendix A introduces some mathematical preliminaries and some basics concepts on power systems and the energy market helpful for the overall understanding of this Ph.D. thesis.
- Appendix B contains all the proofs regarding the distributed algorithm stated in Chapter V.
- Appendix C contains the data of the test networks used in this thesis.

## 1.3 List of Publications

The publications and patents list of Giulio Ferro is hereafter reported.

**Journal Papers-Published**

- Ferro, G., Minciardi, R., Parodi, L., Robba, M., Rossi, M. "Optimal Control of Multiple Microgrids and Buildings by an Aggregator" (2020) *Energies*, 13(5), 1-26.
- G. Ferro, R. Minciardi, M. Robba "A user equilibrium model for electric vehicles: joint traffic and energy demand assignment" (2020), published online *Energy*.
- F Delfino, G Ferro, M Robba, M Rossi "An Energy Management Platform for the Optimal Control of Active and Reactive Powers in Sustainable Microgrids" *IEEE Transactions on Industry Applications 55* (6), 7146-7156
- Ferro, G., Laureri, F., Minciardi, R., Robba,M. "A predictive discrete event approach for the optimal charging of electric vehicles in microgrids" (2019) *Control engineering Practice* 86, pp. 11-23
- Delfino, F., Ferro, G., Minciardi, R., Robba, M., Rossi, M., Rossi, M. "Identification and optimal control of an electrical storage system for microgrids with renewables" (2019) *Sustainable Energy, Grids and Networks*, 17, art. no. 100183, .
- Bracco, S., Delfino, F., Ferro, G., Pagnini, L., Robba, M., Rossi, M. "Energy planning of sustainable districts: Towards the exploitation of small size intermittent renewables in urban areas" (2018) *Applied Energy*, 228, pp. 2288-2297.
- Delfino, F., Ferro, G., Robba, M., Rossi, M." An architecture for the optimal control of tertiary and secondary levels in small-size islanded microgrids" (2018) *International Journal of Electrical Power and Energy Systems*, 103, pp. 75-88.
- Ferro, G., Laureri, F., Minciardi, R., Robba, M. "An optimization model for electrical vehicles scheduling in a smart grid" (2018) *Sustainable Energy, Grids and Networks*, 14, pp. 62-70.

**Journal Papers-Submitted**

- J. Romvary, G, Ferro and A,M. Annaswamy, "A Proximal Atomic Coordination Algorithm for Distributed Optimization" submitted to *IEEE Transactions on Automatic Control* (under review after the first round of revision)

- G. Ferro, M. Robba, M. Rossi and A,M. Annaswamy, "A New Convex Relaxation of the Optimal PowerFlow problem for Distribution Grids" submitted to *IEEE Transactions on Power Systems*
- G. Ferro, M. Robba, R. Haider and A,M. Annaswamy, "A new distributed approach to the Optimal power flow problem in unbalanced and meshed distribution networks " submitted to *IEEE Transactions on Power Systems*
- G. Ferro, R. Minciardi, M. Robba and M. Rossi "A bilevel approach for balancing market' optimization in distribution grids in presence of distributed resources and aggregators " submitted to *IEEE Transactions on Power Systems*
- G. Ferro, M. Robba, R. Sacile "An optimal control strategy for interconnected systems in distribution grids: voltage and frequency regulation for islanded mode operation" Optimal Control, submitted to *Energies*
- G. Ferro, M. Robba, M. Paolucci "Optimal charging and routing of electric vehicles for goods transportation" Optimal Control, submitted to *IEEE Transactions on Smart Grid*

**Conference Papers**

- Ferro G., Minciardi R., Parodi L., Robba M., "Optimal planning of charging stations and electric vehicles traffic assignment: a bi-level approach" Accepted to IFAC 2020 conference
- Ferro G., Minciardi R., Parodi L., Robba M., Rossi M., "Optimal coordination of buildings and microgrids by an aggregator: a bi-level approach" Accepted to IFAC 2020 conference
- Ferro G., Robba M., Dachiardi D., Haider R., Annaswamy A., "A distributed approach to the Optimal Power Flow problem for unbalanced and mesh distribution networks" Accepted to IFAC 2020 conference
- Bracco, S., Brignone, M., Delfino, F., Pampararo, F., Rossi, M., Ferro, G., Robba, M. "An Optimization Model for Polygeneration Microgrids with Renewables, Electrical and Thermal Storage: Application to the Savona Campus" (2018) Proceedings - *2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe, EEEIC/I and CPS Europe 2018*, art. no. 8493965, .
- Ferro, G., Paolucci, M., Robba, M. "An Optimization Model For Electrical Vehicles Routing with time of use energy pricing and partial Recharging" (2018) *IFAC-PapersOnLine*, 51 (9), pp. 212-217.

- Ferro, G., Minciardi, R., Delfino, F., Rossi, M., Robba, M. "A bi-level approach for the management of microgrids" (2018) *IFAC-PapersOnLine*, 51 (28), pp. 309-314.

- Ferro, G., Laureri, F., Minciardi, R., Robba, M. "Optimal Integration of Interconnected Buildings in a Smart Grid: A Bi-level Approach" (2018) *Proceedings - 2017 UKSim-AMSS 19th International Conference on Modelling and Simulation, UKSim 2017*, pp. 155-160.

- Ferro, G., Minciardi, R., Podestà, E., Robba, M. "An optimization model for the sizing of the biomass plants' supply chain" (2018) *IFAC-PapersOnLine*, 51 (5), pp. 114-119.

- Delfino, F., Ferro, G., Minciardi, R., Robba, M., Rossi, M. "Identification and management of an electrical storage system for application in photovoltaic installations" (2017) *IEEE International Conference on Control and Automation, ICCA*, art. no. 8003178, pp. 886-891.

- Ferro, G., Minciardi, R., Robba, M., Rossi, M. "Optimal voltage control and demand response: Integration between Distribution System Operator and microgrids" (2017) *Proceedings of the 2017 IEEE 14th International Conference on Networking, Sensing and Control, ICNSC 2017*, art. no. 8000132, pp. 435-440.

- Ferro, G., Laureri, F., Minciardi, R., Robba, M. "Optimal control of demand response in a smart grid" (2017) 2017 *25th Mediterranean Conference on Control and Automation, MED 2017*, art. no. 7984169, pp. 516-521.

- Delfino, F., Rossi, M., Ferro, G., Minciardi, R., Robba, M. "MPC-based tertiary and secondary optimal control in islanded microgrids" (2015) *1st IEEE International Symposium on Systems Engineering, ISSE 2015* - Proceedings, art. no. 7302507, pp. 23-28.

**Patents**

- "Metodo e sistema per la gestione ottimizzata in tempo reale del trasporto tramite veicoli elettrici" Inventori: FERRO GIULIO, PAOLUCCI MASSIMO, ROBBA MICHELA dep. nr102019000000605 del16/01/2019

- "Metodo per la determinazione rapida dei flussi di potenza all'interno di una generica rete elettrica trifase sbilanciata" Inventori: FERRO GIULIO, ROBBA MICHELA, MANSUETO ROSSI dep. nr102020000005980, 20/03/2020,

# Chapter 2

# The electric system and Balancing Market modelling

In this chapter, the basic concepts of the electric system and the electricity market focused on distribution networks will be introduced. Then, a brief digression will be made to clearly explain what the Balancing Market (BM) is, which are the possible participants, and how it is integrated into the various market sessions. After this section, the mathematical models for each market participant will be presented, highlighting the main characteristics and peculiarities. Like the last section, the test networks that will be be used to test the proposed algorithms are presented. for the market, simulations are presented. In particular, networks derived from a modification of two distribution network benchmarks the IEEE 13 bus and the IEEE 123 bus (i.e. medium voltage networks managed by a distribution system operator).

## 2.1 The Electric System

An electric power system is a network of electrical components deployed to supply, transfer, and use electric power. An example of an electric power system is the grid that provides power to an extended area. An electrical grid power system can be broadly divided into the generators that supply the power, the transmission system that carries the power from the generating centers to the load centers, and the distribution system that feeds the power to nearby homes and industries. Smaller power systems are also found in industry, hospitals, commercial buildings, and homes.

While in the past the electrical grid was mainly characterized by few bulk production plants, in the last ten years many small uncontrolled distributed generators were connected to the distribution grid. These new plants lead to new problems such as voltage fluctuations and peaks, an increase of Joule losses, as described in the introductory section. The electric system can be divided on three levels (see Figure 2.1): transmission, distribution and user grids

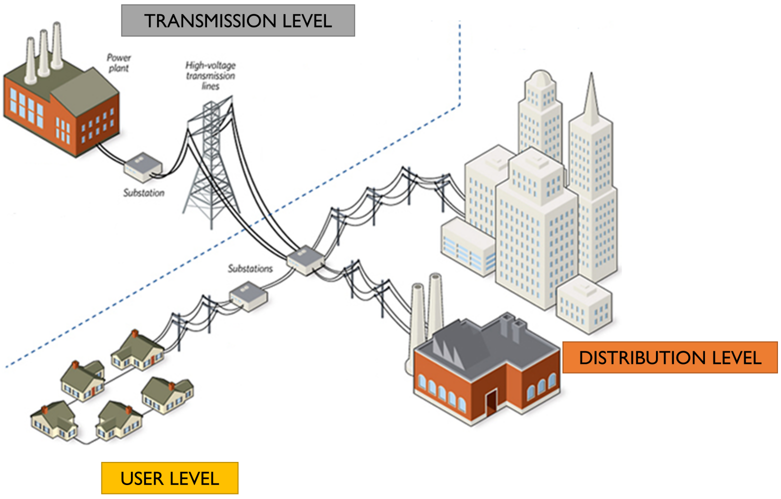


Figure 2.1: Three level power system

At the transmission level, the leading actor is the transmission system operator (TSO) whose goal is to manage the electricity market and take care of the transmission of electricity from large generators to the primary substation, i.e., the connection point with the distribution network.

From the technological point of view, the TSO manages the high voltage (HV) grid that connects the bulk generators to the primary substations that convert power to lower voltages suitable for distribution.

The TSO's role in a wholesale electricity market is to manage the security of the power system in real-time and to determine power flows based on load flow calculations (the optimal combination of generation and reserve for each market trading period) and at prices in each network area, it also handles demand and negotiation concerning power reserves on the network.

The distribution grid refers to the second stage of the electrical grid in which electricity is distributed to homes, industry, and other end-use products. Distribution is the process of reducing power to safe customer-usable levels and delivering electric power to the grid. The entire distribution grid includes lines, transformers, and switching and protection circuits that deliver safe electrical power. The main actor for the distribution grid is the distribution system operator (DSO) that is the owner of the physical infrastructure and manages the overall system in terms of quality of service. In recent years, a large number of distributed generators have been installed in the distribution network, either controllable (e.g. microturbines) or uncontrollable (photovoltaic and wind). This new aspect makes it possible to treat the distribution network as an active network and thus to actively act to improve the overall network service. An example is the introduction of market mechanisms similar to the wholesale energy market, allowing the DSO to dispatch generation and energy reserves through the overall network.

The last level is that of users, which includes all active and passive users served by the distribution network, examples of active users can be microgrids, hybrid systems (photovoltaic with electric storage), or more merely low-power renewable production plants. These users may provide ancillary services to the distribution system to be remunerated; some examples may be frequency adjustment, reactive power regulation, secondary and tertiary reserve, and demand response.

### 2.1.1 The Energy market for distribution grids

One of the central features of the emerging Smart Grid is a highly transformed distribution grid with a high penetration of distributed energy resources (DERs). The electricity grid's hosting capacity is the primary barrier for further increasing distributed generation (DG). The actual infrastructure is designed for mono-directional power flows. As these DERs will be owned and operated by different stakeholders, the efficient and reliable operation of the distribution grid requires an overall market structure that allows the procurement and integration of power generation through DGs, DRs, and storage devices. Despite these recent advances, as the number of DERs grows and renewable generation continues to increase, the wholesale energy market alone may not suffice in realizing efficient and reliable power delivery. Thus, a retail market that oversees the participation of DERs in the distribution grid and implements a suitable mechanism for their scheduling and compensation is highly necessary.

To this end in recent studies, different market mechanisms are proposed to try to optimally manage the distribution network. Many works in literature (better investigated in Chapter 4) define market structures based on iterative processes, whose convergence only is proven by experimental results and not mathematically. Therefore these approaches depend on the case study and the problem structure.

In this thesis, we propose a more general market structure for distribution grid power markets, depicted in Figure 2.2. There are two main market sessions: the Retail Market and the Balancing Market (BM). In the Retail Market session, the dispatch of all the resources is performed mainly based on economic optimization considering the variable prices of the energy acquired from the transmission grid and the forecast of power generation/consumption of the various generators and loads. This market session is performed as a day planning. As previously said, being a purely economic optimization, the various energy prices that the DSO pays to the generators are defined, but no technical considerations are made.

After the first negotiation phase, the economically optimal working point for the system is established. Consequently, balancing energy reserves need to be assessed through the BM session. Indeed, after purely technical evaluation, the DSO can state that the working point obtained during the previous market session is not favorable from the system conditions the point of view (e.g. too high voltages at nodes or lines close to their thermal limit). Differently from the retail market, the BM can be performed both as a day ahead session (after the Retail market negotiation) and as an intra-day market session (e.g. every one hour ) to face with the uncertainties given by the renewable power production and loads forecasting In other words, the BM creates a new working point that is better from the quality of service point of view. This new working point is also advantageous in case of adverse events (e.g. faults on generators/ loads or lines), avoiding undesired conditions during the normal operation.

From a regulatory point of view, each country has its version of the energy market with more or less restrictive constraints due to the various particular political/national situations. The mechanisms proposed in this Ph.D. thesis want to be as generic as possible and therefore be able to accommodate the various particular practical cases.

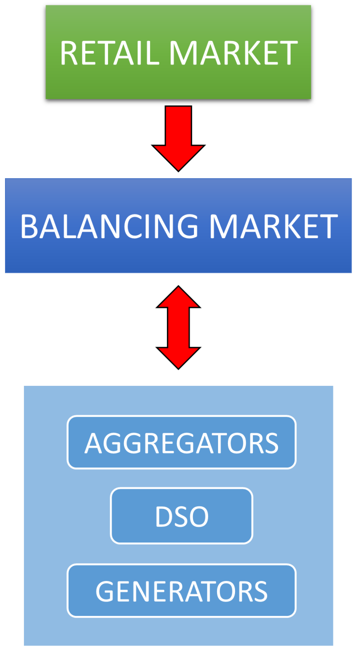


Figure 2.2: The distribution grid market

## 2.2 Balancing market modeling

The topic of electricity BM design has been given relatively little attention from academic researchers, despite its crucial role in both power markets and power system operations. Balance management is a power system operation service vital for ensuring the security of supply through the continuous, real-time balancing of power demand and supply. If the system runs out of balance, power stability and quality will deteriorate, which may trigger the disconnection of system components, and ultimately, power blackouts. The proposed BM is thought to be part of the overall distribution power market and is devoted to obtaining energy reserves for DSOs to balance the system in real-time. Energy reserves can be contracted by the DSO with an associated payment for their availability.

A further important aspect of this Ph.D. thesis is to propose a model also for the figure of the aggregator. This new market player is a purely economic entity, recently introduced from the regulatory point of view, representing a group of active users (generators and loads) that alone could not participate directly in the BM.

Each market participant belongs to the same distribution grid, whose model will be described in the following. As mentioned in the previous section, for this specific market structure, we consider three classes of agents:

- Traditional Generators;
- Aggregators;
- Distribution System Operator.

For generality, no specifications are given regarding the cooperation/competition of the various agents. It is precisely the field of study of this Ph.D. thesis to define new possible market structures, stating the interaction between the various agents.

Any market participants, i.e. the controllable traditional generators and aggregators that want to maximize their profit, taking into account the limits of the flexibility they can sell. Instead, the DSO receives power/price bids from market participants and, decides which bid to accept. In this thesis renewable generators since are not controllable are considered as given inputs. It is important to point out to the reader that the output of the BM optimization problem is the increase/decrease of the amount of active and reactive power of each participant based on technical and economic choices.

In the following, the notation concerning the agents considered for modeling the BM is specified. $N$ is defined as the set of grid nodes. Each traditional generator $g \in GT$, with $GT$ the set of traditional generators, is connected to a single grid node $i = Node(g)$. For any aggregator $a \in A$, with $A$ the set of aggregators, let $C(a)$ the set of customers aggregated by $a$. Let $N_a = card(C(a))$. For any customer $c$ there exists a unique $a$ such that $c \in C(a)$. Moreover let $i = Node(c)$ a map that associates customer $c$ to a grid node $i \in N$. In this way, we can define $Cust(i) = \{c \in C : Node(c) = i\}\ \forall i \in N$.

### 2.2.1 Traditional Generators' Optimization Problem

The introduction of controllable traditional generators within distribution networks has brought a revolution in this paradigm. The possibility of precisely regulating the active and reactive power in a node allows acting in a capillary way minimizing costs and power quality losses of the network. The traditional generator that participates in the BM can offer the flexibility of active and reactive power in both directions since it can increase or decrease both active and reactive power. If the DSO requires flexibility services, the economically optimal working point determined by the retail market is changed to a less efficient one. As the new operating point is economically suboptimal, it implies costs for the generator which requires an economic incentive to the DSO for this service.

We define:

- $P_g^{up}$ and $Q_g^{up}$ the decision variables regarding the increase of active/reactive power for a generator $g$;
- $P_g^{dw}$ and $Q_g^{dw}$ the decision variables regarding the decrease of active/reactive power for a generator $g$;
- $P_g^{da}$ and $Q_g^{da}$ as the working point (in terms of active and reactive power) coming from the Retail market session, considered as input data;
- $C_g^{up}$ and $C_g^{dw}$ as the costs [euro/kW] per unit of active power increased/decreased, considered as input data;
- $\lambda_g^{P,up}(P_g^{up})$ and $\lambda_g^{P,dw}(P_g^{dw})$ as the supply functions representing the prices that generator $g$ requests to the DSO for active power flexibility service.
- $\lambda_g^{Q,up}(Q_g^{up})$ and $\lambda_g^{Q,dw}(Q_g^{dw})$ as the supply functions representing the prices that generator $g$ requests to the DSO for reactive power flexibility service.

The assumption regarding the structure of the supply functions will be clarified in Chapter 4. It is assumed that the costs incurred by each traditional generator are linear both in case of an increase and decrease of produced active power. These costs represent not only the price of fuel of the generator's prime mover but also maintenance cost and efficiency costs given by the change of working point. Instead, as regards the payment for the balancing service we introduce different supply functions (active/reactive power, price) that state the price for the service.

Thus, defining $x_g = \left\{P_g^{up}, P_g^{dw}, Q_g^{up}, Q_g^{dw}\right\}$ as the overall decision variables of a traditional generator $g$, the optimization problem associated to a generator $g \in GT$ is:

$$\max_{x_g}\{\lambda_g^{P,up}(P_g^{up}) - C_g^{up}P_g^{up} - \lambda_g^{P,dw}(P_g^{dw}) + C_g^{dw}P_g^{dw} + \lambda_g^{Q,up}(Q_g^{up}) - \lambda_g^{Q,dw}(Q_g^{dw})\} \tag{2.1}$$

$$0 \leq P_g^{up} \leq \overline{P_g^{up}} \tag{2.2}$$

$$-\underline{P_g^{dw}} \leq P_g^{dw} \leq 0 \tag{2.3}$$

$$Q_g^{dw} \leq 0 \tag{2.4}$$

$$-Q_g^{up} \leq 0 \tag{2.5}$$

$$(P_g^{up} + P_g^{da} + P_g^{dw})^2 + (Q_g^{up} + Q_g^{da} + Q_g^{dw})^2 \leq S_g^2 \tag{2.6}$$

where $\overline{P_g^{up}}$ and $\underline{P_g^{dw}}$ are bounds for flexibility (in both directions), note that active sign convention is used (positive power is generated). Constraints (2.2)-(2.5) define upper and lower bounds over the decision variables, while (2.6) represents the capability constraint (with $S_g^2$ maximum apparent power), that defines the maximum amount of apparent power exchangeable by each traditional generator, considering the BM requirements. In the objective function reactive power has no cost since it is only related to the capability of the generator.

### 2.2.2 The Aggregator's Optimization Problem

The user aggregator is an economic entity that participates in the BM. It represents in the BM a certain energy reserve that is provided by several users but belonging to the same aggregation perimeter. This agent enters into contracts with its customers and provides a remuneration if the balancing service (increasing/decreasing of active and reactive power) is provided. Once the outcomes of the BM are available, the information is shared among the customers. In this case, two possible scenarios can rise: customers reach an agreement without involving the aggregator (while keeping privacy) following constraints on the total resource to be supplied; or the energy is directly requested by the aggregator, which based on the available information distributes the balancing service between the customers. Nevertheless, this topic does not matter of this thesis but it has been studied in other papers of the same authors of this thesis [1].

The working structure of an aggregator is depicted in Figure 2.3. The blue box, representing the BM session, there is the request of balancing service as a result of the bid among the market participants (the bidirectional arrow represents the interaction of the market and the aggregator in terms of bids and service assignment). The second box represents the aggregator that after the market results communicate to its customers the amount of power to be reduced/increased. In the third part of the figure, there is the local optimal management of the customers.

We define for each aggregator $a$:

- $P_c^{up}$ and $Q_c^{up}$ the decision variables regarding the increase of active/reactive power for a customer $c$;
- $P_c^{dw}$ and $Q_c^{dw}$ the decision variables regarding the decrease of active/reactive power for a generator $c$;
- $P_c^{da}$ and $Q_c^{da}$ as the working point (in terms of active and reactive power) coming from the retail marked session for each customer $c$, considered as input data;

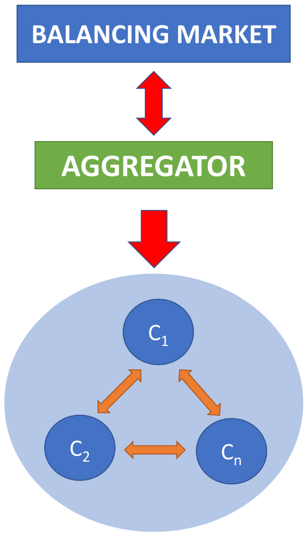


Figure 2.3: Aggregator's interaction with the BM and its customers

- $C_c^{up}$ and $C_c^{dw}$ as the costs [euro/kW] per unit of active power increased/decreased for each customer $c$, considered as input data;
- $\lambda_a^{P,up}(P_c^{up})$ and $\lambda_a^{P,dw}(P_c^{dw})$ as the supply functions representing the prices that aggregator $a$ request to the DSO for active power flexibility service (referring to all customers in $C(a)$).
- $\lambda_g^{Q,up}(Q_c^{up})$ and $\lambda_g^{Q,dw}(Q_c^{dw})$ as the supply functions representing the prices that generator $a$ request to the DSO for reactive power flexibility service (referring to all customers in $C(a)$).

The assumption regarding the structure of the supply functions will be clarified in Chapter 4. It is important to note that the aggregator offers a single bid (i.e. a single supply function) taking into account each customer in an aggregate manner. As said before, this aspect is not the subject of the thesis but has been investigated by the authors of the thesis in other papers. Since both small generators and loads can belong to the same aggregator, these costs for load type customers can represent the dissatisfaction cost of the load's users. Thus, defining $x_c = \left\{P_c^{up}, P_c^{dw}, Q_c^{up}, Q_c^{dw}\right\}$ as the overall decision variables of a customer $c$ of an aggregator $a$, the optimization problem associated to aggregator $a$ is given by:

$$\max_{x_c\ c\in C(a)} \left\{ \sum_{c\in C(a)} \lambda_a^{P,up}(P_c^{up}) - \sum_{c\in C(a)} C_c^{up} P_c^{up} - \sum_{c\in C(a)} \lambda_a^{P,dw}(P_c^{dw}) + \sum_{c\in C(a)} C_c^{dw} P_c^{dw} \right.$$

$$\left. + \sum_{c\in C(a)} \lambda_a^{Q,up}(Q_c^{up}) - \sum_{c\in C(a)} \lambda_a^{Q,dw}(Q_c^{dw}) \right\} \tag{2.7}$$

$$0 \le P_c^{up} \le \overline{P_c^{up}} \quad \forall c \in C(a) \tag{2.8}$$

$$-\underline{P_c^{dw}} \le P_c^{dw} \le 0 \quad \forall c \in C(a) \tag{2.9}$$

$$Q_c^{dw} \le 0 \quad \forall c \in C(a) \tag{2.10}$$

$$-Q_c^{up} \le 0 \quad \forall c \in C(a) \tag{2.11}$$

$$\left(P_c^{up} + P_c^{da} + P_c^{dw}\right)^2 + \left(Q_c^{up} + Q_c^{da} + Q_c^{dw}\right)^2 \le S_c^2 \quad \forall c \in C(a) \tag{2.12}$$

where $\underline{P_c^{dw}}$ and $\overline{P_c^{up}}$ are the minimum and maximum power for the flexibility of customer $c$. Constraints (2.8)-(2.11) define upper and lower bounds over the decision variables, while (2.12) represents the capability constraint (with $S_c^2$ maximum apparent power), that defines the maximum amount of apparent power exchangeable by each customer $c$, considering the BM requirements; note that active sign convention is used. Also for this decision-maker in the objective function reactive power has no cost since it is only related to the capability of each customer. it is important to point out that although the problems of generator and aggregator optimization seem similar, they have a very different practical meaning. The main differences are two: the first is the capillarity of regulation that has the aggregator, in fact, representing more nodes of the network manages to make many offers are accepted by the DSO since it can act directly in more areas of the network; the second instead concerns the price function that is unique for all components of the aggregator, then from an economic point of view the evaluation of the function is more complex (taking into account the characteristics of individual users) than that of generators

### 2.2.3 The DSO Optimization Problem

The DSO is the owner of the distribution network from the primary substation to the end-users, its objective is to optimize the operational management of the network. In a more recent perspective, with a strong presence of intermittent renewable sources on the grid, in addition to the classic problems (minimization of losses and flow of lines on the grid) the DSO has to face with new challenges such as the reversal of power flow and voltage regulation by active users. Moreover, it has to consider power quality indexes of the overall network.

The DSO represents the arbiter of the BM and has complete knowledge of the distribution network and can thus technically evaluate the level of stress of the network components (generators, nodes, lines).

As mentioned at the beginning of this chapter, the DSO can evaluate the quality of the actual working point (through specific indicators) to require flexibility. To this end, we introduce specific Stress Indexes ($SI$) defined as follows:

$$SI_h^G = \frac{P_g^{da} + P_g^{up} + P_g^{dw}}{P_g^{\max}} \quad \forall h \in (GT \cup C_{TOT}) \tag{2.13}$$

$$SI_i^V = \left( \frac{\|V_{ik}\| - \hat{V}_{ik}}{\Delta V_i^*} \right)^2 \forall i \in N \; k \in K \tag{2.14}$$

$$SI_{ij}^C = \left( \frac{I_{ij,k}}{I_{ij}^*} \right)^2 \forall i,j \in Nk \in K \tag{2.15}$$

The terms in (2.13) represents the ratio between the actually generated power and the rated active power of the generator $P_g^{\max}$. The second term (2.14) represents the square of the ratio between the variation of the voltage at node $i$ (with respect to the nominal value specified by $\hat{V}_{ik}$) and a value that expresses a measure of the interval of admissible variations. The third kind of terms (2.15) represents the square of the ratio between the current in branch $(i, j)$ and the rated thermal limit. These simple performance indicators allow evaluating the stress state of the whole system. It is easy to understand that the lower these indices the more easily the system can withstand deep external perturbations, such as possible failures, sudden changes in power production. For example, a sudden voltage peak in a node dramatically increases joule losses with the risk of exceeding the thermal limits of the cables or damaging the generators. Thus, the objective function of the DSO problem is the weighted sum of economic costs, given by the purchased flexibility from the market participants, and the overall $SI$ of the network. That is,

$$\min_{x_g \in GT, x_c c \in C(a), a \in A} \sum_{g \in GT} \left\{ \lambda_g^{P,up}(P_g^{up}) - \lambda_g^{P,dw}(P_g^{dw}) + \lambda_g^{Q,up}(Q_g^{up}) - \lambda_g^{Q,dw}(Q_g^{dw}) \right\}$$
$$+ \sum_{a \in A} \left\{ \sum_{c \in C(a)} \lambda_a^{P,up}(P_c^{up}) - \sum_{c \in C(a)} \lambda_a^{P,dw}(P_c^{dw}) + \sum_{c \in C(a)} \lambda_a^{Q,up}(Q_c^{up}) - \sum_{c \in C(a)} \lambda_a^{Q,dw}(Q_c^{dw}) \right\}$$
$$+ \alpha SI_{sys} \tag{2.16}$$

where

$$SI_{sys} = \sum_{h \in (GT \cup C_{tot})} SI_h^G + \sum_{i \in N} SI_i^V + \sum_{i \in N} \sum_{\substack{j \in N \\ i \neq j}} SI_{ij}^C \tag{2.17}$$

The distribution network is modeled as an unoriented graph $G(N, E)$, with $N$ set of nodes and $E$ set of branches. Since we are facing real distribution networks we need to consider a multiphase unbalanced system. Each variable now is a vector with 3 components (one for each phase $k \in K$ with $K$ the set of phases). The network topology is represented by the admittance matrix $\boldsymbol{Y}$ (more clarifications about the $\boldsymbol{Y}$ matrix construction is given in the Appendix).

The edges are expressed by the 3-phase matrix admittance $y_{i,j} \in \mathbb{R}^{3\times3}$ (representing the cables of a link between two nodes), for a cable between nodes $i$ and $j \in N$, in which each element of the matrix, $y^{k,k'}$, represents the magnetic coupling between phases $k$ and $k'$, such that $y_{k,k'} = g_{k,k'} + ib_{k,k'} \forall k, k' \in K$ where $K = a, b, c$ is the set of phases. That is:

$$y_{ij} = \begin{bmatrix} y_{ij}^{aa} & y_{ij}^{ab} & y_{ij}^{ac} \\ y_{ij}^{ab} & y_{ij}^{bb} & y_{ij}^{bc} \\ y_{ij}^{ac} & y_{ij}^{bc} & y_{ij}^{cc} \end{bmatrix} \quad \forall i, j \in N \tag{2.18}$$

The power flow equations which describe the sinusoidal steady-state equilibrium of a power network are given by:

$$P_i = \sum_{j\in N} V_i V_j (G_{ij} \cos(\delta_i - \delta_j) + B_{ij} \sin(\delta_i - \delta_j)) \quad \forall i \in N \tag{2.19}$$

$$Q_i = \sum_{j\in N} V_i V_j (G_{ij} \sin(\delta_i - \delta_j) - B_{ij} \cos(\delta_i - \delta_j) \quad \forall i \in N \tag{2.20}$$

where $G_{i,j}$ and $B_{i,j}$ are conductance and susceptance parameters for the branch $(i, j)$, $V_i$ and $\delta_i$ are the voltage magnitude and phase at node $i$, respectively. $P_i$ and $Q_i$ represents the power injected at each node that are given by:

$$P_i = \sum_{\substack{g\in GT: \\ i=Node(g)}} \left(P_g^{da} + P_g^{up} + P_g^{dw}\right) + \sum_{a\in A} \sum_{\substack{c\in C(a): \\ i=Node(c)}} \left(P_c^{da} + P_c^{up} + P_c^{dw}\right) \quad \forall i \in N \tag{2.21}$$

$$Q_i = \sum_{\substack{g\in GT: \\ i=Node(g)}} \left(Q_g^{da} + Q_g^{up} + Q_g^{dw}\right) + \sum_{a\in A} \sum_{\substack{c\in C(a): \\ i=Node(c)}} \left(Q_c^{da} + Q_c^{up} + Q_c^{dw}\right) \quad \forall i \in N \tag{2.22}$$

Constraints (2.21) and (2.21) define the power balance at each node that is the sum of the contributions of generators and load at each node considering the increase/decrease of active and reactive power given by the BM session. As can be seen, the above equations (2.19) and (2.20) are nonlinear and nonconvex. The sources of non-convexity are given by the bilinear terms $V_i V_j$ and the trigonometric functions. Nonconvexites are a challenging problem because they increase the computational burden of an optimization problem. Finding a high-quality convex relaxation of the power flow equations represents a good result in terms of reducing the run time of the overall optimization problem. The next chapter will be devoted to the statement of a new convex relaxation for the optimal power flow (OPF) problem in distribution grids.

## 2.3 Distribution test networks

Models and methods developed in this Ph.D. thesis will be tested on test networks based on benchmark IEEE networks and modified for active distribution grids. In this section two of the used test will be described, namely the IEEE 13 bus and the IEEE 123 bus ([2]), posing particular attention to the changes adopted for the study of the BM in the distribution network. Each quantity is expressed in per-unit values [p.u.]. This method, commonly used in power systems studies, offers computational simplicity by eliminating units and expressing system quantities as dimensionless ratios. The bases chosen for this example are 1 [MVA] for the apparent power and 15[kV] for the voltage. All the parameters are reported in the next chapters. The changes were made to highlight the peculiarities of the developed models. In most of the literature, the distribution networks examined are very simple and above all have a radial and balanced typology, which is not true in real systems.

### 2.3.1 Modified IEEE 13 bus network

For this test case, the BM participants are one traditional generator (belonging to bus 646) and two aggregators, while node 650 represents the slack bus. The test network has been adapted to our case employing the following modifications depicted in Figure 2.4:

- at bus 646 there is a traditional generator;
- at bus 684 a small generator belonging to aggregator 1 and a local load are connected;
- bus 634 is composed of a small generator belonging to aggregator 2 and a local load;
- to bus 692 are linked two customers of aggregator 1 and 2;
- two new links are added between nodes 652-680 and 680-692;
- all the switches are in their normal position.

Each quantity is expressed in per unit values [p.u.]. The bases chosen for this example are the same as the original benchmark network. In Table C.1 the apparent powers at each node and phase are reported (as in the benchmark data passive sign convention is used).

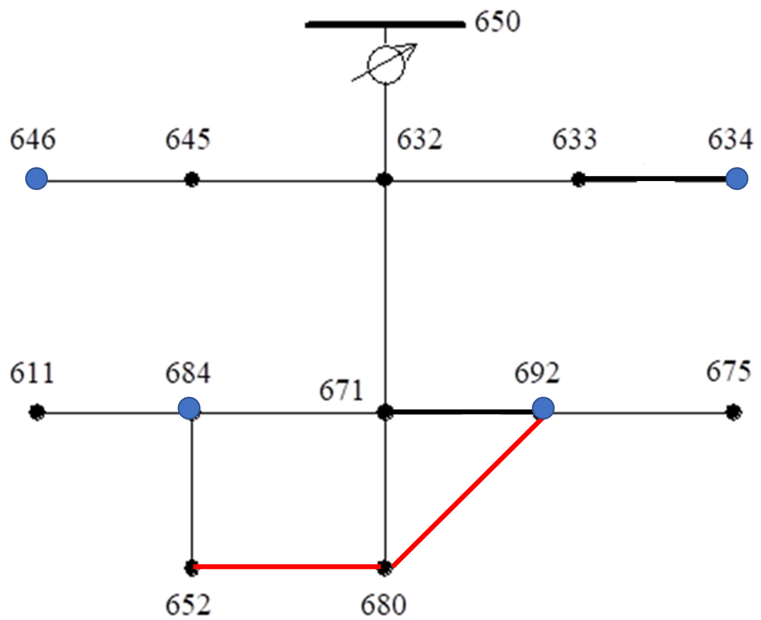


Figure 2.4: The Modified IEEE 13 test network.

### 2.3.2 Modified IEEE 123 bus network

For this second test case, the test network has been adapted to our case using the following modifications depicted in Figure 2.5:

- at bus 41 and 114 there are traditional generators;
- nodes 42-46 belong to the first aggregator;
- nodes 3-6 belong to the second aggregator;
- nodes 31-33 belong to the third aggregator;
- five new links are added between nodes 19-21, 63-66, 90-92 and 90-88;
- all the switches are in their normal position.

Each quantity is expressed in per-unit values [p.u.]. The bases chosen for this example are the same as the original benchmark network. In Table C.2 (Appendix C) the apparent powers at each node and phase are reported (as in the benchmark data passive sign convention is used).

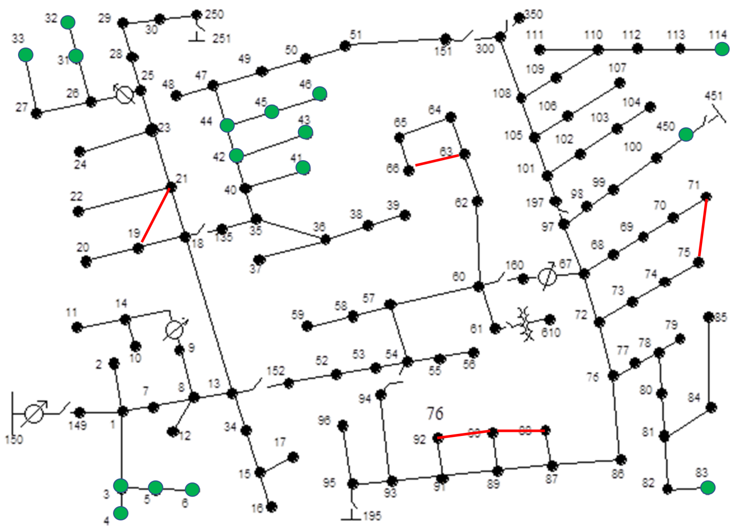

Figure 2.5: The Modified IEEE 123 test network.

# Chapter 3

# A New Convex Relaxation of the Optimal Power Flow problem for Distribution Grids with Unbalanced Structure and Mesh Topology

This chapter aims to address the problem of the non-convexity of power flow equations highlighted in the previous chapter. In fact, before formulating any market structure (competitive or cooperative), a modelling effort is needed to make the power grid model more mathematically simple. The activities presented in this chapter have been partially developed during the research stage at the Active Adaptive Control (AAC) Lab at Massachusetts Institute of Technology (MIT) under the supervison of Dr. Anuradha Annaswamy, and are also described in [3]. The main contribution of this Chapter can be summarized as:

- The statement of a new convex relaxation of the OPF problem based on McCormick envelopes. This new model can afford any type of distribution network with unbalanced structure and mesh topology, without binding working hypothesis.
- The development of an explicit procedure that states optimal upper and lower bounds for injected current increasing the quality of the proposed relaxation.

In order to validate the proposed model an extensive test campaign has been preformed on the most common benchmarks for distribution grids, i.e. the IEEE 13 and 123 bus networks and a more recent test case, the IEEE 8500 nodes. Each of these test cases have been modified (adding distributed generation and new buses to obtain loops) to exploit the peculiarities of the new model, that is the maximum flexibility in terms of unbalances and topology. The test campaign aims at studying the potential capabilities of the model. The first wants to study the possibility to use the proposed relaxation as a load flow tool: the apparent power at each node (except to the slack node) is fixed, all the voltages phasors are thus determined and constraints verified. The second possible purpose of the proposed approach is to be considered as OPF tool considering a voltage minimization objective function.

The organization of this chapter is as follows. Section I investrigates the actual literature of OPF convexification, in Section II the OPF problem is described and the new convex relaxation is defined, Section III describes the preprocessing algorithm designed to obtain optimal upper and lower current bounds for the optimization problem. In Section IV the proposed model has been tested in the modified IEEE 13 and 123 bus and 8500 nodes test networks, in terms of feasibility (as load flow tool) and optimality (as OPF tool).

## 3.1 Literature Review

As reported in Chapter 2, the new paradigm of the distribution network necessitates to develop methods and models for optimizing power injections into the grid. The problem related to power system optimization is called Optimal Power Flow (OPF) [4], which is intrinsically hard to solve since is a nonlinear and non convex problem. In the recent literature many efforts are made to find high quality convex relaxations for the OPF [5]. The most popular models, are the bus injection model (BIM) based on semidefinite programming (SDP) and the branch flow model (BF) based on second order cone programming (SOCP) [5][6],

The main purpose of this chapter is to define a new formulation for the OPF that can be used for any general distribution grid (in terms of network topology and unbalanced design). The motivation to create this model rises from the fact that in the recent literature the most popular models guarantee an optimal performance only in specific cases such as radial topology and balanced structure [7] [8]. All of these hypotheses, from a practical point of view, are never met since the peculiarity of a distribution grid is the unbalanced structure (i.e., the power flow is different for each phase [9]) and mesh topology.

The idea proposed in the chapter consists in transforming all loads and generations into current injection (CI) in order to solve the network as a circuit using Kirchhoff's and Ohm's laws. This concept comes from some methods used in load flow calculation [10] [11]. These methods use the concept of CI to be able to iteratively solve load flow equations using the voltage value obtained from the previous iteration. In this way, the problem is linear in the currents and therefore can be easily solved. In the present chapter, unlike the above mentioned approaches, we want to define an optimization problem. Thus, it is not convenient to use an approach based on an iterative solution. We thus propose to convexify the CI formulation of the OPF through the McCormick envelopes. This technique, firstly introduced in [12], is mainly used to convexify bilinear non linear programming problems. The proposed CI formulation is perfectly suitable to this method, since, adopting the cartesian coordinates for the phasor modelling the only nonlinearities are given by four bilinear terms. In the recent literature other works have applied McCormick enevlopes to the OPF [13] [14] and [15], but in a different way and context from the one presented in this chapter.

The authors of [13] propose the application of the McCormick envelopes to the SDP relaxation of the power flow equations. Differently from our work this application is only for transmission grids, and they propose to relax the trigonometric functions that has a limited validity, only in a certain range of angles. In [14] the authors apply the so called "Meyer and Floudas" envelopes to trilinear monomials formed by the product of the voltage magnitudes and trignometric terms in the polar form of the power flow equations. Finally, in [15] the authors consider a novel general power system modeling approach based on multi-port representation of individual components with Sparse Tableau Formulation of network constraints based on orthogonal components. Differently from the approach proposed in this chapter [14] and [15] propose an application only for transmission grids with balanced structure.

It is important to point out to the reader that, although many works in the literature are more mathematically sounding than the CI model (such as BF model and BI model), they have many limitations from a practical point of view [16] [17]. These restrictions are related to network topology, balancing structure and large presence of distributed generation. As far as network topology is concerned, these models are limited by the fact that they can only solve networks with radial topology. In this case, in order to have a SOCP or SDP relaxation [18][19] it is necessary to consider the square of voltages and currents, completely losing the information about the phases. Instead, with the CI model it is possible to solve networks with any type of topology as the information on the phase is maintained thanks to the presence of the real and imaginary parts of the phasorial variables. From the point of view of phase balancing, most of the works using BF as a model for the distribution network assume that the network is perfectly balanced. This is an extremely strong hypothesis from a practical point of view, since the main peculiarity of a distribution network is the imbalance at the level of nodal voltages, cables and loads/generation. Some recent chapters have tried to model a distribution network modifying the BF adapting it to an unbalanced network as in [20] and [21]. Also in these cases strong hypotheses are made on the reciprocal unbalance of the phases that are valid only in a small range of cases, namely neglecting the relative difference between the phasors of each phase. Other chapters [22] suggest the use of Fortescue's transformation to deal with multiphase systems in the "sequences" context instead of in "phases" ones. This approach is useful in transmission networks but it is often not valid in distribution ones, as the cables are not inverted allowing a decoupling in the admittance matrix. Finally, as a last limitation of the BF, we must mention the inaccuracy of the relaxation given by the large presence of distributed generation at the nodes of the network. A precise description of this restriction is reported in [23], where it is explicitly explained that a condition for which the relaxation brought by the BF is exact only if there is a limited number of DERs.

## 3.2 Formalization of the OPF problem

Electric power transmission needs optimization in order to determine the necessary real and reactive power flows in a system for a given set of loads, as well as the resulting voltages and currents in the system. Power flow studies are used both for analyzing the actual state of the grid and for planning in the day ahead how to face possible disturbances into the system, such as the loss of a line to maintenance and repairs.

The OPF should be performed for a large number of nodes, by using nonlinear equations for the electrical grid representation. Generally, to manage such a complex systems the problem is formalized by the use of simulation and optimization tools, and of suitable system representations. The use of the admittance matrix is particularly convenient because it represents at the same time the network topology and the cable characteristics in order to determine voltages, currents and power flows through the system. In the following, the proposed OPF formulation is described. Note that the real and imaginary part of a phasor $\dot{x}$ are denoted as $x^R$ and $x^I$ while $x^H$ is the Hermitian of vector $x$, finally we use the the overbar $\overline{x}$ and underbar $\underline{x}$ notation to denote upper and lower limits for a variable $x$.

### 3.2.1 General CI-OPF formulation

The distribution network is modelled as an unoriented graph $G(N, E)$, with $N$ set of nodes of the grid and $E$ set of branches. Each variable now is a vector with 3 components (one for each phase $k \in K$ with $K$ the set of phases). The cables are expressed by the 3-phase $3\times3$ matrix admittance in which each component represents the magnetic coupling that each phase (i.e. $k = a, b, c$) has with the other phases (i.e. $k = a, b, c$). That is:

$$\dot{y}_{ij} = \begin{bmatrix} \dot{y}_{ij}^{aa} & \dot{y}_{ij}^{ab} & \dot{y}_{ij}^{ac} \\ \dot{y}_{ij}^{ab} & \dot{y}_{ij}^{bb} & \dot{y}_{ij}^{bc} \\ \dot{y}_{ij}^{ac} & \dot{y}_{ij}^{bc} & \dot{y}_{ij}^{cc} \end{bmatrix} \quad \forall i, j \in N$$

The whole admittance matrix $\boldsymbol{Y}$ is defined as follows:

$$\boldsymbol{Y}_{ij} = \begin{cases} \sum\limits_{m:(i,m)\in E_i} \dot{y}_{im} & \text{if } i = j \\ -\dot{y}_{ij} & \text{if } i \neq j \text{ and } (i,j) \in E \\ 0 & \text{otherwise} \end{cases}$$

The OPF problem aims at minimizing a performance index (related for example to losses, voltage regulation, energy cost etc.), subject to constraints that describe the physics of the power system. In this particular formulation, we keep explicitly the current injections definition at each phase of each node to obtain the CI-OPF formulation. A general formulation of a 3-phase CI-OPF is given by:

$$\min_{(\dot{I},\dot{V},\dot{S})} f\left(\dot{I},\dot{V},\dot{S}\right) \tag{3.1}$$

$$\dot{I} = \mathbf{Y}\dot{V} \tag{3.2}$$

$$\dot{S} = \dot{V}\dot{I}^{H} \tag{3.3}$$

$$\underline{\dot{V}_i} \leq \dot{V} \leq \overline{\dot{V}} \tag{3.4}$$

$$\underline{\dot{S}} \leq \dot{S} \leq \overline{\dot{S}} \tag{3.5}$$

where $\dot{I}$, $\dot{V}$ and $\dot{S}$ are the vectors related to currents, voltages and apparent powers in the grid. In this CI-OPF formulation, (3.1) describes general objective function typical of power systems minimizing a function of current voltages or power, (3.2) is the nodal analysis relation for an AC system, while equation (3.3) is the apparent power definition. Finally equations (3.4)-(3.5) are upper and lower bounds for voltage and apparent power injected at each bus respectively.

In the previous formulation (3.1)-(3.5), notation was abused because all the quantities described are phasors, i.e. complex numbers. There are two notations to describe a complex number: polar coordinates, defining the magnitude and phase $\dot{x} = |x|e^{j\delta}$, and Cartesian coordinates, defining real and imaginary component $\dot{x} = x^R + jx^I$. For this purpose, the Cartesian coordinate representation will be chosen. This representation allows to rewrite the problem (3.1)-(3.5) in a more manageable formulation.

$$\min_{I^R,I^I,V^R,V^I,P,Q} f\left(I^R, I^I, V^R, V^I, P, Q\right) \tag{3.6}$$

$$I^R = \mathrm{Re}\left(\mathbf{Y}V\right) \tag{3.7}$$

$$I^I = \mathrm{Im}\left(\mathbf{Y}V\right) \tag{3.8}$$

$$\underline{V_{ik}}^R \leq V_{ik}^R \leq \overline{V_{ik}}^R \;\; \forall i \in N \;\, k \in K \tag{3.9}$$

$$\underline{V_{ik}}^I \leq V_{ik}^I \leq \overline{V_{ik}}^I \;\; \forall i \in N \;\, k \in K \tag{3.10}$$

$$\underline{P}_{ik} \leq P_{ik} \leq \overline{P}_{ik} \;\; \forall i \in N \;\, k \in K \tag{3.11}$$

$$\underline{Q}_{ik} \leq Q_{ik} \leq \overline{Q}_{ik} \;\; \forall i \in N, \;\, k \in K \tag{3.12}$$

$$P_{ik} = V_{ik}^R I_{ik}^R + V_{ik}^I I_{ik}^I \;\; \forall i \in N \;\, k \in K \tag{3.13}$$

$$Q_{ik} = -V_{ik}^R I_{ik}^I + V_{ik}^I I_{ik}^R \; \forall i \in N \;\, k \in K \tag{3.14}$$

where ${I_{ik}}^R$,${I_{ik}}^I$, ${V_{ik}}^R$,${V_{ik}}^I$, $P_{ik}$ and $Q_{ik}$ are voltages, nodal currents, real and reactive power injected at phase $k$ of node $i$ (active sign convention is used). It is important to note that any objective function in the field of power systems is generally convex. From a practical point of view this assumption is very common and acceptable, because most of the faced problems (volt/VAR regulation, losses minimization, frequency regulation, energy cost minimization, etc.)

are quadratic or linear functions of the variables. As can be seen from (3.6)-(3.12) the sources of non-convexity are given by the bilinear terms $V_i^R I_{ik}^R$, $V_i^I I_{ik}^I$, $V_i^R I_{ik}^I$ and $V_i^I I_{ik}^R$. In the next section a methodology to achieve convex relaxation of a bilinear form will be introduced.

### 3.2.2 Relaxation of the CI-OPF problem

Non convex problems are generally transformed into a convex function by relaxing some parameters. Note that a convex relaxation reduces the computational burden of the overall problem at the cost of obtaining a sub optimal solution. Specifically, the convex problem will provide a lower bound for the optimal solution. Having a tighter relaxation that is still convex will provide a lower bound that is closer to the solution.

McCormick Envelopes [12] are a type of convex relaxation used in bilinear non linear programming problems. McCormick Envelopes provide an envelope that retains convexity while minimizing the size of the new feasible region. This allows the lower bound solution obtained from using these envelopes to be closer to the true solution than if other convex relaxations were used. Tighter envelopes decreases the time needed to solve complex computational problems.

We define the enevlope ($MCE\{w\}$) of a bilinear form $w = xy$ defined over a set $\boldsymbol{S} \subset \mathbb{R}^3 = \{x, y : x \in [\underline{x}, \overline{x}], y \in [\underline{y}, \overline{y}]\}$ as (for a more extensive explanation please see Appendix A):

$$w \geq \underline{x}y + x\underline{y} - \underline{x}\underline{y} \tag{3.15}$$

$$w \geq \overline{x}y + x\overline{y} - \overline{x}\overline{y} \tag{3.16}$$

$$w \leq \underline{x}y + x\overline{y} - \overline{x}\underline{y} \tag{3.17}$$

$$w \leq \overline{x}y + x\underline{y} - \underline{x}\overline{y} \tag{3.18}$$

Through the envelopes described above it is possible to relax the bilinear terms in equations

(3.11) and (3.12) to obtain the following set of constraints:

$$P_{ik} = a_{ik} + b_{ik} \;\; \forall i \in N,\; k \in K \tag{3.19}$$

$$Q_{ik} = -c_{ik} + d_{ik} \;\; \forall i \in N,\; k \in K \tag{3.20}$$

$$a_{ik} \geq \underline{V_i}^R I_{ik}^R + V_i^R \underline{I}_{ik}^R - \underline{V_i}^R \underline{I}_{ik}^R \;\; \forall i \in N,\; k \in K \tag{3.21}$$

$$a_{ik} \geq \overline{V_i}^R I_{ik}^R + V_i^R \overline{I}_{ik}^R - \overline{V_i}^R \overline{I}_{ik}^R \;\; \forall i \in N,\; k \in K \tag{3.22}$$

$$a_{ik} \leq \underline{V_i}^R I_{ik}^R + V_i^R \overline{I}_{ik}^R - \overline{V_i}^R \underline{I}_{ik}^R \;\; \forall i \in N,\; k \in K \tag{3.23}$$

$$a_{ik} \leq \overline{V_i}^R I_{ik}^R + V_i^R \underline{I}_{ik}^R - \underline{V_i}^R \overline{I}_{ik}^R \;\; \forall i \in N,\; k \in K \tag{3.24}$$

$$b_{ik} \geq \underline{V_i}^I I_{ik}^I + V_i^I \underline{I}_i^R - \underline{V_i}^I \underline{I}_{ik}^R \;\; \forall i \in N,\; k \in K \tag{3.25}$$

$$b_{ik} \geq \overline{V_i}^I I_{ik}^I + V_i^I \overline{I}_i^I - \overline{V_i}^I \overline{I}_i^I \;\; \forall i \in N,\; k \in K \tag{3.26}$$

$$b_{ik} \leq \underline{V_i}^I I_{ik}^I + V_i^I \overline{I}_i^I - \overline{V_i}^I \underline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.27}$$

$$b_{ik} \leq \overline{V_i}^I I_{ik}^I + V_i^I \underline{I}_i^R - \underline{V_i}^I \overline{I}_i^I \;\; \forall i \in N,\; k \in K \tag{3.28}$$

$$c_{ik} \geq \underline{V_i}^R I_{ik}^I + V_i^R \underline{I}_i^R - \underline{V_i}^R \underline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.29}$$

$$c_{ik} \geq \overline{V_i}^R I_{ik}^I + V_i^R \overline{I}_i^I - \overline{V_i}^R \overline{I}_i^I \;\; \forall i \in N,\; k \in K \tag{3.30}$$

$$c_{ik} \leq \underline{V_i}^R I_{ik}^I + V_i^R \overline{I}_i^I - \overline{V_i}^R \underline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.31}$$

$$c_{ik} \leq \overline{V_i}^R I_{ik}^I + V_i^R \underline{I}_i^R - \underline{V_i}^R \overline{I}_i^I \;\; \forall i \in N,\; k \in K \tag{3.32}$$

$$d_{ik} \geq \underline{V_i}^I I_{ik}^R + V_i^I \underline{I}_i^R - \underline{V_i}^I \underline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.33}$$

$$d_{ik} \geq \overline{V_i}^I I_{ik}^R + V_i^I \overline{I}_i^R - \overline{V_i}^I \overline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.34}$$

$$d_{ik} \leq \underline{V_i}^I I_{ik}^R + V_i^I \overline{I}_i^R - \overline{V_i}^I \underline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.35}$$

$$d_{ik} \leq \overline{V_i}^I I_{ik}^R + V_i^I \underline{I}_i^R - \underline{V_i}^I \overline{I}_i^R \;\; \forall i \in N,\; k \in K \tag{3.36}$$

It is noteworthy that in the previous statement of the problem (3.6)-(3.14), no limit was defined on the current injected to node $i$. In fact, it is not necessary to have a limit for these variables, as it is possible collect (3.2) and (3.3) to define the AC power as a function of voltage. Defining accurate minimum and maximum limits for current injections is not trivial, as the next section will be dedicated to the statement of an optimization-based approach to define upper and lower bounds for currents.Thus, the overall optimization problem is given by the minimization of (3.6) subject to (3.9)-(3.12) and (3.19)-(3.36).

## 3.3 Current bounds definition

The convex relaxation of the CI-OPF problem through the McCormick envelopes adds additional constraints to the problem, concerning the minimum and maximum limits of injected current at each phase of each node. The purpose of this section is to define a methodology that allows to define correct upper and lower bound for currents, given the limits on voltages and powers.

In this chapter we propose a heuristic method to determine the optimal upper and lower bounds for real and imaginary currents based on the solution of a simple nonlinear optimization problem. For brevity we will present the optimization problem to obtain $\overline{I}_{ik}^{R}$ $\forall i \in N,\ k \in K$. Same considerations has to be done for $\underline{I}_{ik}^{R}$, $\overline{I}_{ik}^{I}$ and $\underline{I}_{ik}^{I}$ $\forall i \in N,\ k \in K$.

Specifically $\overline{I}_{ik}^{R}$ corresponds to the solution of the following optimization problem:

$$\max_{I_{ik}^R} I_{ik}^R \tag{3.37}$$

$$P_{ik} = V_i^R I_{ik}^R + V_i^I I_{ik}^I \tag{3.38}$$

$$Q_{ik} = -V_i^R I_{ik}^I + V_i^I I_{ik}^R \tag{3.39}$$

$$\underline{V}_i^R \leq V_i^R \leq \overline{V}_i^R \tag{3.40}$$

$$\underline{V}_i^I \leq V_i^I \leq \overline{V}_i^I \tag{3.41}$$

$$\underline{P}_{ik} \leq P_{ik} \leq \overline{P}_{ik} \tag{3.42}$$

$$\underline{Q}_{ik} \leq Q_{ik} \leq \overline{Q}_{ik} \tag{3.43}$$

The optimization problem (3.37)-(3.43) can be solved analytically. In the following lemma it is shown the analytical solution of the above problem (for the sake of simplicity we consider a constant load or generator, the same considerations ca be done for variable generator/load).

**Lemma 3.3.1.** *The optimal solution of problem (3.37)-(3.43) corresponds to the maximum value of the points obtained by the combination of the following functions calculated in* $\underline{V}_i^I, \overline{V}_i^I, \underline{V}_i^R, \overline{V}_i^R$*:*

$$\frac{Q_{ik}^2\sqrt{P_{ik}^2+Q_{ik}^2}}{V_{ik}^R\left(Q_{ik}^2+\left(P_{ik}-\sqrt{P_{ik}^2+Q_{ik}^2}\right)^2\right)} \tag{3.44}$$

$$-\frac{P_{ik}^2\sqrt{P_{ik}^2+Q_{ik}^2}}{V_{ik}^I\left(P_{ik}^2+\left(Q_{ik}+\sqrt{P_{ik}^2+Q_{ik}^2}\right)^2\right)} \tag{3.45}$$

$$\frac{P_{ik}V_{ik}^R+Q_{ik}V_{ik}^I}{{V_{ik}^I}^2+{V_{ik}^R}^2} \tag{3.46}$$

*Proof.* We can rewrite (3.37) using (3.38) and (3.39) as

$$I_{ik}^R = \frac{P_{ik}V_{ik}^R + Q_{ik}V_{ik}^I}{\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2} \tag{3.47}$$

Thus (3.37)-(3.43), corresponds at the maximization of (3.47) over the set defined by (3.40)-(3.43). Performing the gradient of (3.47) we have:

$$\frac{\partial I_{ik}^R}{\partial V_{ik}^R} = \frac{P_{ik}\left(\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2\right) - 2V_{ik}^R\left(P_{ik}V_{ik}^R + Q_{ik}V_{ik}^I\right)}{\left(\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2\right)^2} \tag{3.48}$$

$$\frac{\partial I_{ik}^R}{\partial V_{ik}^I} = \frac{Q_{ik}\left(\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2\right) - 2V_{ik}^I\left(P_{ik}V_{ik}^R + Q_{ik}V_{ik}^I\right)}{\left(\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2\right)^2} \tag{3.49}$$

$$\frac{\partial I_{ik}^R}{\partial P_{ik}} = \frac{V_{ik}^R}{\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2} \tag{3.50}$$

$$\frac{\partial I_{ik}^R}{\partial Q_{ik}} = \frac{V_{ik}^I}{\left(V_{ik}^I\right)^2 + \left(V_{ik}^R\right)^2} \tag{3.51}$$

the optimal points could be inside the set or in on the boundary region, to this end, we analyze the points in which the gradient is zero. Solving (3.48) with respect to $V_{ik}^I$ we have two solutions:

$$\frac{V_{ik}^R\left(Q_{ik} - \sqrt{P_{ik}^2 + Q_{ik}^2}\right)}{P_{ik}} \tag{3.52}$$

$$\frac{V_{ik}^R\left(Q_{ik} + \sqrt{P_{ik}^2 + Q_{ik}^2}\right)}{P_{ik}} \tag{3.53}$$

putting (3.52) in (3.49)

$$\frac{\partial I_{ik}^R}{\partial V_{ik}^I} = -\frac{2\left(V_{ik}^R\right)^2}{P_{ik}^2}\left(P_{ik}^2 Q_{ik} - P_{ik}^2\sqrt{P_{ik}^2 + Q_{ik}^2} + Q_{ik}^3 - Q_{ik}^2\sqrt{P_{ik}^2 + Q_{ik}^2}\right) \tag{3.54}$$

as can be seen (3.54) is zero if $V_{ik}^R = 0$ that is out of the region described by (3.40)-(3.43). So the optimal solution lies on the boundary region. The boundary region is given by the combination of the following functions calculated in $\underline{V}_i^I, \overline{V}_i^I, \underline{V}_i^R, \overline{V}_i^R$:

$$\frac{Q_{ik}^2\sqrt{P_{ik}^2 + Q_{ik}^2}}{V_{ik}^R\left(Q_{ik}^2 + \left(P_{ik} - \sqrt{P_{ik}^2 + Q_{ik}^2}\right)^2\right)} \tag{3.55}$$

$$-\frac{P_{ik}^2\sqrt{P_{ik}^2+Q_{ik}^2}}{V_{ik}^I\left(P_{ik}^2+\left(Q_{ik}+\sqrt{P_{ik}^2+Q_{ik}^2}\right)^2\right)} \tag{3.56}$$

$$\frac{P_{ik}V_{ik}^R+Q_{ik}V_{ik}^I}{{V_{ik}^I}^2+V_{ik}^{R2}} \tag{3.57}$$

□

Since it is possible to find offline the optimal of the currents bounds definition problem it does not add any computational burden to the overall OPF problem solution. As regards the imaginary current bounds $I_{ik}^I$ the function to be minimized/maximized over the set defined by (3.40)-(3.43) is

$$I_{ik}^I = \frac{P_{ik}V_{ik}^I - Q_{ik}V_{ik}^R}{(V_{ik}^I)^2+(V_{ik}^R)^2} \tag{3.58}$$

the same considerations can be done for (3.58), the points to be considered are:

$$-\frac{P_{ik}^2\sqrt{P_{ik}^2+Q_{ik}^2}}{V_{ik}^R\left(P_{ik}^2+\left(Q_{ik}-\sqrt{P_{ik}^2+Q_{ik}^2}\right)^2\right)} \tag{3.59}$$

$$\frac{Q_{ik}^2\sqrt{P_{ik}^2+Q_{ik}^2}}{V_{ik}^I\left(Q_{ik}^2+\left(P_{ik}-\sqrt{P_{ik}^2+Q_{ik}^2}\right)^2\right)} \tag{3.60}$$

$$\frac{P_{ik}V_{ik}^I-Q_{ik}V_{ik}^R}{(V_{ik}^I)^2+(V_{ik}^R)^2} \tag{3.61}$$

The proposed procedure can be considered as pre-processing for the optimization problem and has to repeated only when the boundary conditions are modified, for example after a change in the forecast of loads/renewables.

**Numerical example**

In order to show the effectiveness of the presented current limit generator approach a numeric example is provided. We consider a 3 nodes unbalanced network with radial topology shown in Figure3.1. The considered 3-phase adimittance matrices for each link are given by:

$$Y_{12} = \begin{bmatrix} 0.433-1.250i & -0.184+0.462i & -0.1+0.345i \\ -0.184+0.462i & 0.379-1.184i & -0.047+0.263i \\ -0.1+0.345i & -0.047+0.263i & 0.335-1.117i \end{bmatrix}$$

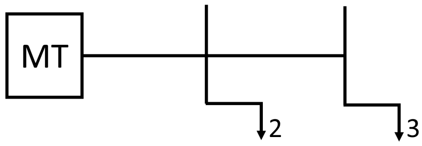


Figure 3.1: 3 nodes test network.

$$Y_{23} = \begin{bmatrix} 0.527-0.675i & -0.141+0.142i & -0.202+0.16i \\ -0.141+0.142i & 0.464-0.668i & -0.103+0.130i \\ -0.202+0.16i & -0.103+0.13i & 0.491-0.671i \end{bmatrix}$$

The overall conductance and susceptance matrix are

$$G = \begin{bmatrix} 0.346 & -0.147 & -0.08 & -0.346 & 0.147 & 0.08 & 0 & 0 & 0 \\ -0.147 & 0.303 & -0.038 & 0.147 & -0.303 & 0.038 & 0 & 0 & 0 \\ -0.08 & -0.038 & 0.268 & 0.08 & 0.038 & -0.268 & 0 & 0 & 0 \\ -0.346 & 0.147 & 0.08 & 0.767 & -0.26 & -0.242 & -0.421 & 0.113 & 0.161 \\ 0.147 & -0.303 & 0.0382 & -0.26 & 0.6747 & -0.121 & 0.113 & -0.371 & 0.082 \\ 0.08 & 0.038 & -0.268 & -0.242 & -0.121 & 0.66 & 0.161 & 0.082 & -0.393 \\ 0 & 0 & 0 & -0.421 & 0.113 & 0.161 & 0.42 & -0.113 & -0.162 \\ 0 & 0 & 0 & 0.111 & -0.371 & 0.082 & -0.113 & 0.371 & -0.082 \\ 0 & 0 & 0 & 0.161 & 0.0823 & -0.393 & -0.162 & -0.082 & 0.392 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.999 & 0.369 & 0.276 & 0.998 & -0.369 & -0.276 & 0 & 0 & 0 \\ 0.369 & -0.946 & 0.21 & -0.369 & 0.946 & -0.211 & 0 & 0 & 0 \\ 0.276 & 0.21 & -0.893 & -0.276 & -0.211 & 0.892 & 0 & 0 & 0 \\ 0.998 & -0.369 & -0.276 & -1.538 & 0.483 & 0.403 & 0.539 & -0.114 & -0.128 \\ -0.369 & 0.9463 & -0.211 & 0.483 & -1.481 & 0.315 & -0.114 & 0.534 & -0.105 \\ -0.276 & -0.211 & 0.892 & 0.403 & 0.315 & -1.429 & -0.128 & -0.105 & 0.536 \\ 0 & 0 & 0 & 0.539 & -0.114 & -0.128 & -0.54 & 0.113 & 0.127 \\ 0 & 0 & 0 & -0.114 & 0.534 & -0.105 & 0.113 & -0.534 & 0.104 \\ 0 & 0 & 0 & -0.128 & -0.105 & 0.536 & 0.127 & 0.1046 & -0.536 \end{bmatrix}$$

Active and reactive power at each node are reported in Table 3.1. The quality of the results obtained from the prepossessing has been tested comparing them with the results of a feasibility optimization problem (no objective function is considered), fixing all the powers at each node except for the stack node whose voltage is fixed to 1 [p.u]. The detailed results of this example are reported in Figures 3.2 and 3.3. The yellow and red lines represent the upper and lower bounds for each variable obtained by the preprocessing, while the cross mark is the actual value of the current (respectively for real and imaginary part). As can be seen from this example, the

| Node | Phase | $P_{ik}[kW]$ | $Q_{ik}[kVAR]$ |
|---|---|---|---|
| 2 | 1 | -200 | -100 |
| 2 | 2 | -300 | -150 |
| 2 | 3 | -500 | -300 |
| 3 | 1 | -200 | -80 |
| 3 | 2 | -500 | -350 |
| 3 | 3 | -100 | -20 |

Table 3.1: Active and Reactive power data for 3 nodes test system

actual currents rely on the bounds obtained through the procedure described in (3.37)-(3.43). The bounds for the first node are wider because the transformer has an operation range not a fixed load/generation.

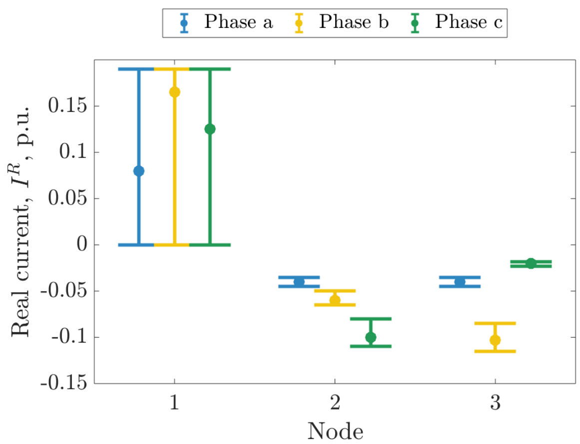


Figure 3.2: Preprocessing results for the 3 node test network, $I^R$.

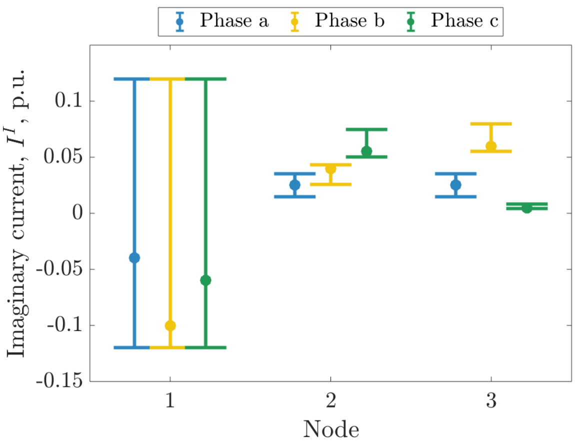


Figure 3.3: Preprocessing results for the 3 node test network, $I^I$.

## 3.4 Case study application

In this section we will test the proposed approach on several case studies. The aim is to show that the CI OPF problem (i.e. the minimization of (3.6) subject to (3.9)-(3.12) and (3.19)-(3.36)) reach a very similar optimal solution compared with the one obtained with the nonlinear solver IPOPT [24] for the nonlinear OPF problem (3.1)-(3.5). Test are made on different networks modifying the IEEE 13 bus, the IEEE 123 bus[25] and the IEEE 8500 nodes [26]. Different topology and the presence of DR and DERs (active network) have been considered, in order to exploit the peculiarities of this model (i.e the maximum flexibility to unbalances ad topology maintaining accuracy in the optimal solution).

The considered networks are:

- IEEE13PR: base case of the IEEE 13 bus system, no changes provided.
- IEEE123PR: base case of the IEEE 123 bus system, no changes provided.
- IEEE8500PR: base case IEEE 8500 nodes
- IEEE13PM: IEEE 13 bus system with mesh topology
- IEEE123PM: IEEE 123 bus system with mesh topology
- IEEE8500PM: IEEE 8500 nodes system mesh topology.
- IEEE13AR: IEEE 13 bus system with DERs
- IEEE123AR: IEEE 123 bus system with DERs
- IEEE8500AR: IEEE 8500 nodes system with DERs.
- IEEE13AM: IEEE 13 combination of active elements and network topology
- IEEE123AM: IEEE 123 combination of active elements and network topology
- IEEE8500AM: IEEE 8500 nodes system combination of active elements and network topology.

These test cases are the most used in literature to test distribution networks and represent a good benchmark to face with. In the present chapter we propose two different classes of test: feasibility, namely the classical power flow analysis, and optimality, the solution of a OPF problem. Both test are important since in future smart grids fast power flow analysis and optimization will be necessary under the massive presence of renewables, to avoid undesired working conditions.

### 3.4.1 Feasibility Test

As a first test it was decided to validate the proposed approach as a feasibility problem, i.e. to solve the load flow problem. To this end, we fix active and reactive power at each node (except for the slack of which the voltage at the node is fixed to 1 [p.u]) obtaining as output each nodal voltage. The aim of this test campaign is to compare the results obtained through the nonlinear model solved through the nonlinear solver IPOPT and the CI-OPF relaxed model solved with Gurobi [27]. In both cases MATLAB with YALMIP [28] were used to build the model. Results obtained on the overall test networks are reported in Table 3.2. Four different metrics are considered: Vmax% and Vmean% are the maximum absolute and the mean percentage error on node voltages respectively, RTNL and RTRX are the run time (expressed in seconds) for the nonlinear and relaxed models.

| Network | Max% | Mean% | RTNL [s] | RTRX [s] |
|---|---|---|---|---|
| IEEE13PR | 0.069 | 0.0527 | 0.19 | 0.03 |
| IEEE123PR | 0.0491 | 0.0399 | 1.191 | 0.1 |
| IEEE8500PR | 0.13 | 0.089 | 73.2 | 2.04 |
| IEEE13PM | 0,0166 | 0.0058 | 0.19 | 0.02 |
| IEEE123PM | 0.0496 | 0.0403 | 1.14 | 0.13 |
| IEEE8500PM | 0.159 | 0.0987 | 85 | 3.5 |
| IEEE13AR | 0.0269 | 0.0161 | 0.26 | 0.04 |
| IEEE123AR | 0.033 | 0.0253 | 1 | 0.22 |
| IEEE8500AR | 0.127 | 0.103 | 73.2 | 3.89 |
| IEEE13AM | 0.0151 | 0.0077 | 0.19 | 0.04 |
| IEEE123AM | 0.0234 | 0.0182 | 0.9 | 0.19 |
| IEEE8500AM | 0.0987 | 0.87 | 73.2 | 3.46 |

Table 3.2: Feasibility test Results

As can be seen from Table 3.2 the two results are very similar, demonstrating that the proposed approach guarantees an excellent performance if it is used as a load flow tool. As a matter of fact, the maximum and average errors are of the magnitude of the $1e{-}2\%$ for all the considered cases. This value is more than acceptable from a practical point of view. From the computational point of view, it can be seen that the relaxed model is solved in one order of magnitude less than the nonlinear one. These results reinforce the effectiveness of the proposed model and the related preprocessing.

As an example in the following Figure 3.4 it is reported the comparison between the real part of voltages obtained with the nonlinear problem and the convex relaxation for the IEEE13PR case. It is noteworthy that figure show that the two solutions are very similar, in fact the solution obtained with the nonlinear solver, indicated with the circle, is superimposed on that obtained with the relaxation indicated with the cross.

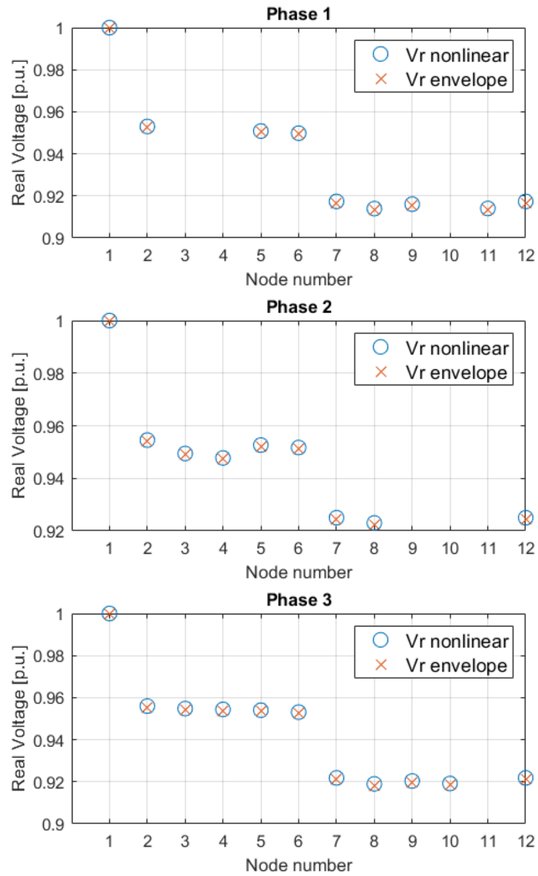


Figure 3.4: Results for IEEE 13PR network

### 3.4.2 Optimality Test

In this section, instead, a comparison will be made regarding the application of the CI-OPF relaxed model to an optimization problem. In particular, the voltage regulation at each grid node through the prosumers' flexibility. Each active user has a regulation range in terms of active and reactive power. This task is a classic problem for the operational management of a DSO, that usually aims to keep the voltage as most possible close the reference value at each node to reduce the overall power losses. The considered objective function is the following:

$$min \sum_{i \in N} \sum_{k \in K} [(V_{ik}^R - \hat{V}_i^R)^2 + (V_{ik}^I - \hat{V}_i^I)^2] \tag{3.62}$$

where $\hat{V}_i^R$ and $\hat{V}_i^I$ are the reference values for the real and imaginary part of the nodal voltage $\forall i \in N$, subject to (3.2)-(3.5)for the nonlinear OPF problem, and (3.9)-(3.12)-(3.19)-(3.36) for the CI OPF relaxed problem.

We put $\hat{V}_i^R = 1$ and $\hat{V}_i^I = 0 \, \forall i \in N$. Together with the metrics introduced for the feasibility test we introduced a new one denoted as $Gap\%$; this metric represent the optimality gap between the result obtained with the CI-OPF model the compared to the "best known" solution obtained with the nonlinear solver IPOPT. We define the following optimality gap:

$$\text{Optimality Gap} = \frac{\text{OPF solution} - \text{CI OPF Relaxed solution}}{\text{CI OPF Relaxed solution}}$$

Results for all the considered test networks are shown in Table 3.3. As can be seen the optimality gap is around $1\%$ for all the simulations. It is noteworthy that the presented lower bound represents a very good approximation of the solution, and can be used in practical applications for the study of distribution grids.

| Network | Max% | Mean% | Gap% | RTNL[s] | RTRX[s] |
|---|---|---|---|---|---|
| IEEE13PR | 0.079 | 0.053 | 0.69 | 0.15 | 0.03 |
| IEEE123PR | 0.051 | 0.04 | 1.11 | 0.1 | 0.09 |
| IEEE8500PR | 0.965 | 0.832 | 0.98 | 73.2 | 3.78 |
| IEEE13PM | 0,0166 | 0.0058 | 0.39 | 0.12 | 0.02 |
| IEEE123PM | 0.0496 | 0.0403 | 1.14 | 0.5 | 0.13 |
| IEEE13AR | 0.0157 | 0.0132 | 1.2451 | 2.60E-01 | 0.04 |
| IEEE123AR | 0.0375 | 0.0227 | 0.9002 | 1 | 0.22 |
| IEEE8500AR | 0.963 | 0.75 | 1.01 | 87.2 | 1.23 |
| IEEE13AM | 0.014 | 0.0091 | 1.1907 | 0.19 | 0.04 |
| IEEE123AM | 0.0338 | 0.0226 | 0.8996 | 0.9 | 0.19 |
| IEEE8500AM | 0.875 | 0.65 | 0.93 | 88 | 2.56 |

Table 3.3: Optimality test results

As far as running time is concerned, from the comparison shown in Table 3.3, it can be seen that the CI OPF model being a quadratic problem solved much faster than the non-linear problem. This computational aspect is very important especially for distributed optimization, as the most popular algorithms of constrained distributed optimization use the Lagrangian to iteratively solve the optimization problem. These techniques need a first optimization step that has the same structure as the original problem, so the simpler the formulation of the problem the faster the convergence.

In summary the proposed convex relaxation has better performances with respect to the approaches in current literature in terms of the wide range of problems it can face with, in fact it can solve both with unbalanced structure and meshed topology, problems that SDP and SOCP approaches cannot solve. Moreover, the computational aspects are very encouraging since the optimality gap obtained is very small and the run times are reduced with respect to a nonlinear solver.

# Chapter 4

# A new bilevel structure for competitive Balancing Markets in distribution grids

This Chapter aims to formulate a competitive structure for the Balancing Market in distribution grids. In this framework, each market player is selfish and wants to maximize its revenues, while the Distribution System Operator (DSO) seeks to improve the quality of service of its network by paying as little as possible. The main contributions of the following Chapter are:

- The statement of a new multi decision-maker architecture to describe BMs for a distribution network, taking into account a new convex relaxation of the electrical grid constraints for general cases such as unbalanced and meshed networks.
- The development of a bi-level BM structure based on multilevel programming, in which every low-level market participant shares its KKT conditions to the upper-level problem to guarantee the optimality of the problem's solution.

A multi-objective optimization problem is formalized for a distribution grid in which are present both traditional generators and aggregators of general prosumers. At the lower level, generators and aggregators want to maximize their profit, taking into account the limits of the flexibility they can sell. At the higher level, the DSO receives power/price bids from market participants and, based on technical and economic considerations, decides the overall power to accept. For the lower-level decision problem, the Karush Khun Tucker (KKT) conditions are developed and included as constraints within the higher-level optimization problem. The resulting optimization problem is applied to the test cases defined in Chapter 2, namely the modified IEEE 13 and 123 bus systems. The organization of this chapter is as follows. Section I investigates the actual literature on the balancing market. Section II defines the equilibrium conditions of each BM participant, in Section III the new DSO optimization problem is formulated. Finally, in Section IV the proposed approach is tested on the two test cases described in Chapter 2.

## 4.1 Literature review

In this chapter, attention is focused on the necessity for the DSO to allocate reserve energy for emergencies and/or improve the quality of service. The grid management, at the distribution level, the uncontrolled installation of renewable energy sources () has led to a lack in the overall hosting capacity (i.e., the overall quantity of renewables affordable by a distribution network without creating power quality losses). The characteristics of this new smart grid require an increase of the power reserve to face the sudden request of active/reactive injection/absorption from a Distribution System Operator (DSO) to compensate for example a sudden drop in the production from a photovoltaic plant. One of the main problems for an electricity market operator that also owns the distribution grid is to avoid congestions and maximize the quality of the service provided [29]. For this reason, it is important to induce the balancing market (BM) sessions, after the closure of the retail energy market (REM). As a matter of fact, after the REM sessions, the DSO should carry out technological evaluations, performing a power flow analysis of the congestions over the network. For instance, the voltage at a certain node may be too high, or that some lines are close to their ampacity limit. One possible solution to the aforementioned problem is to introduce BM sessions through which the market operator acquires flexibility from the participants to avoid the possible congestions over the network [30].

In this framework, demand response (DR), i.e., the possibility for a traditional load to decrease the active power absorption for a certain period, is an effective and reliable strategy for the successful integration of renewable energy sources, handling the demand curve using load flexibility whenever the system requires it [31], [32], [33]. A significant portion of DR programs can be represented by local users and/or prosumers, which, by themselves, cannot participate in the energy market and offer a reduction of load at a certain price. Microgrids and smart buildings represent an opportunity to provide regulation and reserve services in both directions since they can be considered as prosumers [34], [35], [36]. Interconnected buildings can be seen as microgrids or energy communities that can share thermal and electrical power to satisfy comfort, economic and technical exigencies.

A new kind of market entity, namely aggregator, can participate in the BM. This market actor manages different local several plants (customers) which, singularly, could not be able to participate in the BM, providing an aggregate DR or power curtailment ancillary service to the distribution grid [37]. This new framework opens new challenges for the development of new energy management systems, models, and methods.

In the current literature, different optimization problems are formalized for the electricity market but referred specifically to the transmission grid. In [38], a pricing mechanism is proposed, formulating a non-cooperative game among the multiple market participants, which sets the retail prices to maximize their profits. There are several papers regarding the new operational management strategies for the optimization and control of a distribution network in the presence of DR aggregators. In particular, it is necessary to define how to schedule load reduction or production increase among the different customers to achieve an overall load reduction and/or shifting. As well as generators and aggregators, electrical storage systems can participate in the BM, as discussed in [39]. In [40], the energy market is analyzed taking into account the presence of renewables and demand response services. The same authors in [41] propose a distributed optimization approach for the REM based on the dynamic market mechanism (DMM) approach. Similarly to what we present in this paper, a hierarchical approach is proposed in [42]. The problem is solved by developing sensitivity functions for market participants' payoff for their bidding strategies. In [43] the authors study the operation of a retailer that aggregates a group of price-responsive loads and submits block-wise demand bids to the day-ahead and real-time markets, to consider the long term and short term market dynamics. The work presented in [44] proposes a two-stage stochastic model for a large-scale energy resource scheduling problem of aggregators, using Benders' decomposition method is used to reduce the overall problem's computational burden. In [45] customer aggregators are introduced to supply downstream demand most economically. Aggregators' flexible energy demand is incorporated into the centralized energy dispatch model forming a two-level optimization problem where the system operator maximizes social welfare minimizing the energy purchase cost. In [46] and [47] the authors study jointly energy storage systems and DR analyzing their impact and possible use in a smart grid. In [48] the authors focus attention on the integration of residential users because they are in huge numbers and relatively small power loads. It is showed that residential users can receive financial benefits from reserve provision. A hierarchical structure of users is adopted together with a game-theoretic approach. In [49] authors focus attention on plug-in Electric vehicles in a district and on how they can help in demand-side management. In the experiments, the results show that the proposed system can reduce a significant amount of both electricity cost and peak power.

A crucial point is the flexibility potential estimation of different types of consumers for day-ahead and real-time ancillary services provision taking into account customers' reactions to the price signals [50]. Moreover, the use of incentive-based payments as price offer packages is recommended to implement demand response [51]. As well as the scheduling of an aggregator's customers, it is fundamental to design the optimal set of customers for an aggregator. In this framework, authors in [52] study the impact of consumer behavior on the portfolio design of a DR aggregator, developing an optimization model.

Differently from the current literature the proposed market structure, three classes of agents are considered: traditional generators, aggregators, and the market operator that in this case coincides with the DSO. We define a multilevel (two in this case) market architecture. At the upper level, there is the DSO that acquires the minimum amount of ancillary services to have an acceptable operating condition, while at the lower one there are the generators and aggregators that want to maximize their revenues in terms of services provided to the DSO. The main difference with the current literature lies in the accuracy in modeling both the economic entities (through supply functions) and the distribution grid, by the definition of appropriate KPIs for the service quality and accurate power flow modeling described in Chapter 3. The assumption that any lower-level decision maker optimizes its objective is converted into equilibrium constraints

([53]) for the higher optimization problem. More specifically, the Karush Khun Tucker (KKT) conditions for the lower level problems are included as constraints within the higher-level optimization problem. This ensures that the decisions are taken by the DSO take into account the optimal behavior of lower-level decision-makers. It is important to note that, no negotiation in this model is foreseen between and market participants. Actually, in the current literature (see for example ([54])), iterative market structures are proposed. From our point of view, these approaches are not realistic ([55]), since the most actual market structures are based on a single decision process. In summary, the innovation of the proposed approach is :

- The definition of a multiobjective and multi decision-maker market structure for the BM in distribution grids that considers both technical and economical aspects.
- The formalization of a bilevel optimization problem for the balancing reserve definition for general distribution grids taking into account multi-phase and unbalanced settings

## 4.2 The proposed bilevel competitive Balancing Market structure

The scheme of the bi-level architecture is reported in Figure 1. At the lower level, there are the market participants, i.e. the generators and aggregators that want to maximize their profit, taking into account the limits of the flexibility they can sell. At the higher level, power/price bids are received from market participants, and, based on technical and economic considerations, the DSO decides which bid to accept. The proposed approach in this chapter consists of formalizing a unique optimization problem for DSO, assuming that any market participant at the lower level behaves to optimize its performance objective. This is accomplished by providing all parameters that are necessary to write the KKT conditions ensuring the optimality conditions for the lower level problem. This fact makes the market problems considered interrelated, which results in a well-posed equilibrium problem. Specifically, this problem is an instance of the Generalized Nash Equilibrium Problem (GNEP). The GNEP is a generalization of the well known Nash Equilibrium Problem (NEP). In a GNEP each player's set of constraints depends on the rival players' decisions; that is, the actions of other players are not considered fixed for establishing the decisions of each market participant [53]. It is important to point out to the reader that the DSO solves its optimization problem by determining the quantities of power to accept based on technical and economic choices. For this reason, an offer at lower price might be not accepted because it is not in a strategic position for network services' quality.

In the following subsections, the equilibrium constraints for each market participant are obtained. It is important to point out to the reader that until now the supply functions that indicate prices according to power are always left indicated to maintain a generality of the model presented. In literature, the model is not univocal, often quadratic functions are used because they guarantee

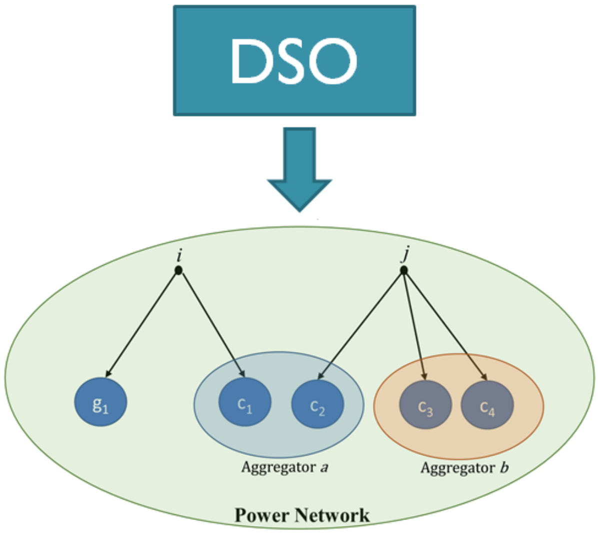

Figure 4.1: The proposed bi-level market structure.

strong convexity properties compared to linear functions, which on the other hand are more reasonable from an engineering point of view. For this reason, the equilibrium constraints will be formalized in both cases.

### 4.2.1 Traditional Generators

As described in Chapter 2, medium size traditional generators can participate in the BM session without an aggregator since the can offer at least the minimum amount of required flexibility. In order to formulate the KKT optimality condition for this market participant we refer to the optimization problem formulated in Chapter 2 (2.1)-(2.6). Defining $\sigma_{c,j}$ as the Lagrange multipliers associated to constraint $j$ of generator $g$, the Lagrange function for this problem is:

$$
\begin{aligned}
&L_g(x_g, \sigma_{g,j}, j = 1, ..., 7) = \\
&-\Big\{\lambda_g^{P,up}(P_g^{up}) - \hat{C}_g P_g^{up} - \lambda_g^{P,dw}(P_g^{dw}) - \check{C}_g P_g^{dw} + \lambda_g^{Q,up}(Q_g^{up}) - \lambda_g^{Q,dw}(Q_g^{dw})\Big\} \\
&-\sigma_{g,1} P_g^{up} + \sigma_{g,2} P_g^{dw} + \sigma_{g,3}(P_g^{up} - \hat{P}_g^{*}) - \sigma_{g,4}(P_g^{dw} + \check{P}_g^{*}) - \sigma_{g,5} Q_g^{up} + \sigma_{g,6} Q_g^{dw} \\
&+\sigma_{g,7}[(P_g^{up} + P_g^{da} + P_g^{dw})^2 + (Q_g^{up} + Q_g^{da} + Q_g^{dw})^2 - S_g^2]
\end{aligned}
\tag{4.1}
$$

The KKT conditions are given by:

$$
\frac{\partial L_g}{\partial P_g^{up}} = -\frac{d\,\lambda_a^{P,up}(P_g^{up})}{d\,P_g^{up}} + \hat{C}_g - \sigma_{g,1} + \sigma_{g,3} + 2\sigma_{g,7}(P_g^{up} + P_g^{da} + P_g^{dw}) = 0 \tag{4.2}
$$

$$
\frac{\partial L_g}{\partial P_g^{dw}} = \frac{d\,\lambda_g^{P,dw}(P_g^{dw})}{d\,P_g^{dw}} - \check{C}_g - \sigma_{g,4} + \sigma_{g,2} + 2\sigma_{g,7}(P_g^{up} + P_g^{da} + P_g^{dw}) = 0 \tag{4.3}
$$

$$\frac{\partial L_g}{\partial Q_g^{up}} = -\frac{d\,\lambda_g^{Q,up}(Q_g^{up})}{d\,Q_g^{up}} - \sigma_{g,5} + 2\sigma_{g,7}(Q_g^{up} + Q_g^{da} + Q_g^{dw}) = 0 \tag{4.4}$$

$$\frac{\partial L_g}{\partial Q_g^{dw}} = \frac{d\,\lambda_g^{Q,dw}(Q_g^{dw})}{d\,Q_g^{dw}} + \sigma_{g,6} + 2\sigma_{g,7}(Q_g^{up} + Q_g^{da} + Q_g^{dw}) = 0 \tag{4.5}$$

$$P_g^{up} \geq 0 \tag{4.6}$$

$$P_g^{dw} \leq 0 \tag{4.7}$$

$$(P_g^{up} - \hat{P}_g^{*}) \leq 0 \tag{4.8}$$

$$-(P_g^{dw} - \check{P}_g^{*}) \leq 0 \tag{4.9}$$

$$(P_g^{up} + P_g^{da} + P_g^{dw})^2 + (Q_g^{up} + Q_g^{da} + Q_g^{dw})^2 - S_g^2 \leq 0 \tag{4.10}$$

$$\sigma_{g,1} P_g^{up} = 0 \tag{4.11}$$

$$\sigma_{g,2} P_g^{dw} = 0 \tag{4.12}$$

$$\sigma_{g,3}(P_g^{up} - \hat{P}_g^{*}) = 0 \tag{4.13}$$

$$\sigma_{g,4}(P_g^{dw} + \check{P}_g^{*}) = 0 \tag{4.14}$$

$$Q_g^{up} \geq 0 \tag{4.15}$$

$$Q_g^{dw} \leq 0 \tag{4.16}$$

$$\sigma_{g,5} Q_g^{up} = 0 \tag{4.17}$$

$$\sigma_{g,6} Q_g^{dw} = 0 \tag{4.18}$$

$$\sigma_{g,7}[(P_g^{up} + P_g^{da} + P_g^{dw})^2 + (Q_g^{up} + Q_g^{da} + Q_g^{dw})^2 - S_g^2] = 0 \tag{4.19}$$

$$\sigma_{g,j} \geq 0 \quad j = 1, .., 7 \tag{4.20}$$

#### 4.2.1.1 Linear supply function

If we consider linear supply function for the balancing services of a generator $g$ we have:

$$\lambda_g^{P,up}(P_g^{up}) = a_{1_g}^{P,up} P_g^{up} + c_{1_g}^{P,up} \implies \frac{d\,\lambda_g^{P,up}(P_g^{up})}{d\,P_g^{up}} = a_{1_g}^{P,up} \tag{4.21}$$

$$\lambda_g^{Q,up}(Q_g^{up}) = a_{1_g}^{Q,up} Q_g^{up} + c_{1_g}^{Q,up} \implies \frac{d\,\lambda_g^{Q,up}(Q_g^{up})}{d\,Q_g^{up}} = a_{1_g}^{Q,up} \tag{4.22}$$

$$\lambda_g^{P,dw}(P_g^{dw}) = a_{1_g}^{P,dw} P_g^{dw} + c_{1_g}^{P,dw} \implies \frac{d\,\lambda_g^{P,dw}(P_g^{dw})}{d\,P_g^{dw}} = a_{1_g}^{P,dw} \tag{4.23}$$

$$\lambda_g^{Q,dw}(Q_g^{dw}) = a_{1_g}^{Q,dw} Q_g^{dw} + c_{1_g}^{Q,dw} \implies \frac{d\,\lambda_g^{Q,dw}(Q_g^{dw})}{d\,Q_g^{dw}} = a_{1_g}^{Q,dw} \tag{4.24}$$

#### 4.2.1.2 Quadratic supply function

In this case, if we consider a quadratic supply function for the balancing services of a generator $g$ we have:

$$\lambda_g^{P,up}(P_g^{up}) = a_{2_g}^{P,up}(P_g^{up})^2 + a_{1_g}^{P,up}P_g^{up} + c_{1_g}^{P,up} \implies \frac{d\,\lambda_g^{P,up}(P_g^{up})}{d\,P_g^{up}} = 2a_{2_g}^{P,up}P_g^{up} + a_{1_g}^{P,up} \quad (4.25)$$

$$\lambda_g^{Q,up}(Q_g^{up}) = a_{2_g}^{Q,up}(Q_g^{up})^2 + a_{1_g}^{Q,up}Q_g^{up} + c_{1_g}^{Q,up} \implies \frac{d\,\lambda_g^{Q,up}(Q_g^{up})}{d\,Q_g^{up}} = 2a_{2_g}^{Q,up}Q_g^{up} + a_{1_g}^{Q,up} \quad (4.26)$$

$$\lambda_g^{P,dw}(P_g^{dw}) = a_{2_g}^{P,dw}(Q_g^{dw})^2 + a_{1_g}^{P,dw}P_g^{dw} + c_{1_g}^{P,dw} \implies \frac{d\,\lambda_g^{P,dw}(P_g^{dw})}{d\,P_g^{dw}} = 2a_{2_g}^{P,up}P_g^{dw} + a_{1_g}^{P,dw} \quad (4.27)$$

$$\lambda_g^{Q,dw}(Q_g^{dw}) = a_{2_g}^{Q,dw}(Q_g^{dw})^2 + a_{1_g}^{Q,dw}Q_g^{dw} + c_{1_g}^{Q,dw} \implies \frac{d\,\lambda_g^{Q,dw}(Q_g^{dw})}{d\,Q_g^{dw}} = 2a_{2_g}^{Q,dw}Q_g^{dw} + a_{1_g}^{Q,dw} \quad (4.28)$$

where $a_{1_g}^{P,up}$, $a_{1_g}^{Q,up}$, $a_{1_g}^{P,dw}$, $a_{1_g}^{Q,dw}$, $c_{1_g}^{P,up}$, $c_{1_g}^{Q,up}$, $c_{1_g}^{P,dw}$ and $c_{1_g}^{Q,dw}$ are known parameters.

### 4.2.2 Aggregators

The aggregator represents a certain energy reserve that is provided by several customers belonging to the same aggregation area. This decision maker establishes contracts with its customers in exchange of a profit for the provided service. In order to formulate the KKT optimality condition for this market participant we refer to the optimization problem formulated in Chapter 2 (2.7)-(2.12).

Defining $\sigma_{c,j}$ as the Lagrange multipliers associated to constraint $j$ of customer $c$, the Lagrange function for the aggregator $a$ is:

$$L_a(x_c, \sigma_{c,j} \quad j = 1, ..., 7, \ \ c \in C(a)) =$$
$$-\left\{ \sum_{c\in C(a)} \lambda_a^{P,up}(P_c^{up}) - \sum_{c\in C(a)} \hat{C}_c P_c^{up} - \sum_{c(a)} \lambda_a^{P,dw}(P_c^{dw}) + \sum_{c\in C(a)} \check{C}_c P_c^{dw} + \sum_{c\in C(a)} \lambda_a^{Q,up}(Q_c^{up}) - \right.$$
$$\left. \sum_{c\in C(a)} \lambda_a^{Q,dw}(Q_c^{dw}) \right\} - \sigma_{c,1}P_c^{up} + \sigma_{c,2}(P_c^{up} - \hat{P}_c^{*}) + \sigma_{c,3}P_c^{dw} - \sigma_{c,4}(P_c^{dw} + \check{P}_c^{*})$$
$$-\sigma_{c,5}Q_c^{up} + \sigma_{c,6}Q_c^{dw} + \sigma_{c,7}[(P_c^{up} + P_c^{da} + P_c^{dw})^2 + (Q_c^{up} + Q_c^{da} + Q_c^{dw})^2 - S_c^2] \quad (4.29)$$

The KKT conditions are

$$\frac{\partial L_a}{\partial P_c^{up}} = -\frac{d\,\lambda_a^{P,up}(P_c^{up})}{d\,P_c^{up}} + \hat{C}_c - \sigma_{c,1} + \sigma_{c,2} + 2\sigma_{c,7}(P_c^{up} + P_c^{da} + P_c^{dw}) = 0 \ \forall c \in C(a) \quad (4.30)$$

$$\frac{\partial L_a}{\partial P_c^{dw}} = \frac{d\,\lambda_a^{P,dw}(P_c^{dw})}{d\,P_c^{dw}} - \check{C}_c - \sigma_{c,4} + \sigma_{c,3} + 2\sigma_{c,7}(P_c^{up} + P_c^{da} + P_c^{dw}) = 0\ \forall c \in C(a) \quad (4.31)$$

$$\frac{\partial L_a}{\partial Q_c^{up}} = -\frac{d\,\lambda_a^{Q,up}(Q_c^{up})}{d\,Q_c^{up}} - \sigma_{c,5} + 2\sigma_{c,7}(Q_c^{up} + Q_c^{da} + Q_c^{dw}) = 0\ \forall c \in C(a) \quad (4.32)$$

$$\frac{\partial L_a}{\partial Q_c^{dw}} = \frac{d\,\lambda_a^{P,dw}(Q_c^{dw})}{d\,Q_c^{dw}} + \sigma_{c,6} + 2\sigma_{c,7}(Q_c^{up} + Q_c^{da} + Q_c^{dw}) = 0\ \forall c \in C(a) \quad (4.33)$$

$$(P_c^{up} + P_c^{da} + Q_c^{dw})^2 + (Q_c^{up} + Q_c^{da} + Q_c^{dw})^2 - S_c^2 \leq 0\ \forall c \in C(a) \quad (4.34)$$

$$-P_c^{up} \leq 0\ \forall c \in C(a) \quad (4.35)$$

$$P_c^{dw} \leq 0\ \ \forall c \in C(a) \quad (4.36)$$

$$Q_c^{dw} \leq 0 \quad (4.37)$$

$$Q_c^{up} \geq 0 \quad (4.38)$$

$$(P_c^{up} - \hat{P}_c^{*})\ \leq 0\ \forall c \in C(a) \quad (4.39)$$

$$-(P_c^{dw} + \check{P}_c^{*})\ \leq 0\ \forall c \in C(a) \quad (4.40)$$

$$\sigma_{c,1} P_c^{up} = 0\ \forall c \in C(a) \quad (4.41)$$

$$\sigma_{c,3} P_c^{dw} = 0\ \ \forall c \in C(a) \quad (4.42)$$

$$\sigma_{c,2}(P_c^{up} - \hat{P}_c^{*})\ = 0\ \forall c \in C(a) \quad (4.43)$$

$$\sigma_{c,4}(P_c^{dw} + \check{P}_c^{*})\ = 0\ \forall c \in C(a) \quad (4.44)$$

$$\sigma_{c,5} Q_c^{up} = 0\ \forall c \in C(a) \quad (4.45)$$

$$\sigma_{c,6} Q_c^{dw} = 0\ \forall c \in C(a) \quad (4.46)$$

$$\sigma_{c,7}[(P_c^{up} + P_c^{da} + P_c^{dw})^2 + (Q_c^{up} + Q_c^{da} + Q_c^{dw})^2 - S_c^2] = 0\ \forall c \in C(a) \quad (4.47)$$

$$\sigma_{c,j}, j = 1, ..., 7,\ \forall c \in C(a) \quad (4.48)$$

#### 4.2.2.1 Linear supply function

If we consider linear supply function for the balancing services of a aggregator $a$ we have:

$$\lambda_a^{P,up}(P_c^{up}) = a_{1_a}^{P,up} P_c^{up} + c_{1_c}^{P,up} \implies \frac{d\,\lambda_a^{P,up}(P_c^{up})}{d\,P_c^{up}} = a_{1_a}^{P,up}\ \forall c \in C(a) \quad (4.49)$$

$$\lambda_a^{Q,up}(Q_c^{up}) = a_{1_a}^{Q,up} Q_c^{up} + c_{1_c}^{Q,up} \implies \frac{d\,\lambda_c^{Q,up}(Q_c^{up})}{d\,Q_c^{up}} = a_{1_a}^{Q,up}\ \forall c \in C(a) \quad (4.50)$$

$$\lambda_a^{P,dw}(P_c^{dw}) = a_{1_a}^{P,dw} P_c^{dw} + c_{1_c}^{P,dw} \implies \frac{d\,\lambda_c^{P,dw}(P_c^{dw})}{d\,P_c^{dw}} = a_{1_a}^{P,dw}\ \forall c \in C(a) \quad (4.51)$$

$$\lambda_a^{Q,dw}(Q_c^{dw}) = a_{1_a}^{Q,dw} Q_c^{dw} + c_{1_c}^{Q,dw} \implies \frac{d\,\lambda_c^{Q,dw}(Q_c^{dw})}{d\,Q_c^{dw}} = a_{1_a}^{Q,dw}\ \forall c \in C(a) \quad (4.52)$$

where $a_{1_a}^{P,up}$, $a_{1_a}^{Q,up}$, $a_{1_a}^{P,dw}$, $a_{1_a}^{Q,dw}$, $c_{1_c}^{P,up}$, $c_{1_c}^{Q,up}$, $c_{1_c}^{P,dw}$ and $c_{1_c}^{Q,dw}$ are known parameters.

#### 4.2.2.2 Quadratic supply function

If we consider a quadratic supply function for the balancing services of a aggregator $a$ we have:

$$\begin{aligned} \lambda_a^{P,up}(P_c^{up}) = a_{2_a}^{P,up}(P_c^{up})^2 + a_{1_a}^{P,up} P_c^{up} + c_{1_a}^{P,up} \implies \\ \frac{d\,\lambda_a^{P,up}(P_c^{up})}{d\,P_c^{up}} = 2a_{2_a}^{P,up} P_c^{up} + a_{1_a}^{P,up} \qquad \forall c \in C(a) \end{aligned} \tag{4.53}$$

$$\begin{aligned} \lambda_a^{Q,up}(Q_c^{up}) = a_{2_a}^{Q,up}(Q_c^{up})^2 + a_{1_a}^{Q,up} Q_c^{up} + c_{1_a}^{Q,up} \implies \\ \frac{d\,\lambda_a^{Q,up}(Q_c^{up})}{d\,Q_c^{up}} = 2a_{2_a}^{Q,up} Q_c^{up} + a_{1_a}^{Q,up} \qquad \forall c \in C(a) \end{aligned} \tag{4.54}$$

$$\begin{aligned} \lambda_a^{P,dw}(P_c^{dw}) = a_{2_a}^{P,dw}(Q_c^{dw})^2 + a_{1_a}^{P,dw} P_c^{dw} + c_{1_a}^{P,dw} \implies \\ \frac{d\,\lambda_a^{P,dw}(P_c^{dw})}{d\,P_c^{dw}} = 2a_{2_a}^{P,up} P_c^{dw} + a_{1_a}^{P,dw} \qquad \forall c \in C(a) \end{aligned} \tag{4.55}$$

$$\begin{aligned} \lambda_a^{Q,dw}(Q_c^{dw}) = a_{2_a}^{Q,dw}(Q_c^{dw})^2 + a_{1_a}^{Q,dw} Q_c^{dw} + c_{1_a}^{Q,dw} \implies \\ \frac{d\,\lambda_a^{Q,dw}(Q_c^{dw})}{d\,Q_c^{dw}} = 2a_{2_a}^{Q,dw} Q_c^{dw} + a_{1_a}^{Q,dw} \qquad \forall c \in C(a) \end{aligned} \tag{4.56}$$

## 4.3 The DSO Optimization Problem

It is now necessary to define the new optimization of the DSO, which is modified compared to the one described in Chapter 2, considering the new distribution grid model described in Chapter 3 and the equilibrium conditions of market participants stated in the previous section. The overall multiobjective optimization problem is given by:

$$\min_{x_g \in GT, x_c c \in C(a),\, a \in A, V_{ik}^R, V_{ik}^I, I_{ik}^R, I_{ik}^I} \sum_{g \in GT} \sum_{k \in K} \left\{ \lambda_{gk}^{P,up}(P_{gk}^{up}) - \lambda_{gk}^{P,dw}(P_{gk}^{dw}) + \lambda_{gk}^{Q,up}(Q_{gk}^{up}) - \lambda_{gk}^{Q,dw}(Q_{gk}^{dw}) \right\}$$
$$+ \sum_{a \in A} \sum_{k \in GT} \left\{ \sum_{c \in C(a)} \lambda_a^{P,up}(P_{ck}^{up}) - \sum_{c \in C(a)} \lambda_a^{P,dw}(P_{ck}^{dw}) + \sum_{c \in C(a)} \lambda_a^{Q,up}(Q_{ck}^{up}) - \sum_{c \in C(a)} \lambda_a^{Q,dw}(Q_{ck}^{dw}) \right\}$$
$$+ \alpha SI_{sys} \tag{4.57}$$

$$SI_{sys} = \sum_{h \in (GT \cup C_{tot})} SI_h^G + \sum_{i \in N} SI_i^V + \sum_{i \in N} \sum_{\substack{j \in N \\ i \neq j}} SI_{ij}^C \tag{4.58}$$

$$SI_h^G = \frac{P_{gk}^{da} + P_{gk}^{up} + P_{gk}^{dw}}{P_{gk}^{\max}} \quad \forall h \in (GT \cup C_{TOT}) \tag{4.59}$$

$$SI_{ik}^V = \left( \frac{\|V_{ik}\| - \hat{V}_{ik}}{\Delta V_i^*} \right)^2 \forall i \in N \;\; k \in K \tag{4.60}$$

$$SI_{ijk}^C = \left( \frac{I_{ijk}}{I_{ij}^*} \right)^2 \forall i, j \in N \;\; k \in K \tag{4.61}$$

$$P_{ik} = \sum_{\substack{g \in GT: \\ i = Node(g)}} \left( P_{gk}^{da} + P_{gk}^{up} + P_{gk}^{dw} \right) + \sum_{a \in A} \sum_{\substack{c \in C(a): \\ i = Node(c)}} \left( P_{ck}^{da} + P_{ck}^{up} + P_{ck}^{dw} \right) \;\; \forall i \in N \;\; k \in K \tag{4.62}$$

$$Q_{ik} = \sum_{\substack{g \in GT: \\ i = Node(g)}} \left( Q_{gk}^{da} + Q_{gk}^{up} + Q_{gk}^{dw} \right) + \sum_{a \in A} \sum_{\substack{c \in C(a): \\ i = Node(c)}} \left( Q_{ck}^{da} + Q_{ck}^{up} + Q_{ck}^{dw} \right) \;\; \forall i \in N \;\; k \in K \tag{4.63}$$

$$I^R = \mathrm{Re}\left(\mathbf{Y}V\right) \tag{4.64}$$

$$I^I = \mathrm{Im}\left(\mathbf{Y}V\right) \tag{4.65}$$

$$P_{ik} = \ a_{ik} + b_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.66}$$

$$Q_{ik} = -c_{ik} + d_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.67}$$

$$a_{ik} \geq \underline{V_{ik}}^R I^R_{ik} + V_{ik}{}^R \underline{I}^R_{ik} - \underline{V_{ik}}^R \underline{I}^R_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.68}$$

$$a_{ik} \geq \overline{V_{ik}}^R I^R_{ik} + V_{ik}{}^R \overline{I}^R_{ik} - \overline{V_{ik}}^R \overline{I}^R_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.69}$$

$$a_{ik} \leq \underline{V_{ik}}^R I^R_{ik} + V_{ik}{}^R \overline{I}^R_{ik} - \overline{V_{ik}}^R \underline{I}^R_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.70}$$

$$a_{ik} \leq \overline{V_{ik}}^R I^R_{ik} + V_{ik}{}^R \underline{I}^R_{ik} - \underline{V_{ik}}^R \overline{I}^R_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.71}$$

$$b_{ik} \geq \underline{V_{ik}}^I I^I_{ik} + V_{ik}{}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \underline{I}^R_{ik} \ \ \forall i \in N,\ \ k \in K \tag{4.72}$$

$$b_{ik} \geq \overline{V_{ik}}^I I^I_{ik} + V_{ik}{}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \overline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.73}$$

$$b_{ik} \leq \underline{V_{ik}}^I I^I_{ik} + V_{ik}{}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \underline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.74}$$

$$b_{ik} \leq \overline{V_{ik}}^I I^I_{ik} + V_{ik}{}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \overline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.75}$$

$$c_{ik} \geq \underline{V_{ik}}^R I^I_{ik} + V_{ik}{}^R \underline{I_{ik}}^I - \underline{V_{ik}}^R \underline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.76}$$

$$c_{ik} \geq \overline{V_{ik}}^R I^I_{ik} + V_{ik}{}^R \overline{I_{ik}}^I - \overline{V_{ik}}^R \overline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.77}$$

$$c_{ik} \leq \underline{V_{ik}}^R I^I_{ik} + V_{ik}{}^R \overline{I_{ik}}^I - \overline{V_{ik}}^R \underline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.78}$$

$$c_{ik} \leq \overline{V_{ik}}^R I^I_{ik} + V_{ik}{}^R \underline{I_{ik}}^I - \underline{V_{ik}}^R \overline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.79}$$

$$d_{ik} \geq \underline{V_{ik}}^I I^R_{ik} + V_{ik}{}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \underline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.80}$$

$$d_{ik} \geq \overline{V_{ik}}^I I^R_{ik} + V_{ik}{}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \overline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.81}$$

$$d_{ik} \leq \underline{V_{ik}}^I I^R_{ik} + V_{ik}{}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \underline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.82}$$

$$d_{ik} \leq \overline{V_{ik}}^I I^R_{ik} + V_{ik}{}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \overline{I_{ik}}^I \ \ \forall i \in N,\ \ k \in K \tag{4.83}$$

$$\underline{V_{ik}}^R \leq I^R_{ik} \leq \overline{V_{ik}}^R \ \ \forall i \in N \ k \in K \tag{4.84}$$

$$\underline{V_{ik}}^I \leq I^I_{ik} \leq \overline{V_{ik}}^I \ \ \forall i \in N \ k \in K \tag{4.85}$$

$$\underline{I_{ik}}^R \leq I^R_{ik} \leq \overline{I_{ik}}^R \ \ \forall i \in N \ k \in K \tag{4.86}$$

$$\underline{I_{ik}}^I \leq I^I_{ik} \leq \overline{I_{ik}}^I \ \ \forall i \in N \ k \in K \tag{4.87}$$

$$\underline{P}_{ik} \leq P_{ik} \leq \overline{P}_{ik} \ \ \forall i \in N \ k \in K \tag{4.88}$$

$$\underline{Q}_{ik} \leq Q_{ik} \leq \overline{Q}_{ik} \ \ \forall i \in N,\ k \in K \tag{4.89}$$

$$-\frac{d\,\lambda_{gk}^{P,up}(P_{gk}^{up})}{d\,P_{gk}^{up}} + \hat{C_{gk}} - \sigma_{gk,1} + \sigma_{gk,3} + 2\sigma_{gk,7}(P_{gk}^{up} + P_{gk}^{da} + P_{gk}^{dw}) = 0\ \forall G \in GT,\ k \in K \tag{4.90}$$

$$\frac{d\,\lambda_{gk}^{P,dw}(P_{gk}^{dw})}{d\,P_{gk}^{dw}} - \check{C_{gk}} - \sigma_{gk,4} + \sigma_{gk,2} + 2\sigma_{gk,7}(P_{gk}^{up} + P_{gk}^{da} + P_{gk}^{dw}) = 0\ \forall G \in GT,\ k \in K \tag{4.91}$$

$$-\frac{d\,\lambda_{gk}^{Q,up}(Q_{gk}^{up})}{d\,Q_{gk}^{up}} - \sigma_{gk,5} + 2\sigma_{gk,7}(Q_{gk}^{up} + Q_{gk}^{da} + Q_{gk}^{dw}) = 0\ \forall G \in GT,\ k \in K \tag{4.92}$$

$$\frac{d\,\lambda_{gk}^{Q,dw}(Q_{gk}^{dw})}{d\,Q_{gk}^{dw}} + \sigma_{gk,6} + 2\sigma_{gk,7}(Q_{gk}^{up} + Q_{gk}^{da} + Q_{gk}^{dw}) = 0\ \forall G \in GT,\ k \in K \tag{4.93}$$

$$P_{gk}^{up} \geq 0\ \forall G \in GT,\ k \in K \tag{4.94}$$

$$P_{gk}^{dw} \leq 0\ \forall G \in GT,\ k \in K \tag{4.95}$$

$$(P_{gk}^{up} - \hat{P_{gk}^{*}})\ \leq 0\ \forall G \in GT,\ k \in K \tag{4.96}$$

$$-(P_{gk}^{dw} - \check{P_{gk}^{*}})\ \leq 0\ \forall G \in GT,\ k \in K \tag{4.97}$$

$$(P_{gk}^{up} + P_{gk}^{da} + P_{gk}^{dw})^2 + (Q_{gk}^{up} + Q_{gk}^{da} + Q_{gk}^{dw})^2 - S_{gk}^2 \leq 0\ \forall G \in GT,\ k \in K \tag{4.98}$$

$$\sigma_{gk,1} P_{gk}^{up} = 0\ \forall G \in GT,\ k \in K \tag{4.99}$$

$$\sigma_{gk,2} P_{gk}^{dw} = 0\ \forall G \in GT,\ k \in K \tag{4.100}$$

$$\sigma_{gk,3}(P_{gk}^{up} - \hat{P_{gk}^{*}}) = 0\ \forall G \in GT,\ k \in K \tag{4.101}$$

$$\sigma_{gk,4}(P_{gk}^{dw} + \check{P_{gk}^{*}}) = 0\ \forall G \in GT,\ k \in K \tag{4.102}$$

$$Q_{gk}^{up} \geq 0\ \forall G \in GT,\ k \in K \tag{4.103}$$

$$Q_{gk}^{dw} \leq 0\ \forall G \in GT,\ k \in K \tag{4.104}$$

$$\sigma_{gk,5} Q_{gk}^{up} = 0\ \forall G \in GT,\ k \in K \tag{4.105}$$

$$\sigma_{gk,6} Q_{gk}^{dw}\ = 0\ \forall G \in GT,\ k \in K \tag{4.106}$$

$$\sigma_{gk,7}[(P_{gk}^{up} + P_{gk}^{da} + P_{gk}^{dw})^2 + (Q_{gk}^{up} + Q_{gk}^{da} + Q_{gk}^{dw})^2 - S_{gk}^2] = 0\ \forall G \in GT,\ k \in K \tag{4.107}$$

$$\sigma_{gk,j} \geq 0 \quad j = 1,..,7\ \forall G \in GT,\ k \in K \tag{4.108}$$

$$-\frac{d\,\lambda_{a}^{P,up}(P_{ck}^{up})}{d\,P_{ck}^{up}} + \hat{C_{ck}} - \sigma_{ck,1} + \sigma_{ck,2} + 2\sigma_{ck,7}(P_{ck}^{up} + P_{ck}^{da} + P_{ck}^{dw}) = 0\ \forall c \in C(a)\ k \in K \tag{4.109}$$

$$\frac{d\,\lambda_{a}^{P,dw}(P_{ck}^{dw})}{d\,P_{ck}^{dw}} - \check{C_{ck}} - \sigma_{c,4} + \sigma_{ck,3} + 2\sigma_{ck,7}(P_{ck}^{up} + P_{ck}^{da} + P_{ck}^{dw}) = 0\ \forall c \in C(a)\ k \in K \tag{4.110}$$

$$-\frac{d\,\lambda_a^{Q,up}(Q_{ck}^{up})}{d\,Q_{ck}^{up}} + \hat{C_{ck}} - \sigma_{ck,5} + 2\sigma_{ck,7}(Q_{ck}^{up} + Q_{ck}^{da} + Q_{ck}^{dw}) = 0\ \forall c \in C(a)\ k \in K \quad (4.111)$$

$$\frac{d\,\lambda_a^{P,dw}(Q_{ck}^{dw})}{d\,Q_{ck}^{dw}} + \sigma_{ck,6} + 2\sigma_{ck,7}(Q_{ck}^{up} + Q_{ck}^{da} + Q_{ck}^{dw}) = 0\ \forall c \in C(a)\ k \in K \quad (4.112)$$

$$(P_{ck}^{up} + P_{ck}^{da} + Q_{ck}^{dw})^2 + (Q_{ck}^{up} + Q_{ck}^{da} + Q_{ck}^{dw})^2 - S_{ck}^2 \leq 0\ \forall c \in C(a)\ k \in K \quad (4.113)$$

$$-P_{ck}^{up} \leq 0\ \forall c \in C(a)\ k \in K \quad (4.114)$$

$$P_{ck}^{dw} \leq 0\ \ \forall c \in C(a)\ k \in K \quad (4.115)$$

$$Q_{ck}^{dw} \leq 0\ \forall c \in C(a)\ k \in K \quad (4.116)$$

$$Q_{c}^{up} \geq 0\ \forall c \in C(a)\ k \in K \quad (4.117)$$

$$(P_{ck}^{up} - \hat{P_{ck}^{*}})\ \leq 0\ \forall c \in C(a)\ k \in K \quad (4.118)$$

$$-(P_{ck}^{dw} + \check{P_{ck}^{*}})\ \leq 0\ \forall c \in C(a)\ k \in K \quad (4.119)$$

$$\sigma_{ck,1}P_{ck}^{up} = 0\ \forall c \in C(a)\ k \in K \quad (4.120)$$

$$\sigma_{ck,3}P_{ck}^{dw} = 0\ \ \forall c \in C(a)\ k \in K \quad (4.121)$$

$$\sigma_{ck,2}(P_{ck}^{up} - \hat{P_{c}^{*}})\ = 0\ \forall c \in C(a)\ k \in K \quad (4.122)$$

$$\sigma_{ck,4}(P_{ck}^{dw} + \check{P_{ck}^{*}})\ = 0\ \forall c \in C(a)\ k \in K \quad (4.123)$$

$$\sigma_{ck,5}Q_{ck}^{up} = 0\ \forall c \in C(a)\ k \in K \quad (4.124)$$

$$\sigma_{ck,6}Q_{ck}^{dw} = 0\ \forall c \in C(a)\ k \in K \quad (4.125)$$

$$\sigma_{ck,7}[(P_{ck}^{up} + P_{ck}^{da} + P_{ck}^{dw})^2 + (Q_{ck}^{up} + Q_{ck}^{da} + Q_{ck}^{dw})^2 - S_{ck}^2] = 0\ \forall c \in C(a)\ k \in K \quad (4.126)$$

$$\sigma_{ck,j}, j = 1, ..., 7,\ \forall c \in C(a)\ k \in K \quad (4.127)$$

It is essential to point out that being a distribution network it is necessary to consider each phase of the system. For this reason, the $k$ subscripts have been added representing the quantities per phase. The optimization problem formalized above is nonlinear due to the KKT conditions and the capability constraints. The convex relaxation proposed in the previous Chapter decreases the number of nonlinearities in the problem replacing the non-convex power flow equations with the set of linear equations defined by the McCormick envelopes.

## 4.4 Case study application

The developed market structure has been tested to the modified benchmark distribution test network presented in Chapter 2.

To test the proposed balancing market approach, the following procedure has been followed: first of all, the active and reactive power was assigned to each node, then the SI was calculated for each network component. Consequently, having verified that the state of the network is unsatisfactory, the BM session is introduced to increase the overall quality of service. Please note that as mentioned above linear supply functions are considered. The nonlinear optimization problem has been solved on a PC Intel i7, 16 GB RAM using the YALMIP [56] interface through the MATLAB language with the IPOPT [57] solver.

### 4.4.1 Modified IEEE 13 Bus system

In the base case, without any BM session the overall $SI_{sys} = 22.0214$ with $SI^G = 2.9733$, $SI^V = 15.4269$ and $SI^C = 3.6211$. The run time is 85 seconds. To understand if the current condition is satisfactory or not, in Figures 4.2, 4.3 and 4.4 are reported the various SI for each phase of the network (node and link numbers are the same as the original network test case).

As can be easily seen from the aforementioned figures, many indices at each node and component are larger or close to the unit. This means that many components are close to or even exceed the limits that allow a good quality of service. For this reason, it is necessary to proceed with the BM session. After the optimization BM session with $\alpha = 1$, the overall $SI_{sys} = 20.3156$ with $SI^G = 2.5649$, $SI^V = 14.2292$ and $SI^C = 3.5215$, at the price 2932€.

The Figures show that the generator of node 634 belonging to the first aggregator is not called to provide a balancing service. This happens because its SI is almost unitary so it does not lead to an improvement in system performance. As far as the whole system is concerned, the overall quality of the service has generally improved, it should be noted that as far as the SI of lines and generators is concerned some of them have slightly increased or unchanged, while all the indicators of nodal voltage have improved.

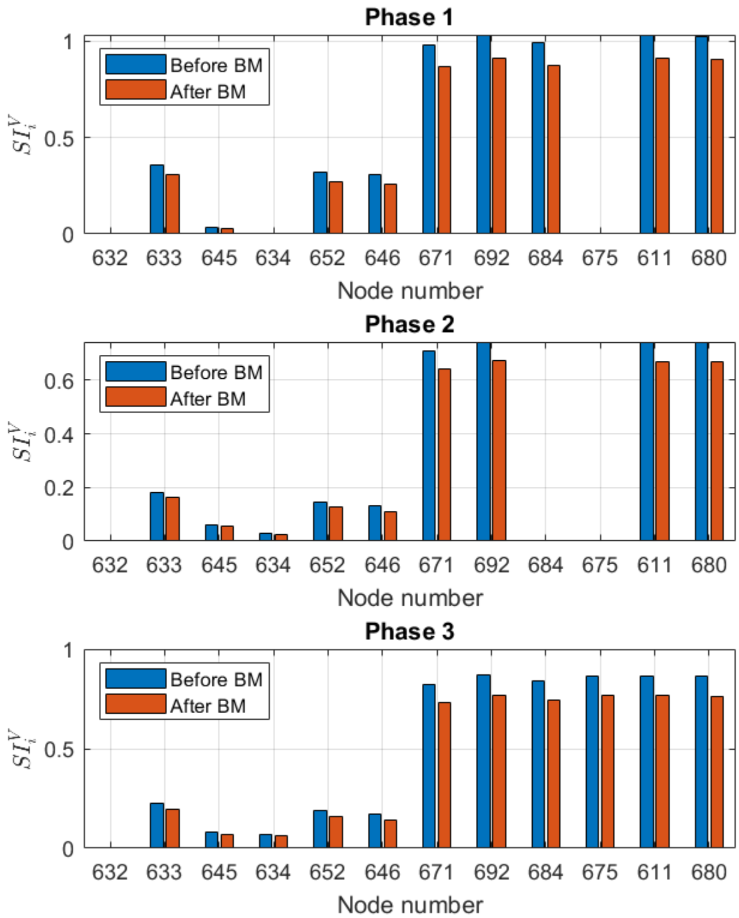


Figure 4.2: $SI_i^V$ before and after the BM session for the Modified IEEE 13 Bus system

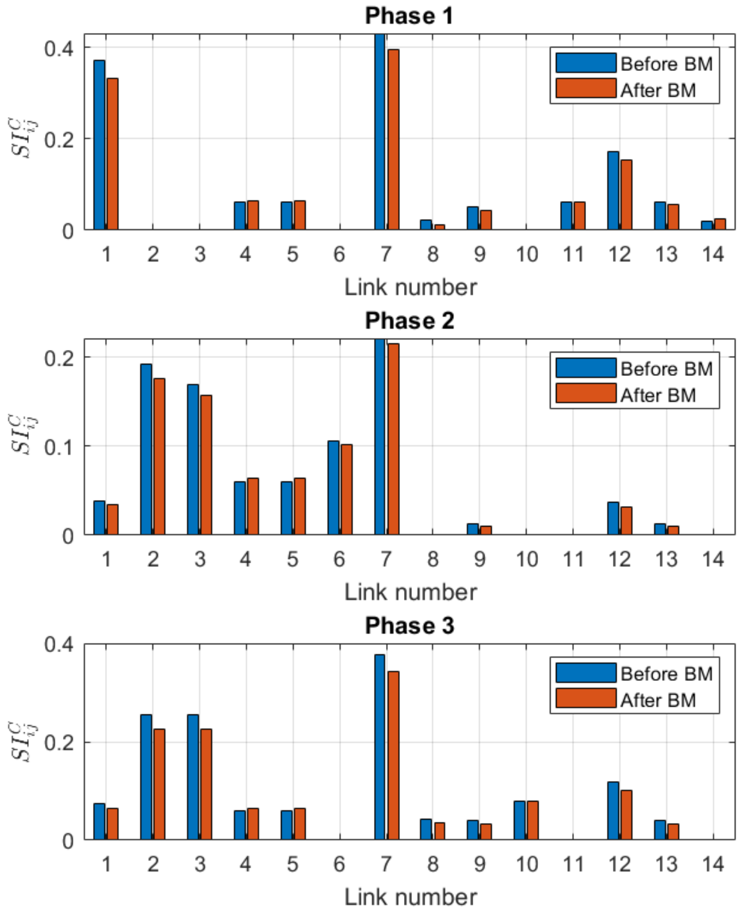


Figure 4.3: $SI_i^V$ before and after the BM session for the Modified IEEE 13 Bus system

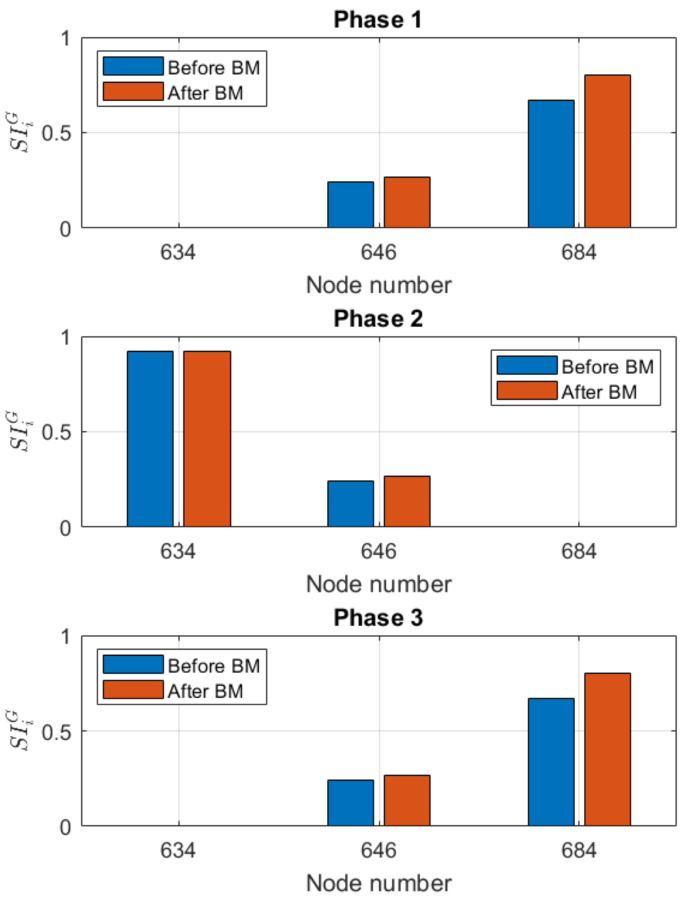


Figure 4.4: $SI_i^V$ before and after the BM session for the Modified IEEE 13 Bus system

#### 4.4.1.1 Sensitivity analysis

The formalized objective function includes basically two not commensurable terms (costs in [€ ] and SI as [p.u.]). The parameter $\alpha$ represents the weight that is given to SI (the higher its value the higher is the importance given to the minimization of the overall stress of the system ). Thus, a sensitivity analysis has been carried out to evaluate the influence of the SI in the cost function. Defining $J_1$ as the economic cost for purchasing balancing services from the market participants and $J_2$ as the overall SI cost, the sensitivity of the solution of the multiobjective problem at the variation of parameter $\alpha$ is presented in Figure 4.5 and Table 4.1.

| $\alpha$ | $J_1$ | $J_2$ |
|---|---|---|
| 0.01 | 2933.3 | 20.262 |
| 0.1 | 2933 | 20.3041 |
| 1 | 2932 | 20.31 |
| 10 | 2929.5 | 20.323 |
| 100 | 2926 | 20.36730 |

Table 4.1: Results of the sensitivity analysis to the parameter $\alpha$

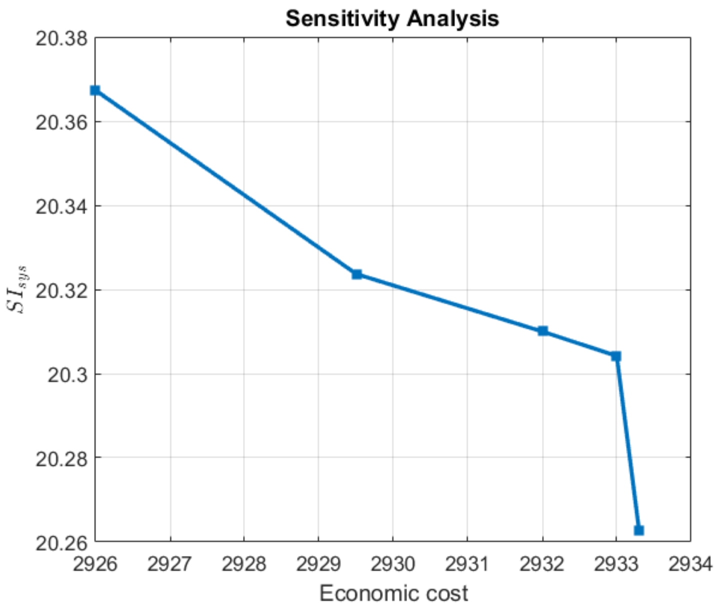


Figure 4.5: Sensitivity analysis to the $\alpha$ parameter

As can be seen from the simulation the two costs are mathematically in conflict, as a matter of fact changing the weight parameter both costs change oppositely (one increases and the other decreases).

### 4.4.2 Modified IEEE 123 Bus system

In the base case, without any BM session the $SI_{sys} = 233.35$ with $SI^G = 1.597$, $SI^V = 214.367789$ and $SI^C = 17.38604$. The run time for the optimization is 326.18 seconds. The comparison between the various indexes before and after the BM session is reported in Figures 4.8, 4.7, and 4.6 (node and link numbers are the same as the original network test ). Also, the global situation of the system improves $SI_{sys} = 204.3237$ with $SI^G = 1.824$, $SI^V = 194.2391$ and $SI^C = 8.2605$, at the price of 3213.42 €. From the figures, it can be seen that the SIs related to voltages and currents decrease while the generators' indices increase or remain unchanged since generation is required by the plants. Namely, the generators of nodes 43 and 44 are not called to provide balancing services. Also in this test case a counterexample was done to prove the correctness of the solution, i.e. the flexibility (in increasing and then in decreasing active power) has been set to the maximum value, in both cases the SI of the whole system increased.

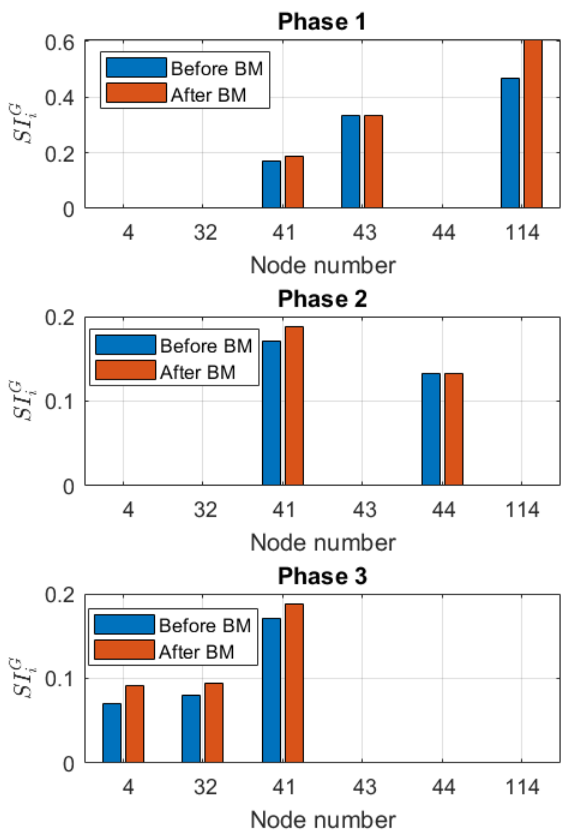


Figure 4.6: $SI_i^G$ before and after the BM session for the Modified IEEE 123 Bus system

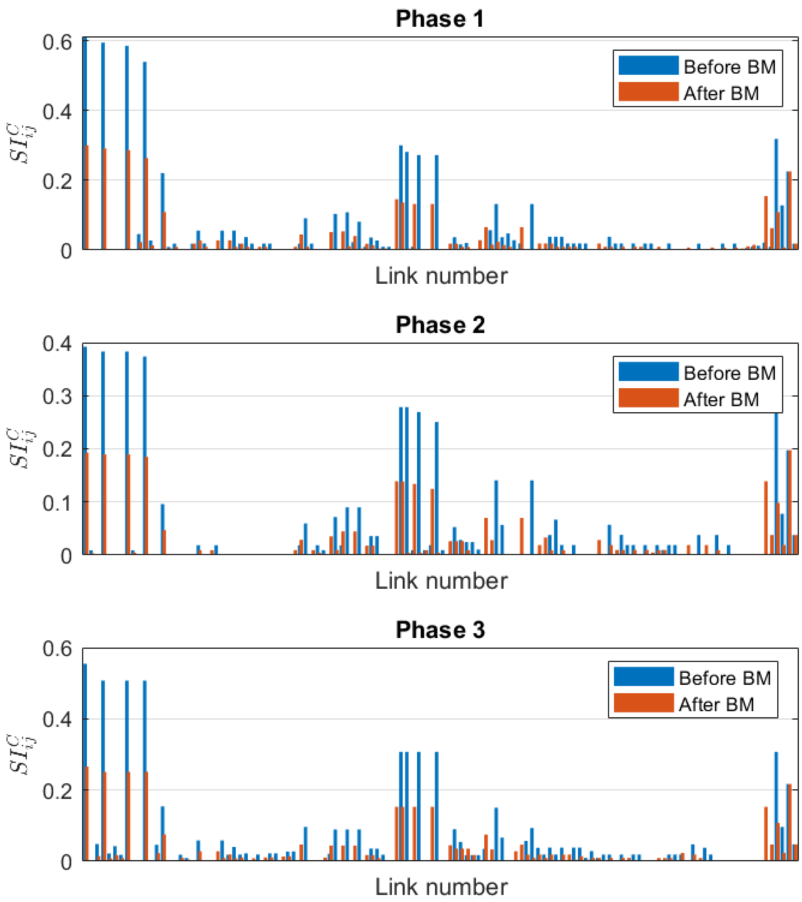


Figure 4.7: $SI_i^C$ before and after the BM session for the Modified IEEE 123 Bus system

In summary, it can be seen that in both cases the quality of service is visibly improved. This new market structure can, therefore, find applicability in real cases being an effective but also simple method it solves a single optimization problem without resorting to iterative mechanisms, whose convergence is not assured, as it happens in current literature.

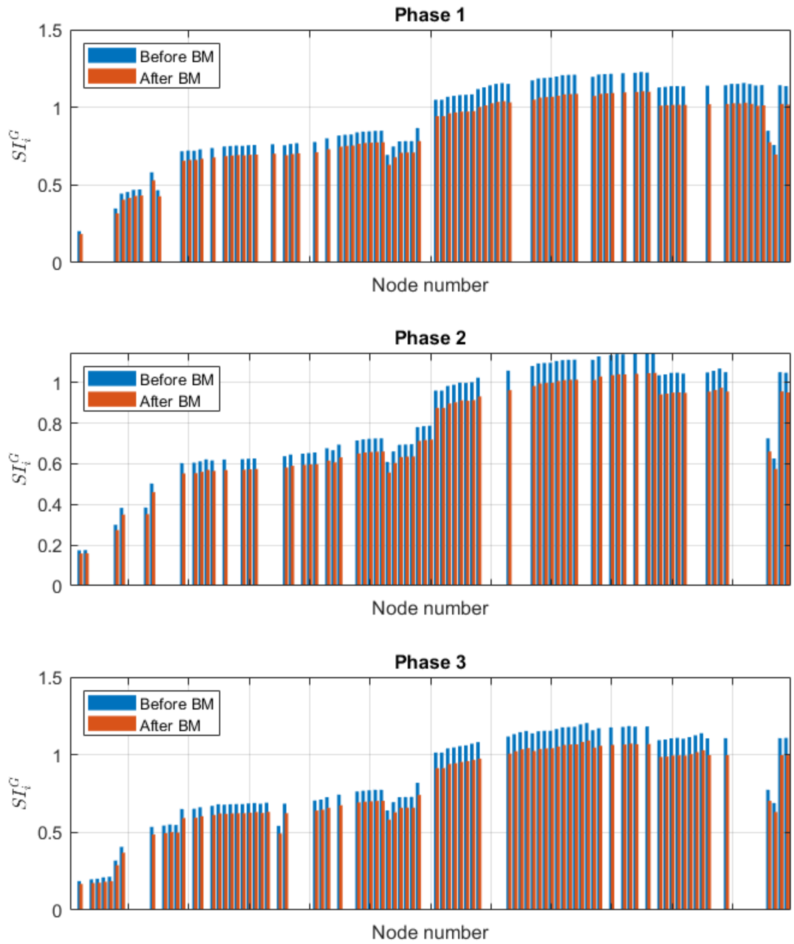


Figure 4.8: $SI_i^V$ before and after the BM session for the Modified IEEE 123 Bus system

# Chapter 5

# A new distributed structure for cooperative Balancing Markets in distribution grids

This Chapter aims to formulate a cooperative structure for the Balancing Market in distribution grids. In this framework, each market participant cooperates with the other actors in order to achieve the so called social welfare. To do this, a distributed optimization algorithm based on the cooperative consensus mechanism has been developed (i.e. the the Proximal Atomic Coordination (PAC) algorithm) . The activities presented in this Chapter have been developed during the research stage at the Active Adaptive Control (AAC) Lab at Massachusetts Institute of Technology (MIT) under the supervison of Dr. Anuradha Annaswamy. The main contribution of the following Chapter are:

- The development of a distributed algorithm, based on the PAC method, with clear articulation of conditions for proof of convergence and the advantages over the popular distributed alternating direction method of multipliers (dADMM) method.
- The application of the proposed method to the BM framework over the modified IEEE 13 and 123 bus systems dining a new cooperative market structure.

In particular, a new distributed algorithm called PAC is presented and discussed. It based on the prox-linear approach and we prove that it achieves convergence in both objective values and distance to feasibility with rate $o(1/\tau)$, where $\tau$ is the number of algorithmic iterations. This convergence rate matches the best known convergence results for distributed algorithms based on the popular alternating direction method of multipliers (ADMM) method. We then discuss the advantages of PAC using the notions of privacy and algorithmic complexity, the latter of which involves complexities in iteration, computation, communication and storage. The algorithm is then applied to create a cooperative Balancing Market structure, exploiting the decomposition profile, and then testing to the modified IEEE 13 and 123 bus systems. It is important to note that in this Chapter only a summary of the theoretical aspects of the algorithm are specified, for more information and plots concerning the convergence properties please see [58] The organization of this Chapter is as follows. Section I investigates the actual literature on the distributed optimization, in Section II the decomposition profile and the PAC algorithm are defined, In Section III PAC properties compared with dADMM are discussed, in Section IV the the cooperative PAC-based market structure is presented and tested over the test networks of Chapter 2 and Chapter 4. Finally in Appendix A the convergence proofs are shown.

## 5.1 Literature review

Optimization formulations feature prominently in many of today's most pressing engineering and data science challenges, including: machine learning [59]; statistics [60]; database queries [61]; power grid regulation [62]; network analysis [63]; and finite-element models [64]. These formulations, however, often involve a large number of variables to optimize over, intricate feasibility regions and complicated objective functions. In fact, it is the large-scale nature of the underlying applications that often render centralized approaches, by far the best understood and most widely implemented class of optimization solvers, to perform poorly.

In an effort to make these large-scale optimization problems more manageable, various distributed optimization approaches have been developed, including many that employ regularization [65], [66], [67], [68], [69], [70] , [71], [72], [73], [74] and [75] . These decentralized approaches solve the original centralized optimization formulation by

- Properly allocating the computational burden of optimizing the problem's variables amongst different processing units
- Exploiting parallel or sequential computation to speed up iteration convergence
- Ensuring that all processing units agree on coupling variables so that the distributed solution matches that of the original global solution.

More broadly, distributed optimization algorithms have been investigated extensively over the years, as presented in the excellent overviews and seminal results of [68], [67], [66], [72]. A key ingredient adopted in distributed optimization is that of dual decomposition, an innovation that allows for the optimization of the primal and dual variables in succession [66]. A popular method that enabled a distributed implementation of dual decomposition is the alternating direction method of multipliers (ADMM), introduced into the optimization literature in [67] wherein a distributed architecture is employed for the computation of primal variables with a centralized solver retained for the update of dual variables. More recently, a fully distributed implementation of ADMM (denoted as dADMM) has been carried out in [69] and [73, 74, 75]. Distributed optimization has also been addressed recently in [76] and [77]. The main trade off of these distributed methods with their centralized relatives is generally higher iteration complexity due to the added responsibility of coordinating the coupling variables [68]. In an effort to decrease this complexity, many methods such as dADMM [69] and the prox-linear algorithm [70] utilize additional coordination variables that are sent along with the minimally required primal and/or dual variable information. This solution, however, adds communication complexity (the number of scalar variables that need to be sent via a communication medium [78]) to the algorithms since now more information has to be transferred between each processing unit. Additionally, the inclusion of coordination variables further adds to each processing unit's primal update computational complexity and associated storage complexity, as well as exposes private information in competitive economic environments. Finally, for these regularized methods, we also have that each primal update presents itself as a sub-minimization of the regularized Lagrangian which necessitates an added computational complexity to solve. In summary, a successful distributed algorithm is one that appropriately balances minimizing all four of these complexities (iteration, computational, communication, and storage). These four complexities, taken together, comprise a given distributed algorithm's algorithmic complexity.

As mentioned earlier, there has been an extensive amount of work done in the framework of distributed optimization solvers and algorithms. The interested reader can consult the excellent overviews ([65] [68]) and the seminal works ([71] [72]) for a more thorough description of the different distributed approaches that have been used to optimize large-scale systems. Most of these methods are based on regularized centralized approaches, such as the method of multipliers (MoM) [79], or the proximal point method ([80][81]). One of the most successful distributed methods is the dADMM method [69], which is based on the Gauss-Seidel inspired centralized ADMM [67]. The dADMM method specifically presents itself as a 2-block ADMM formulation and was shown to exhibit the aforementioned $O\left(1/\tau\right)$ and $O\left(\sqrt{\kappa_f}\log\left(\frac{1}{\epsilon}\right)\right)$ convergence characteristics, matching the best convergence rates for the 2-block centralized ADMM case [82] and comparing favorably with the $O\left(\left(\kappa_f\right)^{\frac{5}{7}}\log\left(\frac{1}{\epsilon}\right)\right)$ convergence rate of the distributed Nesterov gradient descent algorithm in [83].

## 5.2 Proximal Atomic Cordination algorithm

### 5.2.1 A Global Standard Optimization Framework

A vector $y$ is denoted as a column vector and $y_i$ denotes the $i-th$ component of the vector. Overbar $\overline{y}$ and underbar $\underline{y}$ notation denote upper and lower limits for a variable $y$. For a matrix $G$, we write $[G]_i$ to denote the $i-th$ column of matrix $G$, $[G]^i$ to denote the $i-th$ row of matrix $G$, and $G_{ij}$ to denote the entry at $i-th$ row and $j-th$ column. For matrices we let $\lambda_{\min}(G)$, $\hat{\lambda}_{\min}(G)$ and $\lambda_{\max}(G)$ represents it smallest, smallest non-zero and largest eigenvalue, respectively. For a set $Y$, $|Y|$ denotes the number of elements of $Y$. In addition, we denote the elements of a subset $W_{[j]} \subset Y = \{w_k^j\}$, where the subscript $k$ indicates that $w_k^{[j]}$ is the $k$th element of the subset $W_j$ and the superscript $j$ signifies that $w_k^j$ belongs to the $j$th subset.

We start by considering the following global standard optimization (GSO) problem:

$$\begin{array}{c} \min\limits_{y\in\mathbb{R}^Y} \left\{ f(y) \triangleq \sum_{k\in B} f_k(y) \right\} \\ \text{subj. to: } \{\ Gy = 0_M\ \} \end{array} \tag{5.1}$$

where:

- $y \in \mathbb{R}^Y$ represents the $Y$ optimization variables;
- $f_k : \mathbb{R}^Y \to \mathbb{R}$ represents an objective function summand;
- $\sum_{k\in B} f_k(y)$ represents the total objective function;
- $G = \left[g_1^T, ..., g_M^T\right] \in \mathbb{R}^{M\times Y}$ is a matrix s.t. null$(G)$ represents the feasibility region of the GSO and each row $g_i \in \mathbb{R}^Y$ represents the feasibility constraints[1].

We can decompose the GSO into $K$ different coupled sub-optimization problems, called *atoms*, by using the following decomposition profile $\mathcal{D} = (\mathsf{L}, \mathsf{C}, \mathsf{S}, \mathsf{O}, \mathsf{T})$:

- $\mathsf{L} = \{L_j, \quad \forall j \in K\}$ represents the partition of $Y$ with $L_j \subseteq N$ being the index of the components of $y$ that each $j$th atom "owns";
- $\mathsf{C} = \{C_j, \quad \forall j \in K\}$ represents the partition of $M$ with $C_j \subseteq M$ representing the index of the rows of $G$ that each $j$th atom "owns";

[1] It should be pointed out that a general formulation of an optimization problem also includes inequality constraints $hy \leq 0$ that represent the domain of optimization variables. The discussions here can be easily extended to this case as well by suitably restricting the variables' feasible set.

- $\mathsf{S} = \{S_j, \forall j \in K\}$ represents the partition of $B$ with $S_j \subseteq B$ representing the index of the objective summands of $f$ that each $j$th atom "owns";

- $\mathsf{O} = \{O_j \quad \forall j \in K\}$ has each $O_j \subseteq Y$ representing the index of the "copies" of variables of $y$ that each $j$th atom needs to satisfy the *scope* of both $[G_{C_j}]_i$ and $f_{S_j}(y) \triangleq \sum_{k \in S_j} f_k(y)$;

- $\mathsf{T} = \{T_j, \quad \forall j \in K\}$ represents the atomic partitioning of the GSO where each $T_j = L_j \cup O_j \subseteq Y$ represents the total variables each $j$th atom "uses".

Using $\mathcal{D}$, we obtain the following *atomized standard optimization (ASO)*:

$$\begin{aligned} &\min_{a_j \in \mathbb{R}^{|T_j|}} \left\{ \textstyle\sum_{j \in K} \tilde{f}_j(a_j) \right\} \\ &\text{subj. to:} \begin{cases} \tilde{G}_j a_j = 0_{N_j^{\mathsf{C}}}, \text{ for all } j \in K \\ [B]^j a = 0_{N_j^{\mathsf{O}}}, \text{ for all } j \in K \end{cases}, \end{aligned} \tag{5.2}$$

with each atom's variables (both owned and copied) being represented by $a_j \in \mathbb{R}^{|T_j|}$ and:

- $\tilde{f}_j(a_j) = \sum_{k \in |S_j|} f_k\left(\Pi_{T_j} y\right)$ for all $k \in |S_j|$ and $y \in \mathbb{R}^Y$, where $\Pi_{T_j} \in R^{Y \times Y}$ is given by:

  $$\Pi_{T_j} \triangleq \text{diag}\left\{\delta_{1 \in T_j}, \ldots, \delta_{Y \in T_j}\right\},$$

  with:

  $$\delta_{n \in T_j} \triangleq \begin{cases} 1, \text{if } n \in T_j \\ 0, \text{otherwise} \end{cases}.$$

- $\tilde{G}_j \triangleq G_{C_j, T_j} \in \mathbb{R}^{|C_j| \times |T_j|}$ represents the submatrix of $G_{C_j,-}$ obtained by removing all zero columns;

- $B \in \mathbb{R}^{|\mathsf{O}| \times |T|}$ represents the adjacency matrix of the directed graph with nodes signifying the atomic variables with edges $\left(y_k^{[i]}, y_k^{[j]}\right)$, where $y_k^{[i]}$ with $k \in L_i$ is defined to be the variable of $y$ that is "owned" by the $j$th atom and $y_k^{[j]}$ with $k \in O_j$ defined to be the variable of $y$ that is "copied" by the $j$th atom;

- $[B]^j \triangleq B_{\tilde{O}_j,-} \in \mathbb{R}^{|O_j| \times |T|}$ represents the relevant incoming edges of the directed graph with adjacency matrix $B$ of the variables "copied" by the $j$th atom, with $\tilde{O}_j = |O_j| + \sum_{k \in [j-1]} |O_k|$;

- $[B]_j \triangleq B_{-,\tilde{T}_j} \in \mathbb{R}^{|\mathsf{O}| \times |T_j|}$ represents the relevant out-going edges of the directed graph with adjacency matrix $B$ of the variables "owned" by the $j$th atom, with $\tilde{T}_j = |T_j| + \sum_{k \in [j-1]} |T_k|$.

We notice how the ASO of (5.2) is an augmented version of the GSO of (5.1), sharing the same optimal solutions. Specifically, we relate the optimal ASO $a^*$ to optimal GSO $y^*$ via:

$$y^* = \Pi^{\mathsf{L}} a^*,$$

where each $\Pi^{\mathsf{L}} : \mathbb{R}^Y \rightarrow \mathbb{R}^{|T|}$ represents the projection from $a$-space into $y$-space that consists of all owned variables in each $a_{j^*}$ being placed correctly in the resultant $y^{*\prime} = \Pi^{\mathsf{L}} a^*$.

As we will see in the next section, we prefer to use the ASO formulation to solve the underlying problem because it allows us to compute the primal update in a distributed fashion, i.e., by using the ASO we can iterate over all $a_j$ concurrently. The cost of this parallelization, however, is the addition of another constraint, termed *coordination*, that needs to be satisfied for every atom $j \in K$:

$$[B]^j\, a = 0_{N_j^{\circ}}.$$

This coordination constraint can be interpreted as requiring all atomic copied variables in a given $j$th atom (given by $O_j$) to equal the value of their corresponding owned variable in another non $j$-atom.

#### 5.2.1.1 Decomposition Example

In this example the application of the decomposition profile $\mathcal{D}$ over a simple optimization problem is shown. Let's take as an example the following linear problem concerning the optimization of power flows on a simple inductive power grid. This problem is a classic transmission network problem, that for the sake of simplicity has been adapted to a 3-node network shown in Figure 5.1.

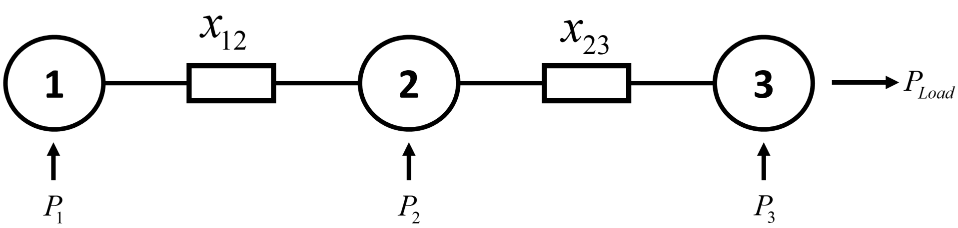


Figure 5.1: Example of a 3 node transmission network

The problem statement is as follows:

$$\min \{C_1 P_1 + C_2 P_2 + C_3 P_3\} \tag{5.3}$$

$$P_{12} - P_1 = 0 \tag{5.4}$$

$$P_{12} + P_2 - P_{23} = 0 \tag{5.5}$$

$$P_{23} + P_3 = P_{Load} \tag{5.6}$$

$$P_{12} = \frac{\delta_1 - \delta_2}{x_{12}} \tag{5.7}$$

$$P_{23} = \frac{\delta_2 - \delta_3}{x_{23}} \tag{5.8}$$

where $P_1$, $P_2$ and $P_3$ are active power injections at each node whose cost is $C_1$, $C_2$, $C_3$ respectively; $P_{12}$ and $P_{23}$ are the active power flows between the network nodes; $x_{12}$ and $x_{23}$ are the cables' inductance and $\delta_1$, $\delta_2$ and $\delta_3$ are the phase associated to each node. Constraints (5.4)-(5.6) define the bower balance at each node and (5.7)-(5.8) represent the so called DC power flow equations commonly used for transmission grid studies. We then collect (5.4)-(5.8) to rewrite the constraints in standard form $Gy = b$ where:

$$y = \begin{bmatrix} P_1 & P_2 & P_3 & \delta_1 & \delta_2 & \delta_3 & P_{12} & P_{23} \end{bmatrix}^T \tag{5.9}$$

$$G = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -\frac{1}{x_{12}} & \frac{1}{x_{12}} & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -\frac{1}{x_{23}} & \frac{1}{x_{23}} & 0 & 1 \end{bmatrix} \tag{5.10}$$

$$b = \begin{bmatrix} 0 & 0 & P_{Load} & 0 & 0 \end{bmatrix}^T \tag{5.11}$$

We then decompose the problem in 3 subproblems applying $\mathcal{D}$

$$\mathsf{L} = \{(1,4,5),(2,7),(3,6,8)\} \tag{5.12}$$

$$\mathsf{C} = \{(1,4),(2),(3,5)\} \tag{5.13}$$

$$\mathsf{S} = \{(1),(2),(3)\} \tag{5.14}$$

$$\mathsf{O} = \{(7),(8),(5)\} \tag{5.15}$$

$$\mathsf{T} = \{(1,4,5,7),(2,7,8),(3,6,8,5)\} \tag{5.16}$$

It is important to point out to the reader that, as described in the previous section, $\mathcal{D}$ defines a set of indexes that represents rows/columns of the matrices of the optimization problem in standard form. Moreover $\mathcal{D}$ is not unique, but depends on how many subproblems has to be formulated. In the following the coordination matrix $B$ is defined

$$B = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{5.17}$$

Using the decomposition profile $\mathcal{D}$ and the definition from the matrix $B$ it is now possible to separate the problem into 3 agents (atoms) on which the algorithm can then be applied. Below the 3 sub-problems are reported

$$\begin{aligned} &\min \{C_1 P_1\} \\ &\tilde{G}_1 a_1 = b_1 \end{aligned} \tag{5.18}$$

with

$$a_1 = \begin{bmatrix} P_1 & \delta_1 & \delta_2 & \hat{P_{12}} \end{bmatrix}^T \tag{5.19}$$

$$\tilde{G}_1 = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 0 & -\frac{1}{x_{12}} & \frac{1}{x_{12}} & 1 \end{bmatrix} \tag{5.20}$$

$$b_1 = \begin{bmatrix} 0 & 0 \end{bmatrix}^T \tag{5.21}$$

$$\begin{aligned} &\min \{C_2 P_2\} \\ &\tilde{G}_2 a_2 = b_2 \end{aligned} \tag{5.22}$$

with

$$a_2 = \begin{bmatrix} P_2 & P_{12} & \hat{P_{23}} \end{bmatrix}^T \tag{5.23}$$

$$\tilde{G}_2 = \begin{bmatrix} -1 & 1 & -1 \end{bmatrix} \tag{5.24}$$

$$b_2 = \begin{bmatrix} 0 \end{bmatrix}^T \tag{5.25}$$

$$\begin{aligned} &\min \{C_3 P_3\} \\ &\tilde{G}_3 a_3 = b_3 \end{aligned} \tag{5.26}$$

with

$$a_3 = \begin{bmatrix} P_3 & \delta_3 & P_{23} & \hat{\delta_2} \end{bmatrix}^T \tag{5.27}$$

$$\tilde{G}_3 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & -\frac{1}{x_{23}} & \frac{1}{x_{23}} & 1 \end{bmatrix} \tag{5.28}$$

$$b_3 = \begin{bmatrix} P_{Load} & 0 \end{bmatrix}^T \tag{5.29}$$

### 5.2.2 PAC Algorithm Specifications

We begin by forming the atomic Lagrangian function for (5.2):

$$\mathcal{L}\left(a,\mu,\nu\right)=\sum_{j\in K}\left[\hat{f}_j\left(a_j\right)+\mu_j^T\tilde{G}_ja_j+\nu_j^T\left[B\right]^j a\right]$$
$$=\sum_{j\in K}\left[\hat{f}_j\left(a_j\right)+\mu_j^T\tilde{G}_ja_j+\nu^T\left[B\right]_j a_j\right]\triangleq\sum_{j\in K}\mathcal{L}_j\left(a_j,\mu_j,\nu\right). \tag{5.30}$$

We can then apply the prox-linear approach of [70] to (5.30) and obtain the *proximal atomic coordination (PAC)* algorithm:

$$a_j\left[\tau+1\right]=\underset{a_j\in\mathbb{R}^{|T_j|}}{\operatorname{argmin}}\left\{\mathcal{L}_j\left(a_j,\bar{\mu}_j\left[\tau\right],\bar{\nu}\left[\tau\right]\right)+\frac{1}{2\rho}\left\|a_j-a_j\left[\tau\right]\right\|_2^2\right\}, \tag{5.31}$$
$$\mu_j\left[\tau+1\right]=\mu_j\left[\tau\right]+\rho\gamma_j\tilde{G}_ja_j\left[\tau+1\right], \tag{5.32}$$
$$\bar{\mu}_j\left[\tau+1\right]=\mu_j\left[\tau+1\right]+\rho\hat{\gamma}_j[\tau+1]\tilde{G}_ja_j\left[\tau+1\right], \tag{5.33}$$
$$\text{Communicate } a_j\forall j\in[K] \text{ within network,} \tag{5.34}$$
$$\nu_j\left[\tau+1\right]=\nu_j\left[\tau\right]+\rho\gamma_j\left[B\right]^j a\left[\tau+1\right], \tag{5.35}$$
$$\bar{\nu}_j\left[\tau+1\right]=\nu_j\left[\tau+1\right]+\rho\hat{\gamma}_j[\tau+1]\left[B\right]^j a\left[\tau+1\right], \tag{5.36}$$
$$\text{Communicate } \bar{\nu}_j\forall j\in[K] \text{ within network,} \tag{5.37}$$

where we utilized the previously mentioned prox-linear method to ensure parallel computation of each primal step. We have chosen a common step-size ($\rho\geq 0$) but different over-relaxation terms $\gamma_j$ and $\hat{\gamma}_j[\tau]$. This constant step-size can be pre-assigned to all agents in the network.

The primal and dual variables are initialized as follows, for $j\in K$:

$$a_j\left[0\right]\in\mathbb{R}^{|T_j|}, \tag{5.38}$$
$$\mu_j\left[0\right]=\rho\gamma_j\tilde{G}_ja_j\left[0\right], \tag{5.39}$$
$$\nu_j\left[0\right]=\rho\gamma_j\left[B\right]^j a\left[0\right], \tag{5.40}$$
$$\bar{\nu}_j\left[0\right]=\nu_j\left[0\right]+\rho\hat{\gamma}_j[0]\left[B\right]^j a\left[0\right]. \tag{5.41}$$

We note the symmetry in the structure between the update equations for $(\mu_j,\bar{\mu}_j)$ and $(\nu_j,\bar{\nu}_j)$. This structure resembles primal updates for accelerated gradient methods [84], in the sense that we "accelerate" our dual variable for our atomic primal updates ($\overline{\mu}_j$ and $\overline{\nu}_j$) while still updating the "true" dual variable ($\mu_j$ and $\nu_j$) in the usual way.

### 5.2.3 Structural Assumptions

We next make the following assumptions on the structure of the GSO and ASO formulations (and components therein), where $\gamma_{min} \triangleq \max_{j \in K} \{\gamma_j\}$:

**Assumption 5.2.1.** *Each $\hat{f}_k \in \left\{ \tilde{f}_j \right\} \forall j \in K$ is a closed, convex and proper (CCP) function with dom $\left( \hat{f}_k \right) = \mathbb{R}^Y$.*

**Assumption 5.2.2.** *There exists a non-trivial optimal GSO solution $y^* \in \mathbb{R}^N$. The optimal ASO solution $a^* \in \mathbb{R}^{|T|}$ is related to $y^* \in \mathbb{R}^N$ via: $y^* = \Pi^{\mathsf{L}} a^*$.*

**Assumption 5.2.3.** *Let the PAC parameters satisfy:*

$$1 > \rho^2 \gamma_{min} \lambda_{max} \left( \tilde{G}^T \tilde{G} + B^T B \right)$$

**Assumption 5.2.4.** *Let the PAC parameters satisfy:*

$$1 \geq \rho^2 \gamma_{min} \lambda_{max} \left( \tilde{G}^T \tilde{G} + B^T B \right)$$

**Assumption 5.2.5.** *Each $\hat{f}_k \in \left\{ \tilde{f}_j \right\} \forall j \in K$ is differentiable, $\alpha$-strongly convex and L-strongly smooth [85].*

**Assumption 5.2.6.** *The over-relaxation terms $\gamma_j$ and $\hat{\gamma}_j[\tau]$ are chosen so that the following inequality is satisfied for all $\tau$ and $j \in [K]$:*

$$\gamma_j > \hat{\gamma}_j[\tau] > \hat{\gamma}_j^{min} > 0$$

*where $\hat{\gamma}_j^{min}$ is a positive constant.*

Assumptions 5.2.1 and 5.2.2 are standard assumptions for optimization problems [85]. Assumptions 5.2.3 and 5.2.4 are needed to guarantee that metric by which we measure convergence for our PAC algorithm is p.d. or p.s.d, respectively. Assumption 5.2.5 is a more stringent assumption for when we have atomic objective functions that are sufficiently “curved” along every atomic variable dimension. Indeed, this latter assumption generally allows for quicker convergence since the curvature of the functions allow for quick primal convergence to the minimum cost solution. Assumption 4.6 is needed to obtain improved privacy, and is discussed in Proposition 5.1.

### 5.2.4 Main Result

Using the above GSO and ASO optimization models, we now state the main result of the chapter (proofs are reported in Appendix B):

**Theorem 5.2.7.** *Let the PAC parameters satisfy $\rho > 0$ and $\gamma_j > 0$ for $1 \leq j \leq K$. Further, let:*

$$a[\tau] = [a_1[\tau]; s; a_K[\tau]],$$

*represent the PAC trajectory of (5.31)-(5.37) under zero-initialization. Then:*

a) *Strictly Sublinear Convergence: If Assumptions 5.2.1, 5.2.2 and 5.2.3 holds, then there exists an optimal ASO solution $a^*$ s.t.:*

$$\lim_{\tau\to\infty} \{a[\tau]\} = a^*,$$

*with convergence rate satisfying, for all $\tau \in \mathbb{N}$:*

$$\left\{\|a[\tau+1] - a[\tau]\|^2_{\tilde{V}_2(\rho,\Gamma)}\right\} \forall\tau \in \mathbb{N} = o\left(\frac{1}{\tau}\right).$$

b) *Ergodic Sublinear Convergence: If Assumptions 5.2.1, 5.2.2 and 5.2.4 holds, then there exists an optimal ASO solution $a^*$ s.t.:*

$$\lim_{\tau\to\infty} \{a[\tau]\} = a^*$$

*s.t.:*

$$\left|\hat{f}(\bar{a}[\tau]) - \hat{f}(a^*)\right| \leq \left(\frac{1}{\tau}\right) \xi^{-1}_{E1}(\rho, \gamma_{min}), \tag{5.42}$$

$$\left\|R^\Gamma \bar{a}[\tau]\right\|^2_2 \leq \left(\frac{1}{\tau}\right) \xi^{-1}_{E2}(\rho, \gamma_{min}), \tag{5.43}$$

*where:*

$$\gamma_{min} = \min_{j\in K} \{\gamma_j\},$$

$$\xi^{-1}_{E1}(\rho, \gamma_{min}) = \frac{2}{\rho}\left[\|a^*\|^2_2 + \frac{4\Phi^2}{\gamma_{min}\hat{\lambda}_{min}(V_1)}\right], \tag{5.44}$$

$$\xi^{-1}_{E2}(\rho, \gamma_{min}) = \frac{2}{\rho^2}\left[\|a^*\|^2_2 + 4\rho^2 + \frac{4\Phi^2}{\gamma_{min}\hat{\lambda}_{min}(V_1)}\right],$$

*with $\Phi \triangleq \sup_{z\in\partial f(a^*)} \{\|z\|_2\}$ and $R^\Gamma$ satisfying:*

$$R^\Gamma R^\Gamma = \tilde{V}_1(\Gamma),^2$$

*and $\bar{a}[\tau] = \frac{1}{\tau}\sum_{v\in[\tau]} a[v]$ being the ergodic average.*

[2] The *root* $R$ exists since $V_1$ is square p.s.d. matrix. Specifically, if $V_1 = U\Sigma U^T$ is a suitable eigenvalue decomposition, then $R = \left[U\Sigma^{\frac{1}{2}}U^T\right]$

c) *Linear Convergence: If Assumptions 5.2.1, 5.2.2, 5.2.4 and 5.2.5 holds, then there exists a unique optimal ASO solution s.t., for all $\tau \in \mathbb{N}$:*

$$\lim_{\tau \to \infty} \{a[\tau]\} = a^*$$

*and:*

$$\|a[\tau] - a^*\| \leq [1 + \xi_P(\rho, \gamma_{min})]^{-\tau} \left( \|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + \|r^*\|_2^2 \right),$$

*where:*

$$\xi_P(\rho, \gamma_{min}) = \frac{2\alpha\rho\gamma_{min}\tilde{\lambda}_{min}(V_1)}{2\alpha L + 2\rho^{-2} + \gamma_{min}\tilde{\lambda}_{min}(V_1)}, \tag{5.45}$$

*with $r^* \in \mathbb{R}^{|T|}$ satisfying:*

$$R^\Gamma r^* + \frac{1}{\rho}\nabla_a \hat{f}(a) = 0.$$

Theorem 5.2.7 details the conditions under which the convergence of PAC is strictly sublinear, sublinear and linear. Essentially, under the regularity conditions of Assumptions 5.2.3 and 5.2.4, we have at least a sublinear convergence in the atomic primal variables so long as we have a convex optimization problem. If we strengthen each atomic objective function $\tilde{f}_j$ to be differentiable, $\alpha$-strongly convex and $L$-strongly smooth, then we can improve this convergence rate to being linear.

We maximize PAC convergence by choosing the following $\rho$ and $\gamma_{min}$ as specified in Theorem 5.2.8 (proofs are reported in Appendix B):

**Theorem 5.2.8.** *If we denote the atomic spectral ratio (ASR) of a particular decomposition profile $\mathcal{D}$ by:*

$$\psi(\mathcal{D}) = \hat{\lambda}_{min}(V_1)(\lambda_{max}(V_1))^{-1}, \tag{5.46}$$

*then the choices of PAC parameters $\rho$ and $\gamma_{min}$ that maximize the corresponding PAC convergence rates are:*

a) *Ergodic Rate: Optimizing (5.44) over $(\rho, \gamma_{min})$:*

$$\rho^*_{E1} = \left(\sqrt{\gamma^*_{E1}\lambda_{max}(V_1)}\right)^{-1}, \tag{5.47}$$

$$\gamma^*_{E1} = \frac{4\Phi^2}{\hat{\lambda}_{min}(V_1)\|a^*\|_2^2}, \tag{5.48}$$

$$\xi^*_{E1} = \frac{\sqrt{\psi(\mathcal{D})}}{2\Phi\|a^*\|_2}. \tag{5.49}$$

*b) Strong Primal Rate: Optimizing (5.45) over $(\rho, \gamma_{min})$:*

$$\rho_P^* = \left(\sqrt{\gamma_P^* \lambda_{max}(V_1)}\right)^{-1}, \tag{5.50}$$

$$\gamma_P^* = \frac{2\alpha L}{2\lambda_{max}(V_1) + \hat{\lambda}_{min}(V_1)}, \tag{5.51}$$

$$\xi_P^* = \left(\frac{1}{2\sqrt{\alpha^{-1}L}}\right)\sqrt{\frac{2\hat{\lambda}_{min}(V_1)\,\psi(\mathcal{D})}{2\lambda_{max}(V_1) + \hat{\lambda}_{min}(V_1)}}. \tag{5.52}$$

### 5.2.5 PAC Privacy Framework

In competitive economic environments, a successful distributed algorithm also needs to preserve privacy and limit the dissemination of any sensitive information between processing units. In the current context, we consider sensitive information to correspond to any information about a processing unit's computations that can be used by others to their advantage. In this chapter, we limit the objective of rogue processing units to be that of using this sensitive information in order to sabotage the global convergence properties of the overall distributed solver, in contrast to other adversarial scenarios. Like the decreasing of iteration complexity through the introduction of regularized coordination variables discussed previously, the limiting of sensitive information does introduce additional algorithmic complexity that will need to be appropriately balanced. Protection of any sensitive information corresponds to privacy. An advantage of distributed optimization algorithms could be to keep private the coordination between the agents. In the current context, we consider sensitive information to correspond to any information about $k$-th atom that $j$-th atom can make use of to their advantage. In particular, a rogue $j$-th atom may seek to utilize information about $k$-th atom to recover the information about the trajectory of $\nu_j$. Such information may be utilized to sabotage the overall convergence property of the PAC through deliberate manipulations of the coordination constraints.

The privacy framework is illustrated by Figure 5.2. The "true" valuation of the coordination constraint is given by the *blue* circles and the $\nu_i^{[j\to k]}[\tau]$ points. This valuation is computed and kept private by $k$-th atom. The "protected" valuation of the coordination constraint is given by the *black* triangles and is communicated from $k$-th atom to $j$-th atom. In this configuration, atom-$j$ only sees the *black* trajectory of the "protected" valuation and not the *blue* trajectory of the "true" valuation. More mathematical details concerning the privacy of PAC and its comparison with dADMM will be dealt with in the following sections.

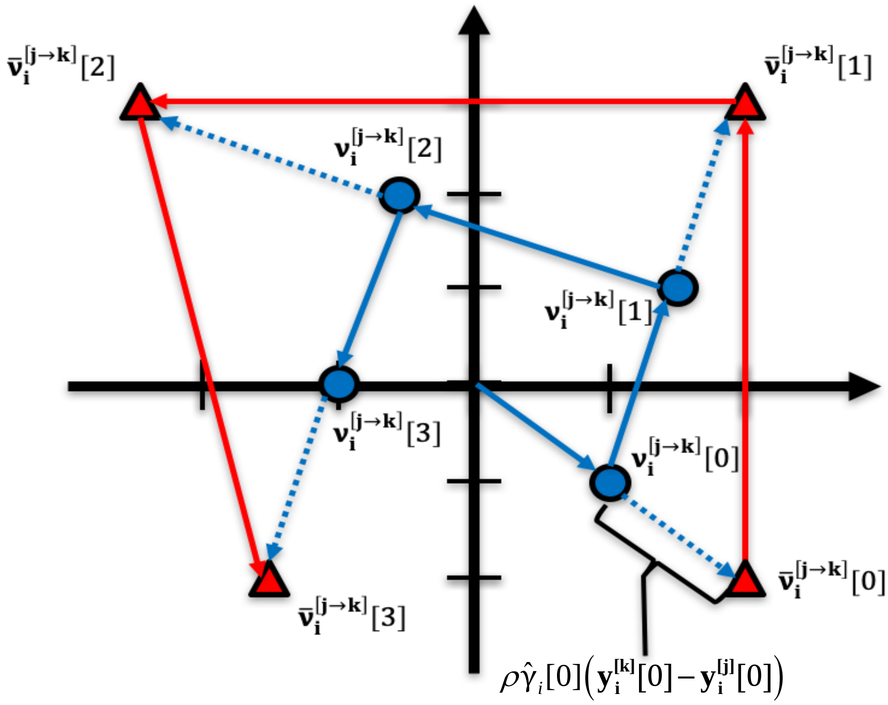


Figure 5.2: Example of the dual variable trajectory for $\left\{\nu_i^{[j\to k]}[\tau]\right\}_{\tau=0}^{3}$ and $\left\{\bar{\nu}_i^{[j\to k]}[\tau]\right\}_{\tau=0}^{3}$.

## 5.3 Advantages of PAC over dADMM

In this section, we carry out a detailed comparison between PAC and dADMM in terms of privacy features and algorithmic complexities. While we focus our comparisons on the dADMM algorithm in [69], the overall observations extend to other variations such as those in [73, 74, 75] as well.

### 5.3.1 Derivation of dADMM

To begin, we adapt dADMM from [69] to our general ASO formulation of (5.2) to obtain the following algorithm:

$$a_j[\tau+1] = \underset{a_j \in \mathbb{R}^{|T_j|}}{\operatorname{argmin}} \left\{ \mathcal{L}_j\left(a_j, \mu_j[\tau], \nu[\tau]\right) + \frac{\rho}{2}\left\|\tilde{G}_j a_j\right\|_2^2 + \frac{\rho}{2}\left\|[B]_j\left(a_j - a_j[\tau]\right) + w[\tau]\right\|_2^2 \right\}, \tag{5.53}$$

$$\mu_j[\tau+1] = \mu_j[\tau] + \rho\gamma\left(\tilde{G}_j a_j[\tau+1]\right), \tag{5.54}$$

$$\text{Communicate } \{a_j[\tau+1]\}\, \forall j \in [K] \text{ within network,} \tag{5.55}$$

$$w_j[\tau+1] = \left(\frac{1}{1+d_j}\right)[B]^j\, a[\tau+1], \tag{5.56}$$

$$\nu_j[\tau+1] = \nu_j[\tau] + \rho w_j[\tau+1] \tag{5.57}$$

$$\text{Communicate } \left\{w_j[\tau+1], \nu_{j[\tau+1]}\right\} \forall j \in [K] \text{ within network,} \tag{5.58}$$

where $d_j \triangleq |\{i \in K|\, L_i \cap O_j \neq \emptyset\}|$ in (5.56) represents the *in-degree* of atom-$j$ w.r.t. dependencies on other atoms (e.g., how many other atoms have "owned" indices that the $j$th atom needs a local "copy" of).

### 5.3.2 Privacy of PAC

This section aims to show that, with respect to dADMM, PAC keeps the $\nu_j$ trajectory, associated to the coordination between the atoms, private at each iteration. In what follows, we will mathematically show that the PAC algorithm prevents such a recovery. We define $\hat{\nu}_{j\to k}[\tau]$ with $k \neq j$ as the estimation of the $\nu_k$ trajectory at iteration $\tau$ by a rogue $j$-th atom. A repeated application of (5.35) from iteration 0 to iteration $\tau$ leads to:

$$\nu_j[\tau] = \rho\gamma_j\,[B]^j\left[\sum_{s=0}^{\tau} a[s]\right]. \tag{5.59}$$

Similarly, from (5.36), we obtain

$$\bar{\nu}_j[\tau] = \rho\gamma_j \left[B\right]^j \left[\sum_{s=0}^{\tau} a[s]\right] + \rho\hat{\gamma}_j[\tau] \left[B\right]^j a.[\tau] \tag{5.60}$$

Using (5.59), we propose that the rogue atom's estimation is carried out as

$$\hat{\nu}_{j\to k}\left[\tau\right] = \rho\widehat{\gamma_k \left[B\right]^k} \left[\sum_{s=0}^{\tau} a[s]\right], \tag{5.61}$$

where $\widehat{\gamma_k \left[B\right]^k}$ is an estimate of $\gamma_k \left[B\right]^k$. It can be seen that such an estimation is possible at every atom $j \neq k$ since the information on the right hand side of (5.61) is globally broadcasted.

**Proposition 5.3.1.** *Given the algorithm (5.31)-(5.37) with initialization as in (5.38)-(5.41), when the rogue agent estimation is carried out as in (28), then*

$$\|\hat{\nu}_{j\to k}\left[\tau\right] - \nu_j[\tau]\|_2 \not\to 0. \tag{5.62}$$

In contrast to the PAC, it is easy to see from (20)-(25) that dADMM does not preserve privacy as $a_j$, $w_j$ and $\nu_j$ are broadcast to everyone.

### 5.3.3 Algorithmic Complexity

In this section we compare the various complexities required at each atom for both the PAC algorithm of (5.31)-(5.37) and the dADMM algorithm of (5.53)-(5.58). Assume for illustrative purposes that each atomic objective function is a convex quadratic of the form:

$$\tilde{f}_j\left(a_j\right) = a_j^T \tilde{Q}_j a_j + \tilde{b}_j^T a_j + \tilde{c}_j, \tag{5.63}$$

where $\tilde{Q}_j \succeq 0 \in \mathbb{R}^{|T_j|\times|T_j|}$ and $\tilde{b}_j, \tilde{c}_j \in \mathbb{R}^{|T_j|}$. Then, we have the following complexities:

- **Iteration Complexity**: As we show in Theorem 5.2.8, PAC exhibits strictly sublinear convergence $o\left(\frac{1}{\tau}\right)$ while, as [69] demonstrates, dADMM exhibits sublinear convergence $O\left(\frac{1}{\tau}\right)$. The dADMM method specifically presents itself as a 2-block ADMM formulation and was shown to exhibit $O\left(1/\tau\right)$ and $O\left(\sqrt{\kappa_f}\log\left(\frac{1}{\epsilon}\right)\right)$ convergence characteristics, where $\kappa_f$ is the condition number, leading to an $\epsilon$-optimal solution[82, 83].

- **Computational Complexity**: Using the first-order condition for optimality, we have that the primal PAC updates of (5.31) can be written in closed-form as:

$$a_j\left[\tau+1\right] = \left(\tilde{Q}_j + \frac{1}{\rho}I_{N_j^{\mathrm{T}}}\right)^{-1}\left(\frac{1}{\rho}a_j\left[\tau\right] - \tilde{G}_j^T\tilde{\mu}_j\left[\tau\right] - \left[B\right]_j^T\tilde{\nu}\left[\tau\right] - \tilde{b}_j\right), \tag{5.64}$$

and that the primal dADMM updates of (5.53) can be written in closed-form as:

$$\left(\tilde{Q}_j + \rho\tilde{G}_j^T\tilde{G}_j + \rho\left[B\right]_j^T\left[B\right]_j\right)a_j\left[\tau+1\right] =$$
$$\left(\rho\left[B\right]_j^T\left[B\right]_j a_j\left[\tau\right] - \tilde{G}_j^T\mu_j\left[\tau\right] - \left[B\right]_j^T\nu\left[\tau\right] - \rho\left[B\right]_j^T w\left[\tau\right] - \tilde{b}_j\right) \quad (5.65)$$

We note that the matrix inversion in (5.64) only needs to be done once and can be considered a pre-processing step, so the atomic update complexity of PAC is only that of matrix multiplication: $O\left((|T_j|)^2\right)$. The update complexity of dADMM, on the other hand, necessitates solving a linear system of equations and so takes $O\left((|T_j|)^3\right)$ operations.[3] Hence we see that the PAC has superior update complexity over that of dADMM.

- **Communication Complexity**: For PAC, each atom is required to communicate primal $a_j$ and dual $\tilde{\nu}_j$, leading to a communication complexity of $O\left(|T_j| + |O_j|\right)$. In comparison, for dADMM, we have that each atom is required to communicate primal $a_j$ and dual $\{\nu, w_j\}$, leading to a communication complexity of $O\left(|T_j| + |O_j|\right)$. Hence we see that PAC has the same communication complexity over that of dADMM.

- **Storage Complexity**: For PAC, each atom is required to locally store and maintain primal $a_j$ and dual $\{\mu_j, \nu_j, \tilde{\nu}_j\}$,[4] for a storage requirement of $|T_j| + |C_j| + 2|O_j|$. In comparison, for dADMM, we have that each atom is required to locally store and maintain primal $a_j$ and dual $\{\mu_j, \nu_j, w_j\}$, for a storage requirement of $|T_j| + |C_j| + 2|O_j|$. Hence we see that PAC and dADMM have the same storage complexity requirements.

[3] We cannot invert the LHS matrix of (5.3.3) like we can in (5.64) since the LHS matrix is only guaranteed to be p.s.d., not p.d.

[4] We can equivalent state $\tilde{\mu}\left[\tau\right]$ in (5.31) as $\mu_j\left[\tau\right] + \rho\gamma\tilde{G}_j a_j\left[\tau\right]$.

## 5.4 Cooperative Balancing market framework

This section will show how the distributed algorithm described in the previous section can be transformed into a cooperative market mechanism. In this framework, unlike the approach described in Chapter 4, each market participant is not selfish but rather participates together with others to achieve the so-called social optimum. In mathematical terms, this concept translates into the union of optimization problems related to each agent, to create a unique one without any DSO dominance over the other market players. This concept is depicted in Figure 5.3, as a matter of fact using the PAC algorithm each market player shares some variables to reach the convergence to the optimal point while keeping the coordination information private.

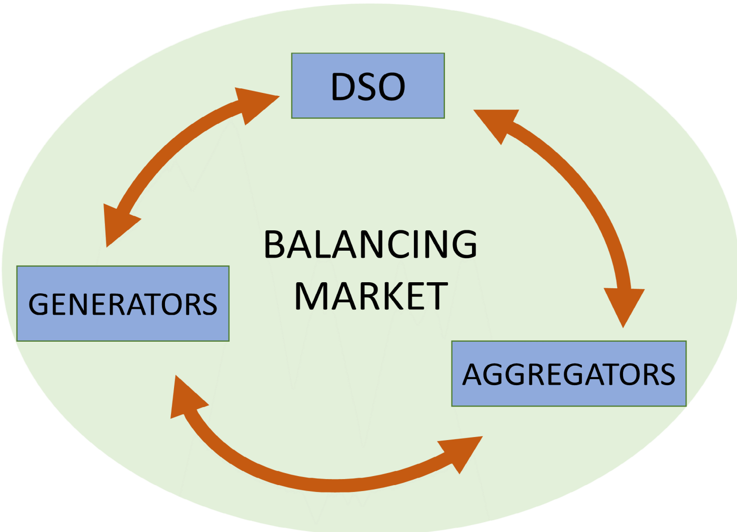


Figure 5.3: Cooperative market structure.

### 5.4.1 The overall Optimization problem

In this subsection the cooperative optimization problem is defined which is different from the problems formalized in Chapters 2 and Chapter 4, considering the new distribution grid model described in Chapter 3 and the equilibrium conditions of market participants stated in the previous section. The overall optimization problem aims to minimize the social welfare objective function

$$\min_{x_g \in GT, x_c c \in C(a), a \in A, V^R_{ik}, V^I_{ik}, I^R_{ik}, I^I_{ik}} (2.1) + (2.7) + (2.16) \tag{5.66}$$

Under constraints

$$SI_{sys} = \sum_{h \in (GT \cup C_{tot})} SI_h^G + \sum_{i \in N} SI_i^V + \sum_{i \in N} \sum_{\substack{j \in N \\ i \neq j}} SI_{ij}^C \tag{5.67}$$

$$SI_h^G = \frac{P_{gk}^{da} + P_{gk}^{up} + P_{gk}^{dw}}{P_{gk}^{\max}} \quad \forall h \in (GT \cup C_{TOT}) \tag{5.68}$$

$$SI_{ik}^V = \left( \frac{\|V_{ik}\| - \hat{V}_{ik}}{\Delta V_i^*} \right)^2 \forall i \in N \;\; k \in K \tag{5.69}$$

$$SI_{ijk}^C = \left( \frac{I_{ijk}}{I_{ij}^*} \right)^2 \forall i, j \in N \;\; k \in K \tag{5.70}$$

$$P_{ik} = \sum_{\substack{g \in GT: \\ i = Node(g)}} \left( P_{gk}^{da} + P_{gk}^{up} + P_{gk}^{dw} \right) + \sum_{a \in A} \sum_{\substack{c \in C(a): \\ i = Node(c)}} \left( P_{ck}^{da} + P_{ck}^{up} + P_{ck}^{dw} \right) \;\; \forall i \in N \;\; k \in K \tag{5.71}$$

$$Q_{ik} = \sum_{\substack{g \in GT: \\ i = Node(g)}} \left( Q_{gk}^{da} + Q_{gk}^{up} + Q_{gk}^{dw} \right) + \sum_{a \in A} \sum_{\substack{c \in C(a): \\ i = Node(c)}} \left( Q_{ck}^{da} + Q_{ck}^{up} + Q_{ck}^{dw} \right) \;\; \forall i \in N \;\; k \in K \tag{5.72}$$

$$I^R = \mathrm{Re}\left( \mathbf{Y} V \right) \tag{5.73}$$

$$I^I = \mathrm{Im}\left( \mathbf{Y} V \right) \tag{5.74}$$

$$P_{ik} = \; a_{ik} + b_{ik} \;\; \forall i \in N, \;\; k \in K \tag{5.75}$$

$$Q_{ik} = -c_{ik} + d_{ik} \;\; \forall i \in N, \;\; k \in K \tag{5.76}$$

$$a_{ik} \geq \underline{V_{ik}}^R I_{ik}^R + V_{ik}{}^R \underline{I}_{ik}^R - \underline{V_{ik}}^R \underline{I}_{ik}^R \;\; \forall i \in N, \;\; k \in K \tag{5.77}$$

$$a_{ik} \geq \overline{V_{ik}}^R I_{ik}^R + V_{ik}{}^R \overline{I}_{ik}^R - \overline{V_{ik}}^R \overline{I}_{ik}^R \;\; \forall i \in N, \;\; k \in K \tag{5.78}$$

$$a_{ik} \leq \underline{V_{ik}}^R I_{ik}^R + V_{ik}{}^R \overline{I}_{ik}^R - \overline{V_{ik}}^R \underline{I}_{ik}^R \;\; \forall i \in N, \;\; k \in K \tag{5.79}$$

$$a_{ik} \leq \overline{V_{ik}}^R I_{ik}^R + V_{ik}{}^R \underline{I}_{ik}^R - \underline{V_{ik}}^R \overline{I}_{ik}^R \;\; \forall i \in N, \;\; k \in K \tag{5.80}$$

$$b_{ik} \geq \underline{V_{ik}}^I I_{ik}^I + V_{ik}{}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \underline{I}_{ik}^R \;\; \forall i \in N, \;\; k \in K \tag{5.81}$$

$$b_{ik} \geq \overline{V_{ik}}^I I_{ik}^I + V_{ik}{}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \overline{I_{ik}}^I \;\; \forall i \in N, \;\; k \in K \tag{5.82}$$

$$b_{ik} \leq \underline{V_{ik}}^I I_{ik}^I + V_{ik}{}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \underline{I_{ik}}^I \;\; \forall i \in N, \;\; k \in K \tag{5.83}$$

$$b_{ik} \leq \overline{V_{ik}}^I I_{ik}^I + V_{ik}{}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \overline{I_{ik}}^I \;\; \forall i \in N, \;\; k \in K \tag{5.84}$$

$$c_{ik} \geq \underline{V_{ik}}^R I_{ik}^I + V_{ik}{}^R \underline{I_{ik}}^I - \underline{V_{ik}}^R \underline{I_{ik}}^I \;\; \forall i \in N, \;\; k \in K \tag{5.85}$$

$$c_{ik} \geq \overline{V_{ik}}^R I_{ik}^I + V_{ik}{}^R \overline{I_{ik}}^I - \overline{V_{ik}}^R \overline{I_{ik}}^I \;\; \forall i \in N, \;\; k \in K \tag{5.86}$$

$$c_{ik} \leq \underline{V_{ik}}^R I_{ik}^I + V_{ik}^R \overline{I_{ik}}^I - \overline{V_{ik}}^R \underline{I_{ik}}^I \ \ \forall i \in N, \ k \in K \tag{5.87}$$

$$c_{ik} \leq \overline{V_{ik}}^R I_{ik}^I + V_{ik}^R \underline{I_{ik}}^I - \underline{V_{ik}}^R \overline{I_{ik}}^I \ \ \forall i \in N, \ k \in K \tag{5.88}$$

$$d_{ik} \geq \underline{V_{ik}}^I I_{ik}^R + V_{ik}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \underline{I_{ik}}^I \ \ \forall i \in N, \ k \in K \tag{5.89}$$

$$d_{ik} \geq \overline{V_{ik}}^I I_{ik}^R + V_{ik}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \overline{I_{ik}}^I \ \ \forall i \in N, \ k \in K \tag{5.90}$$

$$d_{ik} \leq \underline{V_{ik}}^I I_{ik}^R + V_{ik}^I \overline{I_{ik}}^I - \overline{V_{ik}}^I \underline{I_{ik}}^I \ \ \forall i \in N, \ k \in K \tag{5.91}$$

$$d_{ik} \leq \overline{V_{ik}}^I I_{ik}^R + V_{ik}^I \underline{I_{ik}}^I - \underline{V_{ik}}^I \overline{I_{ik}}^I \ \ \forall i \in N, \ k \in K \tag{5.92}$$

$$\underline{V_{ik}}^R \leq I_{ik}^R \leq \overline{V_{ik}}^R \ \ \forall i \in N \ k \in K \tag{5.93}$$

$$\underline{V_{ik}}^I \leq I_{ik}^I \leq \overline{V_{ik}}^I \ \ \forall i \in N \ k \in K \tag{5.94}$$

$$\underline{I_{ik}}^R \leq I_{ik}^R \leq \overline{I_{ik}}^R \ \ \forall i \in N \ k \in K \tag{5.95}$$

$$\underline{I_{ik}}^I \leq I_{ik}^I \leq \overline{I_{ik}}^I \ \ \forall i \in N \ k \in K \tag{5.96}$$

$$\underline{P}_{ik} \leq P_{ik} \leq \overline{P}_{ik} \ \ \forall i \in N \ k \in K \tag{5.97}$$

$$\underline{Q}_{ik} \leq Q_{ik} \leq \overline{Q}_{ik} \ \ \forall i \in N, \ k \in K \tag{5.98}$$

$$P_{gk}^{up} \geq 0 \ \forall G \in GT, \ k \in K \tag{5.99}$$

$$P_{gk}^{dw} \leq 0 \ \forall G \in GT, \ k \in K \tag{5.100}$$

$$(P_{gk}^{up} - \hat{P_{gk}^{*}}) \ \leq 0 \ \forall G \in GT, \ k \in K \tag{5.101}$$

$$-(P_{gk}^{dw} - \check{P_{gk}^{*}}) \ \leq 0 \ \forall G \in GT, \ k \in K \tag{5.102}$$

$$(P_{gk}^{up} + P_{gk}^{da} + P_{gk}^{dw})^2 + (Q_{gk}^{up} + Q_{gk}^{da} + Q_{gk}^{dw})^2 - S_{gk}^2 \leq 0 \ \forall G \in GT, \ k \in K \tag{5.103}$$

$$-P_{ck}^{up} \leq 0 \ \forall c \in C(a) \ k \in K \tag{5.104}$$

$$P_{ck}^{dw} \leq 0 \ \ \forall c \in C(a) \ k \in K \tag{5.105}$$

$$Q_{ck}^{dw} \leq 0 \ \forall c \in C(a) \ k \in K \tag{5.106}$$

$$Q_{c}^{up} \geq 0 \ \forall c \in C(a) \ k \in K \tag{5.107}$$

$$(P_{ck}^{up} - \hat{P_{ck}^{*}}) \ \leq 0 \ \forall c \in C(a) \ k \in K \tag{5.108}$$

$$-(P_{ck}^{dw} + \check{P_{ck}^{*}}) \ \leq 0 \ \forall c \in C(a) \ k \in K \tag{5.109}$$

$$(P_{ck}^{up} + P_{ck}^{da} + Q_{ck}^{dw})^2 + (Q_{ck}^{up} + Q_{ck}^{da} + Q_{ck}^{dw})^2 - S_{ck}^2 \leq 0 \ \forall c \in C(a) \ k \in K \tag{5.110}$$

We then use the PAC algorithm to solve the BM problem of (5.67)-(5.110), with the social welfare objective function. All simulations were performed using a 2.3 GHz Intel Core i7 with MATLAB [86] and the YALMIP interface [28] using the GUROBI solver [27]. In order to be able to use PAC it is necessary to define the new vector of decision variables increased by the variables needed to reach consensus. In this respect, in order to limit the number of sub-problems, it was decided to define a problem for each node of the network without separating the phases. The new vector of decision variables is given by:

$$a_i = \begin{bmatrix} I_i^R & I_i^I & V_i^R & V_i^I & P_g^{up} & P_g^{dw} & Q_g^{up} & Q_g^{dw} & P_c^{up} & P_c^{dw} & Q_c^{up} & Q_c^{dw} & a_i & b_i & c_i & d_i & \hat{V}_j^R & \hat{V}_j^I \end{bmatrix}^T$$
$$\forall i, j \in N \quad \forall G \in GT \quad \forall c \in C(a) \qquad (5.111)$$

where $\hat{V}_j^R$ and $\hat{V}_j^I$ are the only coordination variables needed for this optimization problem, as a matter of fact are needed for the nodal analysis equations to couple the system through the power flow and the branch current can be expressed as a function of the voltages between two nodes.

### 5.4.2 Application to the modified IEEE 13 Bus system

For this simulations the chosen PAC parameters are $\gamma = 1.931$ and $\rho = 0.2873$. The comparison between the various indexes before and after the BM session is reported in Figures 5.4, 5.5 and 5.6. Also in this case the global situation of the system improves ( $SI_{sys} = 21.7979$ with $SI^G = 3.432$, $SI^V = 14.9751$ and $SI^C = 3.3908$, at the price of 838.54 €). From the figures, it can be seen that the SIs related to voltages and currents decrease while the generators' indices are called to increase (node 684) and decrease (node 646) their power production. The generator of node 634 is not called to provide balancing services.

Further analysis has been made on the performance of the PAC for this particular application. In Figure 5.7, we evaluate the algorithm's performances as measured by two different metrics. These are the distance to global feasibility: $\left\|\tilde{\mathbf{G}}\mathbf{a}\left[\tau\right]\right\|_2$; and the distance to coordination: $\left\|\mathbf{B}\mathbf{a}\left[\tau\right]\right\|_2$. These metrics describe how feasible the solution is (whether power flow equations are satisfied), and how far the variable copies are from their true values. Results show the algorithm converges to the optimal solution in less than 2000 iterations. The maximum and average time required for each primal update for PAC is 0.0096s and 0.0023s, respectively.

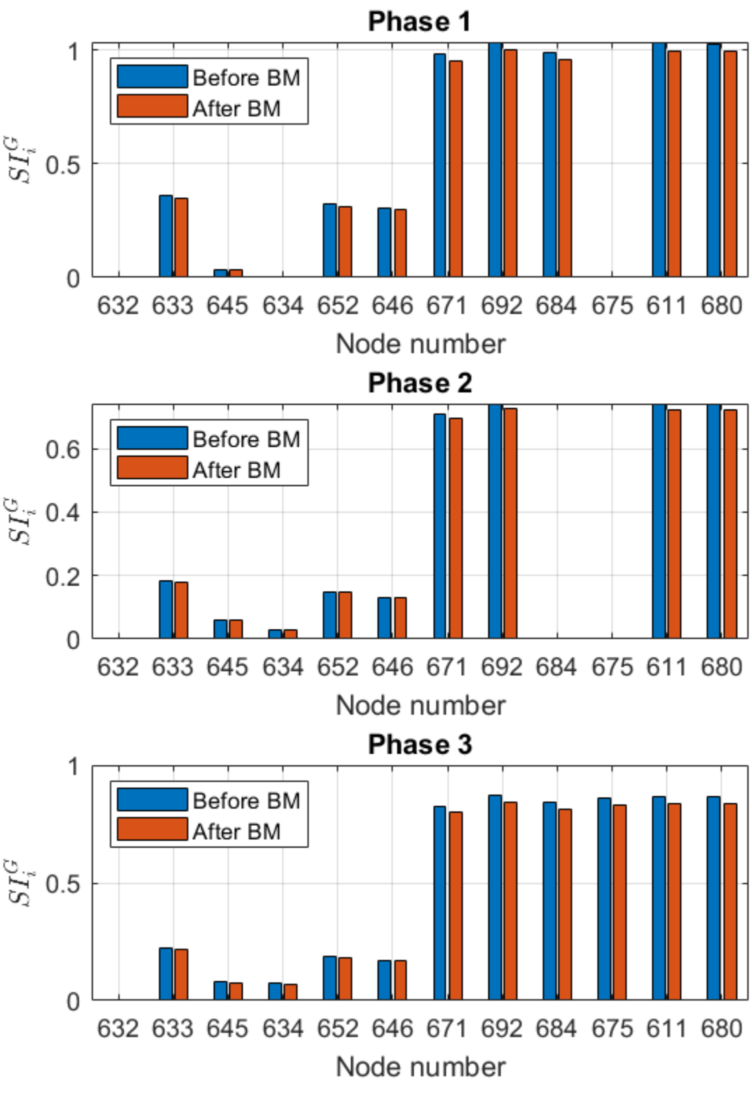


Figure 5.4: $SI_i^V$ before and after the BM session for the Modified IEEE 13 Bus system

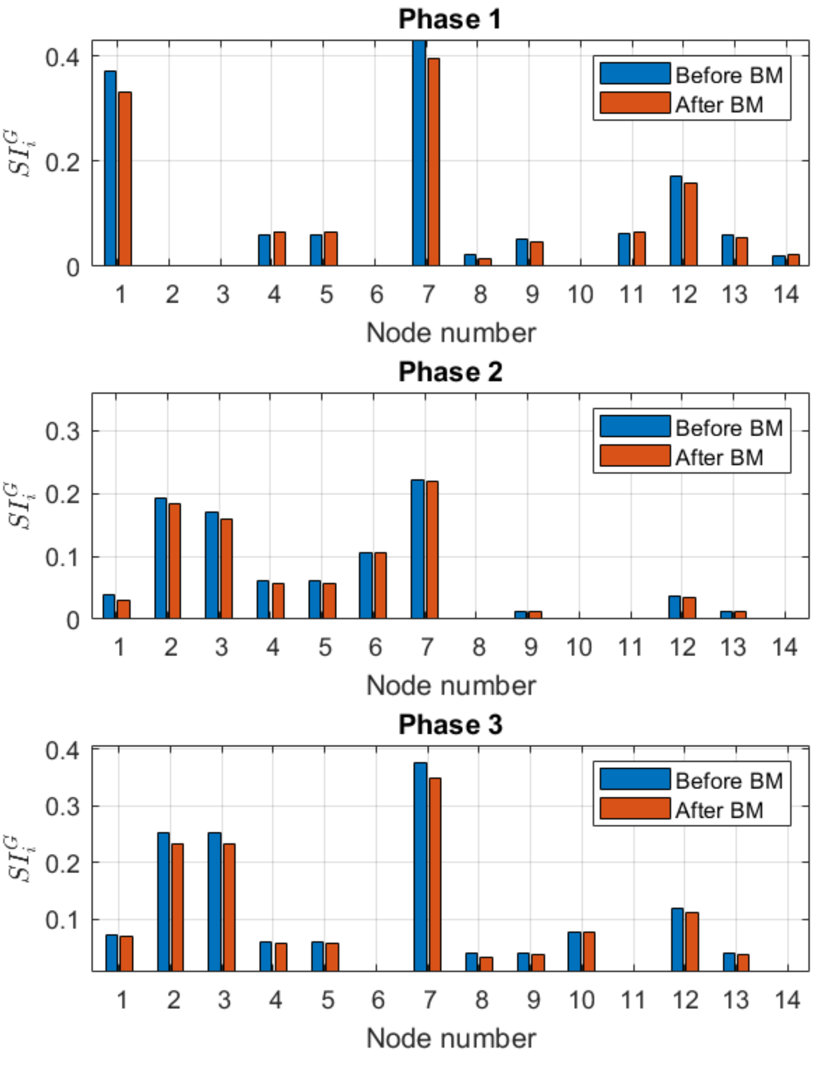


Figure 5.5: $SI_i^C$ before and after the BM session for the Modified IEEE 13 Bus system

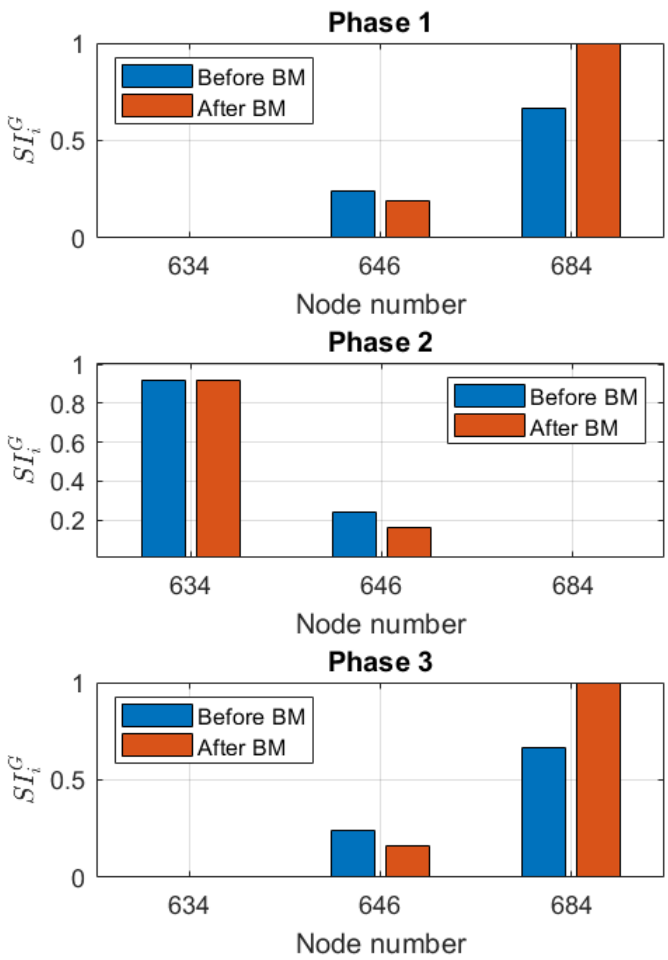


Figure 5.6: $SI_i^G$ before and after the BM session for the Modified IEEE 13 Bus system

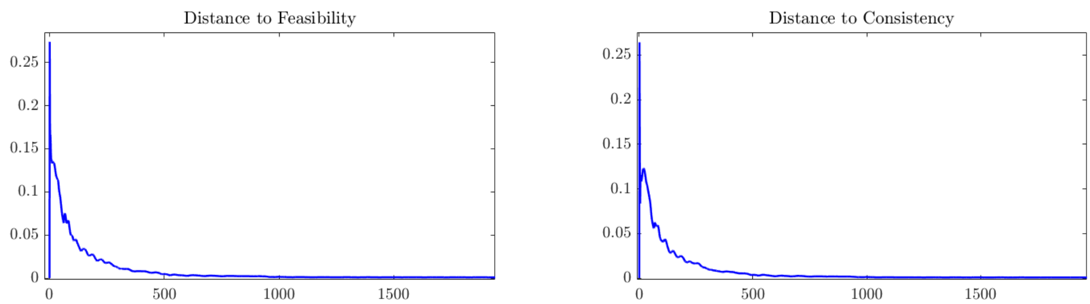


Figure 5.7: Algorithms performances.

### 5.4.3 Application to the modified IEEE 123 Bus system

In this second market simulation, the chosen PAC parameters are $\gamma = 2.031$ and $\rho = 0.585$. Numerical experiment show the for this particular case PAC converges to the optimal solution in less than 3500 iterations. The maximum and average time required for each primal update for PAC is 0.0105s and 0.0054s, respectively. For this last simulation (with $\alpha = 1$) the overall SI is decreased to 219.351 with $SI^G = 2.2845$, $SI^V = 208.57$ and $SI^C = 8.49$, at the price of 1409.67 €).

From a first analysis, the competitive structure compared to the cooperative one seems to be more effective in terms of improving the quality of the distribution network service. It must be taken into account that the two decision-making problems are different and the two market structures have different properties, from an architectural and security point of view the distributed structure is more efficient.

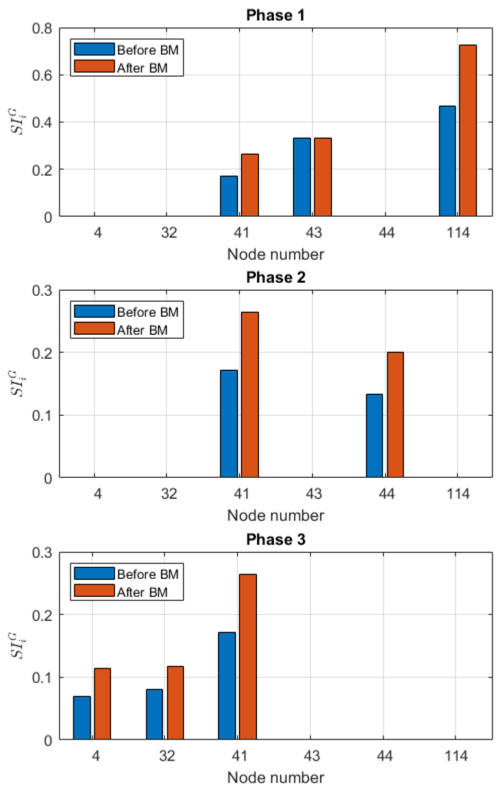


Figure 5.8: $SI_i^G$ before and after the BM session for the Modified IEEE 123 Bus system

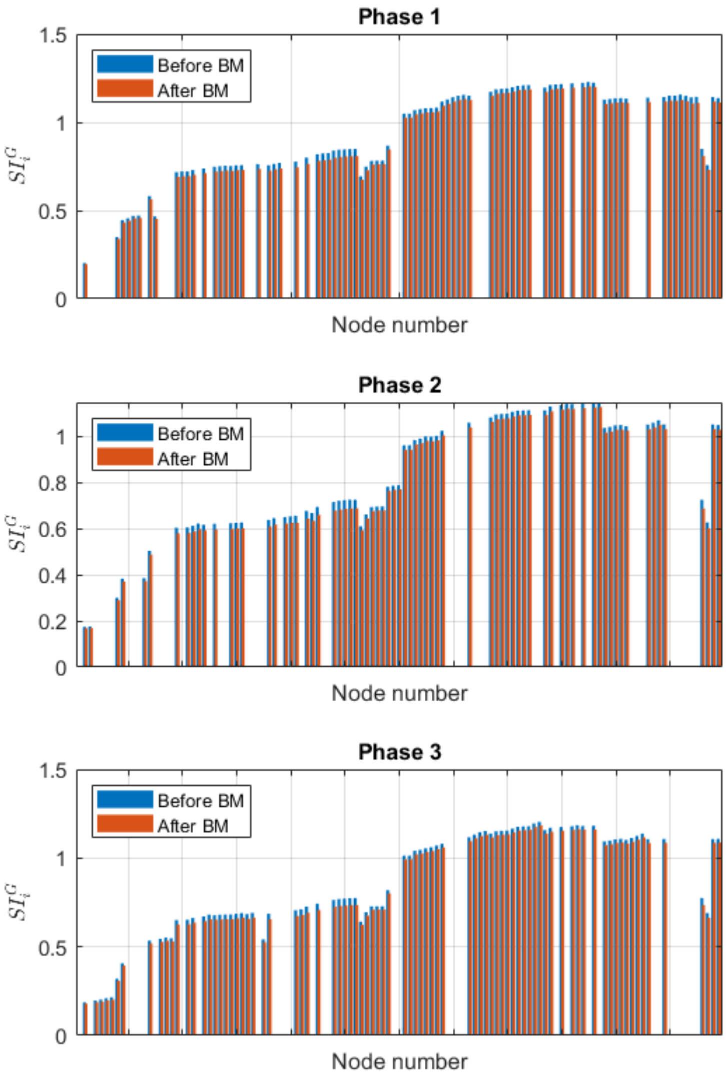


Figure 5.9: $SI_i^V$ before and after the BM session for the Modified IEEE 123 Bus system

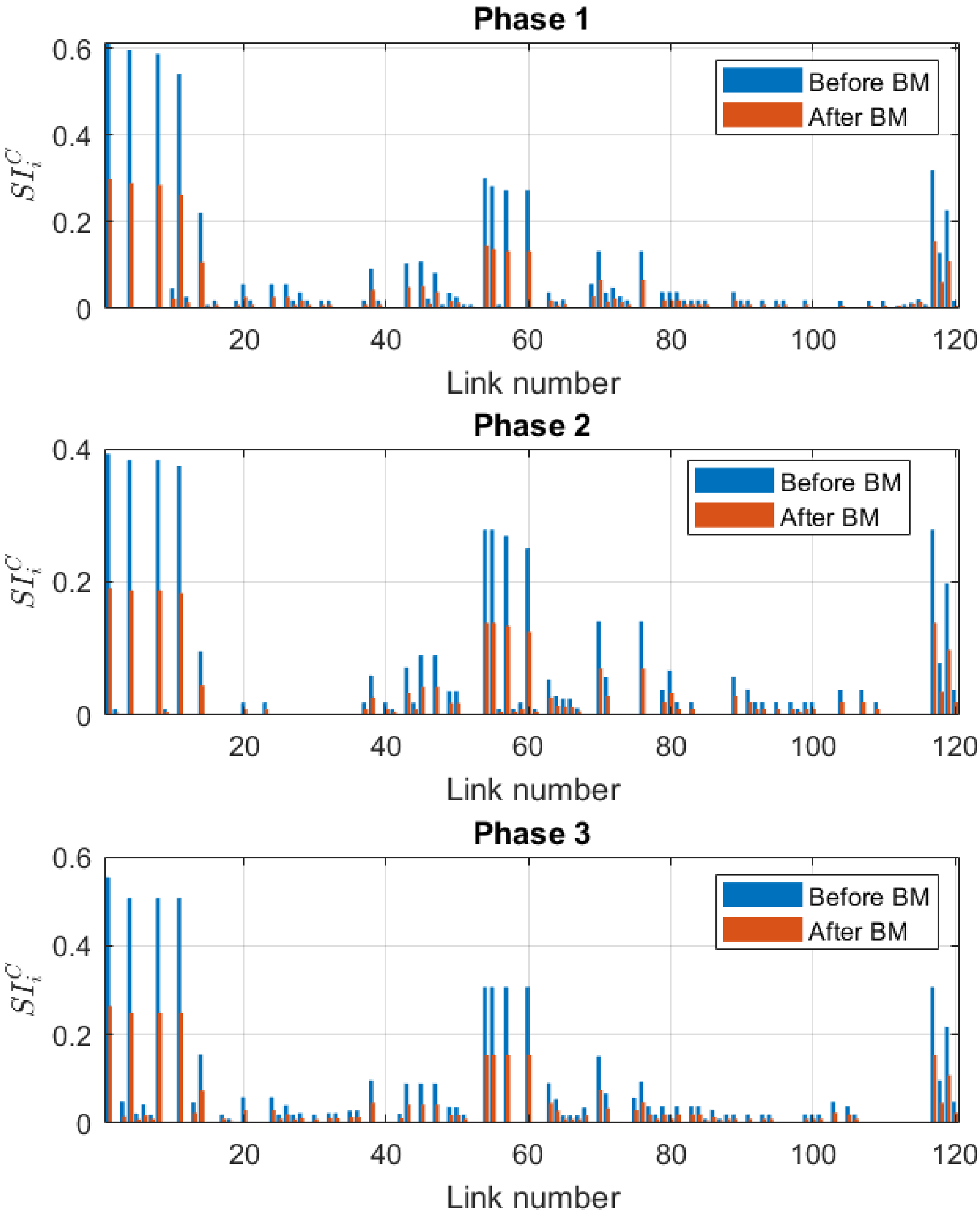


Figure 5.10: $SI_i^C$ before and after the BM session for the Modified IEEE 123 Bus system

# Chapter 6

# Conclusions and future developments

The Ph.D. thesis concerns the development and application of new methods for the optimization of network systems (with multi-decision makers) with particular attention to the case of power distribution networks. In particular, the definition of new competitive and cooperative market structures in distribution power grids is addressed. This need arises from the fact that in the recent future the possibility of active participation of loads and generators (aggregates and non-aggregates) in a balancing market for distribution networks will be introduced. In Chapter 2 the overall modeling framework is defined, stating the optimization problems for the market actors, namely generators, aggregators, and DSO. The first critical points concerning the accurate modeling of the electricity grid, i.e. the nonconvexity of the power flow equations, have been highlighted.

In Chapter 3 a new convex relaxation for the OPF problem for the distribution networks is presented. The main idea is to reformulate the problem as CI instead of power injections/absorption. The non-convexity of the model due to bilinear terms has been codified through the McCormick envelopes, obtaining a linear relaxation of these constraints. New constraints on current injections to improve the quality of the relaxation are obtained optimally through an explicit rule based on simple optimization problems, which do not add any computational burden to the centralized solver. With this new formulation, it is possible to solve OPF problems of any kind without any restriction of topology, network type, and technology. This aspect is extremely innovative compared to the current state of the art because most of the models are not as flexible as the one presented. The new model has been applied to the main IEEE testbeds for distribution networks (IEEE 13 and 123 bus, and IEEE 8500 nodes), suitably modified to exploit the model's potential. The results show an excellent quality of the lower bound obtained and considerably shorter calculation times compared to the non-linear model. Future developments will concern the application of the model in the context of distributed and multi-period optimization, as well as the development of new heuristics for further improvement of relaxation.

In Chapter 4, a multi-level approach for the modeling and optimization of transmission networks' BM has been developed and tested. The proposed BM scheme has been described as a bilevel and multi-decision maker optimization problem. At the highest level, there is the market arbiter, i.e. the DSO, who wants to keep the quality of the network service as high as possible but spending as little as possible in buying flexibility from the market participants. Instead, at the lower level, there are the generators and aggregators (which can operate on multiple network nodes at the same time) that want to maximize their profits by selling flexibility to the DSO. Each market participant has been modeled through its own cost and profit functions and its own set of constraints. To take into account these participants into the DSO problem, optimality constraints have been derived, which ensure that each lower-level decision maker behaves optimally. The resulting optimization problem has been solved for the modified IEEE 13 and 123 bus systems described in Chapter 2. The results show that the quality of the network service is improved after the market session. Future developments will include new BM participants such as electric vehicle aggregators, intermittent renewables, electric storage systems, and even perform multi-period optimization.

In Chapter 5 the PAC algorithm for distributed convex optimization is presented and discussed. We first detailed how we can partition a centralized formulation into an atomized one. We then derived the PAC algorithm using a prox-linear approach and showed how the PAC algorithm achieves an $\epsilon$-solution in at most $o\left(\frac{1}{\tau}\right)$ iterations. The distributed algorithm has then been applied to the BM optimization problem cooperatively defined in Chapter 2 to create a new market structure with applications to the modified IEEE 13 and 123 bus systems. The results show that also in this case the quality of the network service is improved after the market session. Future research directions will include the application of PAC under non-convex constraints and/or objective functions, to prove the PAC convergence to local minima in the presence of non-convexities. Ongoing research also concerns the development of new methodologies to speed up the time required for solving the proximal function inherent in each PAC primal update.

# Appendices

# Appendix A

# Preliminaries

In this section, some basic concepts that are fundamental to the the overall understanding of this PhD thesis are presented. That are:

- Complex numbers and phasors
- Optimization problems
- Convex sets and convex functions
- Karush-Kuhn-Tucker Conditions
- Multilevel programming
- McCormick envelopes
- Introduction to graph theory
- Consensus Based distributed optimization
- Power systems basics

Furthermore, in order to help the reader to understand these concepts, some examples are provided.

## A.1 Complex numbers and phasors

The complex numbers are the field $\mathbb{C}$ of numbers in the form $a + ib$, where $a, b \in \mathbb{R}$, and $i$ is a solution of the equation

$$x^2 = -1 \tag{A.1}$$

Since no real number satisfies this equation, $i$ is called an imaginary number. For the complex number $z = a+ib$, $a$ is called the real part, and $b$ is called the imaginary part. This representation of a number $z \in \mathbb{C}$ is called the rectangular coordinate form of a complex number, since it can be represented in a Cartesian plane like in Figure A.1, where the $x$-axis, denoted as $Re$, represents the real axis, while the y-axis $Im$ represents the imaginary axis.

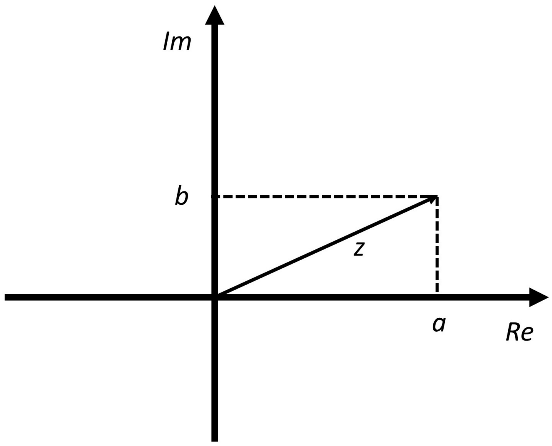


Figure A.1: Cartesian representation of a complex number.

The $\mathbb{C}$ set can be defined as the algebraic extension of $\mathbb{R}$ adding the imaginary number $i$. This means that complex numbers can be added, subtracted, and multiplied, as polynomials in the variable $i$, according to the rule

$$i^2 = -1 \tag{A.2}$$

Another way to represent a complex number $z$ through the use of the polar coordinates (called polar form). It is possible to state the real and imaginary components of a complex number $z$ in terms of $r$ and $\theta$. Namely, $r$ is the magnitude of the vector associated with $z$ and $\theta$ is the angle (phase) made by the vector with the real axis. Using Euler's formula we can represents a complex number in terms of magnitude and phase as:

$$z = re^{i\theta} = r(\cos\theta + i\sin\theta) \tag{A.3}$$

An important property of the complex numbers is the complex conjugate. The complex conjugate of a complex number is the complex number with an equal real part and an imaginary part equal in magnitude but opposite in sign:

- In rectangular from, the complex conjugate of $a + ib$ is $a - ib$.
- In polar form, the complex conjugate of $re^{i\theta}$ is $re^{-i\theta}$.

- As regards matrices and vectors we define $x^H$ as the Hermitian of the matrix or vector $x$, that is the complex conjugate and transpose of $x$.

In power systems engineering, a phasor (denoted as $\dot{A}$), is a complex number representing a sinusoidal function whose amplitude $A$, frequency $\omega$, and phase $\theta$ are time-invariant. Euler's formula indicates that sinusoids can be represented mathematically as the sum of two complex-valued functions:

$$\dot{A} = A\cos(\omega t + \theta) = A\frac{e^{i(\omega t+\theta)} + e^{-i(\omega t+\theta)}}{2} \tag{A.4}$$

## A.2 Optimization problems

An optimization problem or mathematical programming problem is a mathematical entity that allows maximizing or minimizing a certain objective (i.e., objective function) subject to diverse restrictions, typically in the form of equality or inequality constraints. An optimization problem has the general form

$$\min_{x} f\left(x\right) \tag{A.5}$$

$$\text{s. t.}$$

$$h(x) = 0 \tag{A.6}$$

$$g(x) \leq 0 \tag{A.7}$$

where $x \in \mathbb{R}^n$ is the decision variable vector, $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function to be minimized, $h(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n_E}$ are the functions for the equality constraints and $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^{n_I}$ for the inequality constraints. The following observations valid in this context:

- Any $x \in \mathbb{R}^n$ is called a solution, while a solution meeting (A.6)-(A.7) is called a feasible solution, and a feasible solution minimizing (A.5) is called an optimal solution or a minimizer. The objective function value at the minimizer is the objective function optimal value.
- The set of solutions meeting both $h(x) = 0$ and $g(x) \leq 0$ constitutes the feasible region of the optimization problem.
- Minimizing $f(x)$ is equivalent to maximizing $-f(x)$.

**Example 1 : Linear optimization problem**

In the following, an example of linear programming (both objective function and constraints are linear) problem is provided.

$$\max_{x,y} \left(x + y - 50\right) \tag{A.8}$$

$$50x + 24y \leq 2400 \tag{A.9}$$

$$30x + 33y \leq 2100 \tag{A.10}$$

$$x \geq 45 \tag{A.11}$$

$$y \geq 5 \tag{A.12}$$

$$x, y \in \Re \tag{A.13}$$

The optimal solution (obtanied though a mathematical programming solver) is given by $x = 45$, $y = 6.25$ and the value of the objective function is $1.25$.

## A.3 Convex sets and convex functions

In this section an important property of sets and functions is defined, i.e., the convexity. This property is essential for nonlinear optimization theory over this kind of functions and sets. As a mater of fact, the convexity property (of the objective function and the feasibility region) implies the uniqueness of the optimal solution [87]. In the following, the formal definition of convex set and convex functon are provided.

- A set $C$ is *convex* if the line segment between any two points in $C$ lies in $C$, i.e., if for any $x_1, x_2 \in C$ and any $\theta$ with $0 \leq \theta \leq 1$, we have $\theta x_1 + (1-\theta)x_2 \in C$.
- A function $f : \mathbb{R}^n \to \mathbb{R}$ is *convex* if **dom** $f$ is a convex set and if for all $x, y \in \mathbf{dom} f$, and $\theta$ with $0 \leq \theta \leq 1$, we have $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$.

More informally, a set defined on an $n$-dimensional interval is called convex if the line segment between any two points on the set of the function lies on the set itself (as shown in Figure A.2). Equivalently, a function is convex if its epigraph (the set of points on or above the graph of the function) is a convex set. For a twice-differentiable function of a single variable, if its second derivative is always nonnegative on its entire domain, then the function is convex.

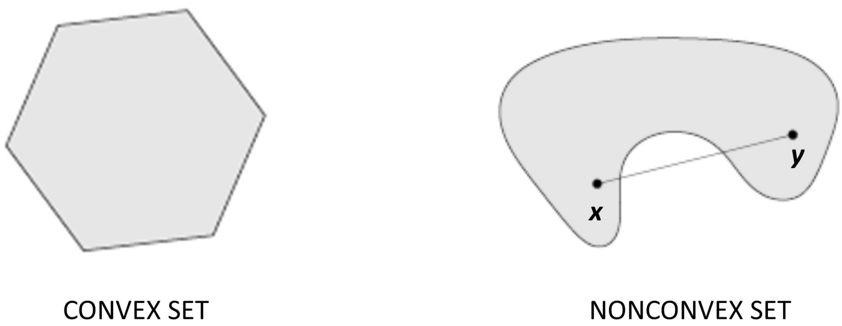


Figure A.2: Convex and nonconvex sets

Well-known examples of convex functions of a single variable include the linear functions ($cx$ where $c$ is a known constant), the square function $x^2$ and the exponential function $e^x$. On the contrary, the sinusoidal functions i.e. $cos(x)$ and $sin(x)$ are non convex.

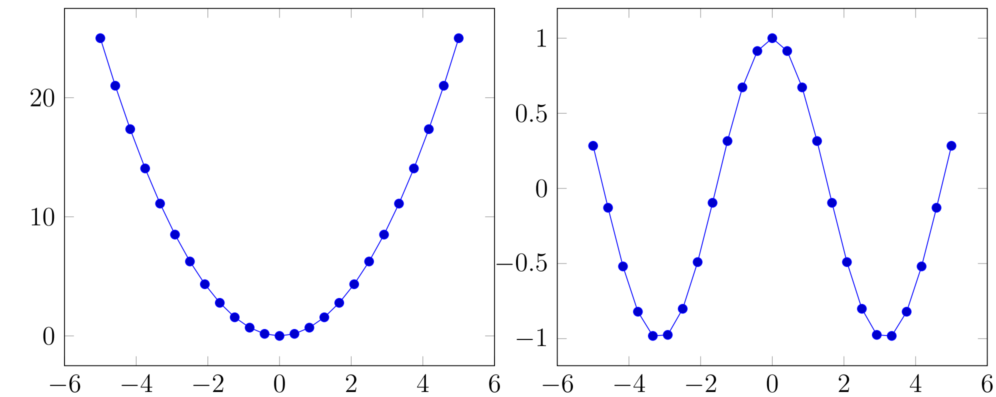


Figure A.3: Convex and nonconvex functions

## A.4 Karush-Kuhn-Tucker Conditions

The Karush-Kuhn-Tucker (KKT) conditions are first-order conditions such that the optimal solutions of a broad range of optimization problems should satisfy. Furthermore, KKT conditions can be necessary but not sufficient conditions feasible solutions meeting them are not necessarily optimal but optimal solutions need to meet them, these conditions become necessary and sufficient under convexity (both feasibility set and objective function) hypothesis. KKT conditions are first-order conditions, i.e., conditions that are formulated using first derivative vectors and matrices (gradients and Jacobians). To formulate the KKT conditions it is convenient to define the Lagrange function.

The Lagrange function of problem (A.5)-(A.7) is

$$\mathcal{L} = f(x) + \lambda^T h(x) + \mu^T g(x) \tag{A.14}$$

where $f(x)$, $h(x)$ and $g(x)$ are continuously differentiable in the feasibility set ($x \in \{x|h(x) = 0, g(x) \leq 0\}$). The KKT conditions of problem (A.5)-(A.7) are

$$\nabla_x f(x) + \lambda^T \nabla_x h(x) + \mu^T \nabla_x g(x) = 0 \tag{A.15}$$

$$h(x) = 0 \tag{A.16}$$

$$g(x) \leq 0 \tag{A.17}$$

$$\mu^T g(x) = 0 \tag{A.18}$$

$$\mu \geq 0 \tag{A.19}$$

where $\lambda \in \mathbb{R}^{n_E}$ and $\mu \in \mathbb{R}^{n_I}$ are respectively, the equality and inequality Lagrange multiplier vectors, and $\nabla_x$ denotes the gradient with respect to $x$.

Constraint (A.15) states that the gradient of the Lagrangian function of (A.14) at an optimal solution $x$ should be zero. Constraint (A.16) states the equality constraints. Constraint (A.17) enforces the inequality constraints. Constraint (A.18) states that inner product of the multiplier vector of the inequality constraints and the inequality constraint vector is zero. Constraint (A.19) states that the multiplier vector of the inequality constraints is nonnegative. Constraints (A.17)-(A.19) are known as complementarity conditions and are equivalently written as $0 \leq \mu \perp g(x) \leq 0$ where $\perp$ indicates complementarity.

**Example 2 : The KKT conditions of problem (A.8)- (A.13)**

First of all, it is important to define the Lagrange function of the optimization problem (according to (A.14)):

$$\mathcal{L}(x,y,\mu) = (x + y - 50) + \mu_1(50x + 24y - 2400) + \mu_2(30x + 33y - 2100) + \mu_3(-x + 45) + \mu_4(-y + 5) \tag{A.20}$$

The set of KKT condition id given by:

$$\frac{\partial \mathcal{L}}{\partial x} = 1 + 50\mu_1 + 30\mu_2 - \mu_3 \tag{A.21}$$

$$\frac{\partial \mathcal{L}}{\partial y} = 1 + 24\mu_1 + 33\mu_2 - \mu_4 \tag{A.22}$$

$$50x + 24y - 2400 \leq 0 \tag{A.23}$$

$$30x + 33y - 2100 \leq 0 \tag{A.24}$$

$$-x + 45 \leq 0 \tag{A.25}$$

$$-y + 5 \leq 0 \tag{A.26}$$

$$\mu_1(50x + 24y - 2400) = 0 \tag{A.27}$$

$$\mu_2(30x + 33y - 2100) = 0 \tag{A.28}$$

$$\mu_3(-x + 45) = 0 \tag{A.29}$$

$$\mu_4(-y + 5) = 0 \tag{A.30}$$

$$\mu_1 \geq 0 \tag{A.31}$$

$$\mu_2 \geq 0 \tag{A.32}$$

$$\mu_3 \geq 0 \tag{A.33}$$

$$\mu_4 \geq 0 \tag{A.34}$$

Since problem (A.8)- (A.13) is linear (therefore the objective function and the feasibility set is convex), the optimal solution is unique (the same as in Example 1).

## A.5 Multilevel programming

A multilevel programming problem (MP) is an optimization problem involving constraints that represent equilibrium conditions in other words, an MP is an optimization problem whose constraints include other interrelated optimization problems. The formulation of an MP is provided below

$$\operatorname*{Min}_{x} \; f\left(x, x_1, ..., x_N\right) \tag{A.35}$$

$$s.t$$

$$g\left(x, x_1, ..., x_N\right) = 0 \tag{A.36}$$

$$h\left(x, x_1, ..., x_N\right) \leq 0 \tag{A.37}$$

$$\begin{cases} \operatorname*{Min}_{x_1} \; f_1\left(x, x_1, ..., x_N\right) \\ s.t \\ g_1\left(x, x_1, ..., x_N\right) = 0 \\ h_1\left(x, x_1, ..., x_N\right) \leq 0 \end{cases} \tag{A.38}$$

$$\vdots$$

$$\begin{cases} \operatorname*{Min}_{x_N} \; f_N\left(x, x_1, ..., x_N\right) \\ s.t \\ g_N\left(x, x_1, ..., x_N\right) = 0 \\ h_N\left(x, x_1, ..., x_N\right) \leq 0 \end{cases} . \tag{A.39}$$

Observe the nested structure of problem (A.35)-(A.37) above, i.e., an optimization problem constrained by a set of $N$ optimization problems. Vector $x \in \mathbb{R}^{n^o}$ includes the set of optimization variables of the upper level problem, vector $x_i \in \mathbb{R}^{n^i}$ includes the set of optimization variables of constraining problem $i$, and $\lambda_i \in \mathbb{R}^{n^i}$ and $\mu_i \in \mathbb{R}^{n^i}$ are the dual variable vectors of problem $i$ associated with the equality and inequality constraints.

The overall MP problem con be solved providing an analytical solution of the lower level optimization problem to be added at the upper level one. As an example, KKT conditions for lower level decision problems can be used. Therefore, in this case we will solve an "augmented" version of the upper level optimization adding as new constraints the KKT conditions of each lower level problem. The overall optimization problem is given by:

$$\underset{x}{\text{Min}} \; f\left(x, x_1, ..., x_N\right) \tag{A.40}$$

$$s.t$$

$$g\left(x, x_1, ..., x_N\right) = 0 \tag{A.41}$$

$$h\left(x, x_1, ..., x_N\right) \leq 0 \tag{A.42}$$

$$KKT_1 \tag{A.43}$$

$$\vdots$$

$$KKT_N \tag{A.44}$$

**Example 3: Multilevel programming problem**

The following example has been taken from [53]

$$\underset{\pi, f_1, s_1, s_2}{\text{Min}} \quad -\pi(s_1 + f_1) + a_1 + b_1(s_1 + f_1) \tag{A.45}$$

s.t

$$\begin{cases} \underset{s_1}{\text{Min}} \quad -\pi s_1 + a_1 + b_1(s_1 + f_1) \\ s.t \\ -\pi + \gamma - \beta(s_1 + s_2 + f_1 + f_2) = 0 \\ s_1 - s_1^{max} \leq 0 \end{cases} \tag{A.46}$$

$$\begin{cases} \underset{s_2}{\text{Min}} \quad -\pi s_2 + a_2 + b_2(s_2 + f_2) \\ s.t \\ -\pi + \gamma - \beta(s_1 + s_2 + f_1 + f_2) = 0 \\ s_2 - s_2^{max} \leq 0 \end{cases} \tag{A.47}$$

Using the KKT conditions of the two lower-level problems the overall problem becomes

$$\underset{\pi, f_1, s_1, s_2, \lambda_1, \mu_1, \lambda_2, \mu_2}{\text{Min}} \quad -\pi(s_1 + f_1) + a_1 + b_1(s_1 + f_1) \tag{A.48}$$

s. t.

$$-\pi + b_1 - \lambda_1 \beta + \mu_1 = 0 \tag{A.49}$$

$$-s_1 - \lambda_1 = 0 \tag{A.50}$$

$$-\pi + \gamma - \beta(s_1 + s_2 + f_1 + f_2) = 0 \tag{A.51}$$

$$0 \leq \mu_1 \perp (s_1 - s_1^{max}) \leq 0 \tag{A.52}$$

$$-\pi + b_2 - \lambda_2 \beta + \mu_2 = 0 \tag{A.53}$$

$$-s_2 - \lambda_2 = 0 \tag{A.54}$$

$$-\pi + \gamma - \beta(s_1 + s_2 + f_1 + f_2) = 0 \tag{A.55}$$

$$0 \leq \mu_2 \perp (s_2 - s_2^{max}) \leq 0 \tag{A.56}$$

## A.6 McCormick envelopes

In this section we will study bilinear forms and its convexification. This approach is useful in the definition of the new convex relaxation of the OPF problem that is one of the main cores of this PhD thesis. We define a bilinear form as $w = xy$ with $x, y \in \mathbb{R}$. This is a classical nonconvex and nonlinear function (as can be seen in Figure A.4) recurring in optimization problems.

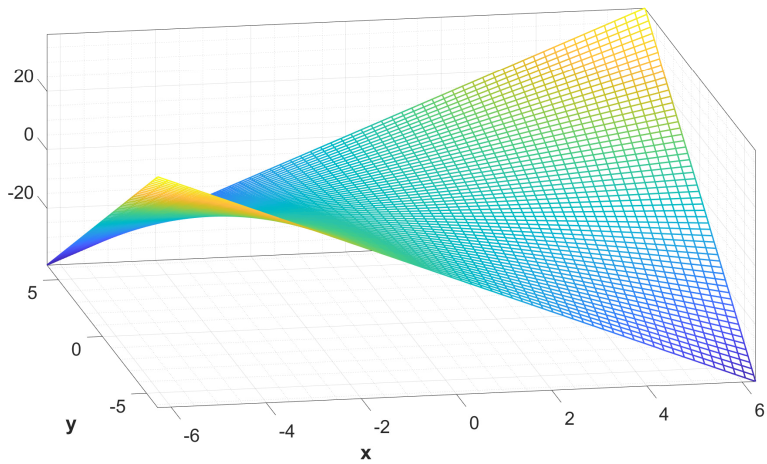


Figure A.4: A bilinear form

McCormick Envelopes (MCE) is a type of convex relaxation used in bilinear Non-Linear Programming problems. Solving nonconvex problems is a complicated task. First, the nonconvex function is transformed into a convex function by relaxing the parameters on the problem.

Relaxing the bounds through a convex relaxation decreases the computational difficulty of solving the problem at the cost of introducing solutions that do not correspond to the original objective function i.e. the optimal solution in the relaxed problem will not always be the optimal solution to the objective problem. Solving the convex problem will provide a lower bound for the optimal solution. Having a tighter relaxation that is still convex will provide a lower bound that is closer to the solution.

The enevlope of a bilinear form $w = xy$ defined over a set $\boldsymbol{S} \subset \mathbb{R}^3 = \{x, y : x \in [\underline{x}, \overline{x}], y \in [\underline{y}, \overline{y}]\}$ is stated as:

$$w \geq \underline{x}y + x\underline{y} - \underline{x}\underline{y} \tag{A.57}$$

$$w \geq \overline{x}y + x\overline{y} - \overline{x}\overline{y} \tag{A.58}$$

$$w \leq \underline{x}y + x\overline{y} - \overline{x}\underline{y} \tag{A.59}$$

$$w \leq \overline{x}y + x\underline{y} - \underline{x}\overline{y} \tag{A.60}$$

McCormick envelopes are obtained through simple mathematical steps. In order to obtain (A.57), we proceed defining the following quantities:

$$\hat{w}_1 = x - \underline{x} \tag{A.61}$$

$$\hat{w}_2 = y - \underline{y} \tag{A.62}$$

since $\hat{w}_1\hat{w}_2 \geq 0$ we have

$$\hat{w}_1\hat{w}_2 = (x - \underline{x})(y - \underline{y}) \tag{A.63}$$

$$= xy - \underline{x}y - x\underline{y} + \underline{x}\underline{y} \geq 0 \tag{A.64}$$

$$\implies w \geq \underline{x}y + x\underline{y} - \underline{x}\underline{y} \tag{A.65}$$

thus (A.57) is obtained. the other relations, namely (A.58), (A.59) and (A.60), can be obtained defining

$$\hat{w}_3 = \overline{x} - x \tag{A.66}$$

$$\hat{w}_4 = \overline{y} - y \tag{A.67}$$

and and making all the following combinations:

$$\hat{w}_3\hat{w}_4 = (\overline{x} - x)(\overline{y} - y) \geq 0 \tag{A.68}$$

$$= xy - \overline{x}y - x\overline{y} + \overline{x}\overline{y} \geq 0 \tag{A.69}$$

$$\implies w \geq \overline{x}y + x\overline{y} - \overline{x}\overline{y} \tag{A.70}$$

$$\hat{w}_2\hat{w}_3 = (y - \underline{y})(\overline{x} - x) \leq 0 \tag{A.71}$$

$$= -xy + \overline{x}y + x\underline{y} - \overline{x}\underline{y} \leq 0 \tag{A.72}$$

$$\implies w \leq \overline{x}y + x\underline{y} - \overline{x}\underline{y} \tag{A.73}$$

$$\hat{w}_1\hat{w}_4 = (x - \underline{x})(\overline{y} - y) \leq 0 \tag{A.74}$$

$$= -xy - \underline{x}y + x\overline{y} + \underline{x}\overline{y} \leq 0 \tag{A.75}$$

$$\implies w \leq \underline{x}y + x\overline{y} - \underline{x}\overline{y} \tag{A.76}$$

**Example 4: The McCormick envelope of $\mathbf{x^2}$**

In this example the envelope of the bilinear form $w = x^2$ over a set $-1 \leq x \leq 2$. The envelope is given by:

$$(2 - x)(2 - x) = w - 4x + 4 \geq 0 \tag{A.77}$$

$$(x + 1)(2 - x) = -w + x + 2 \geq 0 \tag{A.78}$$

$$(x + 1)(x + 1) = w + 2x + 1 \geq 0 \tag{A.79}$$

## A.7 Introduction to graph theory

In this section, we will introduce some basic principles about graph theory that are useful for the basic modeling of a distribution network and the definition of the distributed optimization algorithm.

A simple graph $G(V, E)$ consists of a non-empty finite set $V$ of elements called vertices (or nodes), and a finite set $E$ of distinct unordered pairs of distinct elements of $V$ called edges. We call $V(G)$ the vertex set and$E(G)$ the edge set of $G$. Two main categories of graphs are:

- **Directed graph**: a graph where the edges have a direction associated with them (Figure A.5).
- **Undirected graph**: a graph where all the edges are bidirectional (Figure A.6).

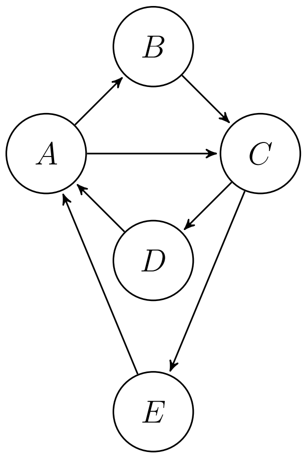


Figure A.5: A directed graph

A useful representation of a graph is the adjacency matrix. For a graph with vertex set $V$, the adjacency matrix is a square $|V| \times |V|$ matrix $A$ defined as follows:

$$\begin{cases} A_{ij} = 1, & \text{if } \ ij \in E \\ A_{ij} = 0, & \text{otherwise} \end{cases} \tag{A.80}$$

For a graph with no self-loops, the adjacency matrix must have 0s on the diagonal. For an undirected graph, the adjacency matrix is symmetric. The adjacency matrices of the graphs

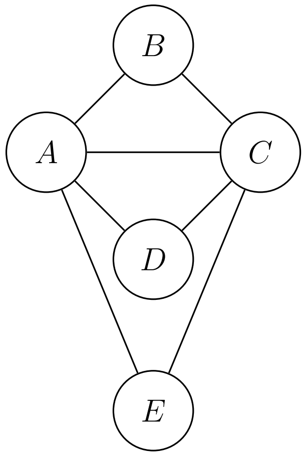


Figure A.6: An undirected graph

depicted in Figures A.6 and A.5 are:

$$A_{ud} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{A.81}$$

$$A_{d} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{A.82}$$

## A.8 Consensus Based distributed optimization

Distributed consensus optimization is an important family of optimization problems which is formulated as

$$\min_{x} \sum_{i=1}^{N} f_i(x) \tag{A.83}$$

s. t.

$$h(x) = 0 \tag{A.84}$$

$$g(x) \leq 0 \tag{A.85}$$

where $x \in \mathbb{R}^n$ and $f_i(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ are convex functions, $h(x)$ and $g(x)$ describe a convex set of solutions. $N$ agents want to optimize the sum of local objective functions, $f_i(x)$, $i \in \{1, ..., N\}$ over a global variable vector $x$. In order to solve the problem in a distributed way, it is necessary to decompose the problem into sub-problems, the resolution of which in an iterative way through the exchange of information with other agents leads to a global solution of the problem.

One possible strategy to decompose the problem, and solve it in a completely decentralised way is to use the consensus method. This methodology proposes to define an augmented version of the problem by introducing the so-called "copy" variables. In this way it is possible to duplicate the variables needed to define each sub-problem; the price to pay will result in increasing the overall number of variables and introducing a new set of constraints:

$$\min_{\tilde{x} = x \bigcup \hat{x}} \sum_{i=1}^{N} f_i(\tilde{x}) \tag{A.86}$$

s. t.

$$h(\tilde{x}) = 0 \tag{A.87}$$

$$g(\tilde{x}) \leq 0 \tag{A.88}$$

$$x - \hat{x} = 0 \tag{A.89}$$

As can be seen from the new version of the problem you have a new set of variables $\tilde{x} = x \bigcup \hat{x}$, where $\hat{x}$ is the set of copied variables needed to fully decompose the problem. The equivalence of the global solution is guarantee by adding the new constraint (A.89) that states the equivalence between the "main" variables and their copy.

## A.9 ADMM

In this section the popular Alternating Direction Method of Multipliers (ADMM) algorithm is briefly presented. ADMM is an algorithm that is intended to unite the decomposability of dual ascent with the superior convergence properties of the method of multipliers. The algorithm solves problems in the form:

$$\min_{x,z} \; f(x) + g(z) \tag{A.90}$$

$$\text{s. t.}$$

$$Ax + Bz = c \tag{A.91}$$

with variables $x \in \mathbb{R}^n$ and $z \in \mathbb{R}^m$, where $A \in \mathbb{R}^{p \times n}$, $B \in \mathbb{R}^{p \times m}$, and $c \in \mathbb{R}^p$. We will assume that $f$ and $g$ are convex. The only difference from the general linear equality-constrained version of problem (A.5)-(A.7) is that the variable, called $x$ there, has been split into two parts, called $x$ and $z$ here, with the objective function separable across this splitting. To solve the problem, we form the augmented Lagrangian

$$L(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + (\rho/2)\|Ax + Bz - c\|_2^2 \tag{A.92}$$

where $y$ is the Lagrange multiplier associated to (A.91), and $\rho$ is the step size parameter. The algorithm performs the following iterations, until convergence:

$$x[\tau + 1] = \underset{x}{\text{argmin}} \; L(x, z[\tau], y[\tau]) \tag{A.93}$$

$$z[\tau + 1] = \underset{z}{\text{argmin}} \; L(x[\tau + 1], z, y[\tau]) \tag{A.94}$$

$$y[\tau + 1] = y[\tau] + \rho(Ax[\tau + 1] + Bz[\tau + 1] - c) \tag{A.95}$$

The convergence proof of ADMM algorithm is proven in [67]

## A.10 Power systems basics

### A.10.1 Admittance matrix

In power engineering, nodal admittance matrix or $Y$ Matrix is an $N \times N$ matrix describing a power system with $N$ buses. For a network with $N$ buses, the admittance between a bus $k$, and another bus $i$ is described by $y_{ik} = g_{ik} + ib_{ik}$, where $g_{ik}$ is the line conductance and $b_{ik}$ is the line susceptance. The general mathematical expression of $Y$ is:

$$Y_{ij} = \begin{cases} \sum\limits_{m:(i,m)\in E_i} y_{im} & \text{if } i = j \\ -y_{ij} & \text{if } i \neq j \text{ and } (i,j) \in E \\ 0 & \text{otherwise} \end{cases} \tag{A.96}$$

It is important to note that $y_{ki}$ is non-zero only where a physical connection exists between two buses. From the general case where $N$ is greater than 2, it is desirable to solve these equations as a system, namely through matrix algebra. The general $Y$ matrix appears as follows:

$$Y = \begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1n} \\ Y_{21} & Y_{22} & \cdots & Y_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ Y_{n1} & Y_{n2} & \cdots & Y_{nn} \end{bmatrix} \tag{A.97}$$

**Example 5: Admittance matrix of a three-nodes system**

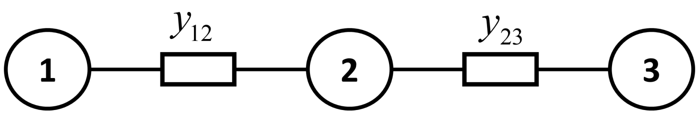


Figure A.7: 3 nodes example

Considering the network depicted in Figure A.7 according to A.96 the $Y$ matrix is:

$$Y = \begin{bmatrix} y_{12} & -y_{12} & 0 \\ -y_{12} & y_{22} + y_{23} & -y_{23} \\ 0 & -y_{23} & y_{23} \end{bmatrix} \tag{A.98}$$

### A.10.2 Three phase systems

An unbalanced three-phase circuit is one that contains at least one source or load that does not possess three-phase symmetry. A source with the three source-function magnitudes unequal and/or the successive phase displacements different from $\frac{2}{3}\pi$ will make the overall network unbalanced. Similarly, a three-phase load with different power request per phase make the overall network unbalanced. In this case, the cables have to be expressed by the 3-phase admittance matrix $y_{r,s} \in \mathbb{R}^{3\times3}$, for a cable between nodes $r$ and $s \in N$, in which each element of the matrix, $y_{k,k'}$, represents the magnetic coupling between phases $k$ and $k'$, such that $y_{k,k'} = g_{k,k'} + ib_{k,k'} \forall k, k' \in K$ where $K = a, b, c$ is the set of phases. That is:

$$y_{ij} = \begin{bmatrix} y_{ij}^{aa} & y_{ij}^{ab} & y_{ij}^{ac} \\ y_{ij}^{ab} & y_{ij}^{bb} & y_{ij}^{bc} \\ y_{ij}^{ac} & y_{ij}^{bc} & y_{ij}^{cc} \end{bmatrix} \ \forall i, j \in N \tag{A.99}$$

### A.10.3 Nodal Analysis

In power systems analysis, the nodal analysis method is a way to determine the voltage between the nodes of a power network, in terms of the injected currents at each node.

Nodal analysis writes an equation at each electrical node, according to Kirchhoff's current law (KCL), requiring that the sum of the branch currents at each node must be zero. The branch currents are written in terms of the network's node voltages. Nodal analysis is possible when all the circuit elements' branch constitutive relations have an admittance representation. The nodal analysis produces a compact set of equations for the network, which can be quickly solved using linear algebra.

In general, for a network with $N$ nodes, the node-voltage equations obtained by nodal analysis can be written in a matrix form as derived in the following.

For any node $k \in N$, KCL states:

$$\sum_{j\neq k\in N} Y_{jk}(v_k - v_j) = 0 \ \forall k \in N \tag{A.100}$$

where $Y_{kj} = Y_{jk}$ is the negative of the sum of the admittances between nodes $k$ and $j$, and $v_k$ is the voltage of node $k$. This implies

$$\begin{aligned} 0 = \sum_{j\neq k\in N} Y_{jk}(v_k - v_j) &= \sum_{j\neq k\in N} Y_{jk} v_k - \sum_{j\neq k\in N} Y_{jk} v_j \\ &= Y_{kk} v_k - \sum_{j\neq k\in N} Y_{jk} v_j \ \forall k \in N \end{aligned} \tag{A.101}$$

where $Y_{kk}$ is the sum of conductances connected to node $k$. We note that the first term contributes linearly to the node $k$ via $Y_{kk}$, while the second term contributes linearly to each node $j$connected to the node $k$ via $Y_{jk}$ with a minus sign. If an independent current source/input $i_k$ is also attached to node $k$, the above expression is generalized to $i_k = Y_{kk} v_k - \sum_{j\neq k} Y_{jk} v_j$. It is readily to show that one can combine the above node-voltage equations for all $N$ nodes, and write them down in the following matrix form:

$$\begin{pmatrix} Y_{11} & Y_{12} & \cdots & Y_{1N} \\ Y_{21} & Y_{22} & \cdots & Y_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{N1} & Y_{N2} & \cdots & Y_{NN} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_N \end{pmatrix} = \begin{pmatrix} i_1 \\ i_2 \\ \vdots \\ i_N \end{pmatrix} \tag{A.102}$$

or simply

$$I = YV \tag{A.103}$$

In the network presented in Figure A.7 relation A.102 becomes

$$\begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} y_{12} & -y_{12} & 0 \\ -y_{12} & y_{22} + y_{23} & -y_{23} \\ 0 & -y_{23} & y_{23} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \tag{A.104}$$

# Appendix B

# Proofs of Chapter 5

**Lemma B.0.1.** *Suppose* $D = diag\left(d_1, \ldots, d_M\right) \succ 0$ *and* $A \in \mathbb{R}^{M\times N}$*. Let* $\left\{\lambda_1\left(B\right), \ldots, \lambda_N\left(B\right)\right\}$ *be the eigenvalues of matrix* $B \in \mathbb{R}^{N\times N}$ *with* $\lambda_i\left(B\right) \leq \lambda_j\left(B\right)$ *for* $1 \leq i \leq j \leq N$*. Then:*

*a)* $A^T DA \succeq 0$*;*

*b)* $d_{min}\lambda_j\left(A^T A\right) \leq \lambda_j\left(A^T DA\right) \leq d_{max}\lambda_j\left(A^T A\right)$*;*

*Proof of Lemma B.0.1.*

a) Let $x \in \mathbb{R}^N$. Then:

$$x^{\mathrm{T}}A^{\mathrm{T}}DAx = x^{\mathrm{T}}A^{\mathrm{T}}D^{\frac{1}{2}}D^{\frac{1}{2}}Ax = \left(D^{\frac{1}{2}}Ax\right)^{\mathrm{T}}\left(D^{\frac{1}{2}}Ax\right) = \left\|D^{\frac{1}{2}}Ax\right\|_2^2 \geq 0 \tag{B.1}$$

b) Recall the following from Courant-Fischer-Weyl Min-Max Principle:

$$\lambda_j\left(A^T DA\right) = \min_{U:\dim(U)=j}\left\{\max_{x\in U:\|x\|_2^2=1}\left\{x^T A^T DAx\right\}\right\} \tag{B.2}$$

$$\lambda_j\left(A^T DA\right) = \max_{U:\dim(U)=N-j+1}\left\{\min_{x\in U:\|x\|_2^2=1}\left\{x^T A^T DAx\right\}\right\}. \tag{B.3}$$

Next, if $x \in \mathbb{R}^{N\times N}$, then:

$$x^{\mathrm{T}}A^{\mathrm{T}}DAx = \left\|D^{\frac{1}{2}}Ax\right\|_2^2 = \sum_{j=1}^{M}\left(\sqrt{d_j}\,[A]^j\,x\right)^2 = \sum_{j=1}^{M}d_j\left([A]^j\,x\right)^2, \tag{B.4}$$

from which we get:

$$x^{\mathrm{T}}A^{\mathrm{T}}DAx \leq d_{\max}\left(x^{T}A^{\mathrm{T}}Ax\right), \tag{B.5}$$

$$x^{\mathrm{T}}A^{\mathrm{T}}DAx \geq d_{\min}\left(x^{T}A^{\mathrm{T}}Ax\right). \tag{B.6}$$

We then get the desired inequalities by applying (B.5) to (B.2) and (B.6) to (B.3) and factoring out the constants $d_{\max}$ and $d_{\min}$, respectively.

□

*Proof of Theorem 5.2.7.* Defining the following quantities:

$$\tilde{G} = \text{ diag}\left\{\tilde{G}_1, \ldots, \tilde{G}_K\right\} \tag{B.7}$$

$$\Gamma^{\tilde{G}} = \text{ diag}\left\{\gamma_1 I_{N_1^{\mathsf{C}}}, \ldots, \gamma_K I_{N_K^{\mathsf{C}}}\right\} \tag{B.8}$$

$$\Gamma^{B} = \text{ diag}\left\{\gamma_1 I_{N_1^{\mathsf{O}}} \ldots, \gamma_K I_{N_K^{\mathsf{O}}}\right\} \tag{B.9}$$

$$\hat{\Gamma}^{G}[\tau] = \text{ diag}\left\{\hat{\gamma}_1[\tau] I_{N_1^{C}}, ..., \hat{\gamma}_K[\tau] I_{N_K^{C}}\right\} \tag{B.10}$$

$$\hat{\Gamma}^{B}[\tau] = \text{ diag}\left\{\hat{\gamma}_1[\tau] I_{N_1^{C}}, ..., \hat{\gamma}_K[\tau] I_{N_K^{C}}\right\} \tag{B.11}$$

$$\hat{\Lambda}^{G}[\tau] = \Gamma^{G}[\tau] - \hat{\Gamma}^{G}[\tau] \tag{B.12}$$

$$\hat{\Lambda}^{B}[\tau] = \Gamma^{B}[\tau] - \hat{\Gamma}^{B}[\tau] \tag{B.13}$$

$$V_1 = \tilde{G}^T\tilde{G} + B^T B, \tag{B.14}$$

$$\tilde{V}_1\left(\Gamma\right) = \tilde{G}^T\Gamma^{\tilde{G}}\tilde{G} + B^T\Gamma^{B}B \tag{B.15}$$

$$\tilde{V}_2\left(\rho, \Gamma\right) = \frac{1}{\rho^2} I_{N^{\mathsf{T}}} - \tilde{V}_1\left(\Gamma\right) \tag{B.16}$$

$$\tilde{V}_3\left(\hat{\Lambda}[\tau]\right) = \tilde{G}^T\hat{\Lambda}^{\tilde{G}}[\tau]\tilde{G} + B^T\hat{\Lambda}^{B}[\tau]B \tag{B.17}$$

We shall proceed in a similar fashion to the proofs of Lemmas 2.1 and 2.2 in [70] for a). We will then proceed in a similar fashion to the proof strategies of Theorems 1 and 2 and Theorem 3 of [69] for b) and c), respectively.

a): From dual update equations of (5.32) and (5.35), we have:

$$\mu[\tau] = \rho\left(\Gamma^{\tilde{G}}\tilde{G}\right)\left[\sum_{s=0}^{\tau} a[s]\right], \tag{B.18}$$

$$\nu[\tau] = \rho\left(\Gamma^{B}B\right)\left[\sum_{s=0}^{\tau} a[s]\right]. \tag{B.19}$$

From the definition of the iteration-varying $\hat{\gamma}_j[\tau]$ given in Section IV.A we have

$$\hat{\gamma}_j[\tau] = \beta_{\mathbf{j}}\gamma_j \Rightarrow \beta_{\mathbf{j}}[\tau] < 1, \tag{B.20}$$

$$\hat{\gamma}_j[\tau] - \gamma_j = 1 - \beta_{\mathbf{j}}[\tau] \Rightarrow 1 - \beta_{\mathbf{j}}[\tau] < 0. \tag{B.21}$$

Using both (B.18) , (B.19) and defining $\hat{\Lambda}[\tau] = \operatorname{diag}\left\{(1-\beta_{\mathbf{1}}[\tau])\, I_{|C_1|}, ..., (1-\beta_{\mathbf{K}}[\tau])\, I_{|C_k|}\right\} = \hat{\Gamma}[\tau] - \Gamma$ in conjunction with the necessary condition for optimality of (5.31), we have:

$$\begin{aligned} a\left[\tau+1\right] = \underset{a \in \mathbb{R}^{|T|}}{\operatorname{argmin}} &\left\{ \begin{array}{c} \hat{f}\left(a\right) + \left(\bar{\mu}\left[\tau\right]\right)^T \tilde{G}a \\ \left(\bar{\nu}\left[\tau\right]\right)^T Ba + \frac{1}{2\rho}\left\|a - a\left[\tau\right]\right\|_2^2 \end{array} \right\} \Rightarrow 0_{N^{\mathsf{T}}} = c\frac{1}{\rho}c_{\hat{f}}^{[\tau+1]} + \tilde{V}_1\left(\Gamma\right)\left[\sum_{s=0}^{\tau+1} a[s]\right] \\ &+\tilde{V}_2\left(\rho,\Gamma\right)\left(a\left[\tau+1\right] - a\left[\tau\right]\right) + \tilde{V}_3\left(\hat{\Lambda}[\tau]\right) a\left[\tau\right] \\ &< \frac{1}{\rho}c_{\hat{f}}^{[\tau+1]} + \rho\hat{V}_1\left(\Gamma\right)\sum_{s=0}^{\tau} a[s] + \tilde{V}_2\left(\rho,\Gamma\right)\left(a\left[\tau+1\right] - a\left[\tau\right]\right) \end{aligned} \tag{B.22}$$

where $c_{\hat{f}}^{[\tau+1]} \in \partial\hat{f}\left(a\left[\tau+1\right]\right)$. We note that $\tilde{V}_1\left(\Gamma\right)$ is a p.s.d. matrix and $\tilde{V}_3\left(\hat{\Lambda}\right)$ is negative definite since $1 - \beta_{\mathbf{j}}[\tau] < 0$ , while Assumption 5.2.3 gives us:

$$1 > \rho^2 \lambda_{\max}\left(\tilde{V}_1\left(\Gamma\right)\right) \Rightarrow \lambda_{\min}\left(\frac{1}{\rho^2} I_{N^{\mathsf{T}}} - \tilde{V}_1\left(\Gamma\right)\right) > 0 \Rightarrow \lambda_{\min}\left(\tilde{V}_2\left(\rho,\Gamma\right)\right) > 0 \tag{B.23}$$

from which we get that $\tilde{V}_2\left(\rho,\Gamma\right)$ is p.d.. Since $\tilde{V}_1\left(\Gamma\right)$ is p.s.d when $\gamma_j \succ 0$ for each $j \in K$, we can then define its symmetric "square root" $R^{\Gamma} \in \mathbb{R}^{|T|\times|T|}$ according to:

$$R^{\Gamma}R^{\Gamma} = \tilde{V}_1\left(\Gamma\right) = \tilde{G}^T\Gamma^{\tilde{G}}\tilde{G} + B^T\Gamma^B B^1.$$

Connecting $R^{\Gamma}$ to $a^*$ and the ASO-feasibility of (5.2), we have:

$$a \text{ is ASO-feasible iff } R^{\Gamma}a = 0_{N^{\mathsf{T}}}. \tag{B.24}$$

Sufficiency of (B.24) follows directly while necessity of (B.24) follows from:

$$\begin{aligned} R^{\Gamma}a = 0_{N^{\mathsf{T}}} \Rightarrow R^{\Gamma}R^{\Gamma}a = 0_{N^{\mathsf{T}}} \Rightarrow \tilde{G}^T\Gamma^{\tilde{G}}\tilde{G}a = -B^T\Gamma^B Ba \Rightarrow \left\|\tilde{G}a\right\|_{\Gamma^{\tilde{G}}}^2 = -\left\|Ba\right\|_{\Gamma^B}^2 \\ \Rightarrow \tilde{G}a = 0_{N^{\mathsf{C}}} \text{ and } Ba = 0_{N^{\mathsf{O}}} \Rightarrow a \text{ is ASO-feasible.} \end{aligned}$$

Using (B.24), we see that the ASO of (5.2) is equivalent to the following *auxiliary standard optimization (XSO)* problem:

$$\min_{\rho R^{\Gamma}a = 0_{\tilde{N}}} \left\{\hat{f}\left(a\right)\right\}, \tag{B.25}$$

[1] We note that such a *root matrix* exists for $\tilde{V}_1\left(\Gamma\right)$ since the latter is square p.s.d. matrix. Specifically, if $\tilde{V}_1\left(\Gamma\right) = U\Sigma U^T$ is a suitable eigenvalue decomposition, then $R^{\Gamma} = \left[U\Sigma^{\frac{1}{2}}U^T\right]$. See the 7th footnote on pg. 5086 of [69] for more information.

whose necessary condition for optimality satisfies:

$$R^{\Gamma} r^{*} + \frac{1}{\rho} c_{\hat{f}}^{*} = 0_{N^{\intercal}}, \tag{B.26}$$

where $r^*$ is the optimal dual variable, $a^*$ is ASO-optimal, and $c_{\hat{f}}^{*} \in \partial \hat{f}(a^*)$. Since (B.25) is equivalent to (5.2), we have that the dual summands of the associated Lagrangians for both are equal for all $\tau > 0$ and, using (B.18) and (B.19), we have:

$$\begin{aligned}(r[\tau])^T \left(\rho R^{\Gamma}\right) a[\tau] &= (\mu[\tau])^T \tilde{G} a[\tau] + (\nu[\tau])^T B a[\tau] = \left(\sum_{s=0}^{\tau} a[s]\right)^T \left(\rho \tilde{V}_1(\Gamma)\right) a[\tau] \\ &= \left(\sum_{s=0}^{\tau} a[s]\right)^T \left(\rho R^{\Gamma} R^{\Gamma}\right) a[\tau] = \left(R^{\Gamma}\left[\sum_{s=0}^{\tau} a[s]\right]\right)^T \left(\rho R^{\Gamma}\right) a[\tau].\end{aligned} \tag{B.27}$$

from which we conclude that $r[\tau] \triangleq R^{\Gamma}\left(\sum_{s=0}^{\tau} a[s]\right)$. Using (B.26) and noting that:

$$\begin{aligned}(a^* - a[\tau+1])^T c_{\hat{f}}^{*} &= (a^* - a[\tau+1])^T \left[-\left(\rho R^{\Gamma}\right) r^*\right] \\ &= \rho (r^*)^T R^{\Gamma} a[\tau+1] = \rho (r^*)^T (r[\tau+1] - r[\tau])\end{aligned} \tag{B.28}$$

we can then exploit the monotonicity property of the subdifferential operator $\partial \hat{f}$ and use (B.22) to obtain:

$$\begin{aligned}0 &\leq (a[\tau+1] - a^*)^T \left(\frac{1}{\rho} c_{\hat{f}}^{[\tau+1]} - \frac{1}{\rho} c_{\hat{f}}^{*}\right) + (r^*)^T (r[\tau+1] - r[\tau]) \\ &\leq (a[t+1] - a^*)^T \left[-\tilde{V}_1(\Gamma)\left[\sum_{s=0}^{\tau+1} a[s]\right]\right. \\ &\quad \left. - \tilde{V}_2(\rho, \Gamma)(a[\tau+1] - a[\tau])\right] + (r^*)^T (r[\tau+1] - r[\tau]) \\ &= -(r[\tau+1] - r[\tau])^T r[\tau+1] - (a[\tau+1] - a^*)^T \tilde{V}_2(\rho, \Gamma)(a[\tau+1] - a[\tau]) \\ &\quad + (r^*)^T (r[\tau+1] - r[\tau]) \\ &= (a[\tau+1] - a[\tau])^T \tilde{V}_2(\rho, \Gamma)(a^* - a[\tau+1]) + (r[\tau+1] - r[\tau])^T I_{N^{\intercal}} (r^* - r[\tau+1]).\end{aligned} \tag{B.29}$$

Using the identity for $Z \in \mathbb{R}^{n \times n}$ and $z_1, z_2, z_3 \in \mathbb{R}^n$:

$$2(z_1 - z_2)^T Z (z_3 - z_1) = \|z_2 - z_3\|_Z^2 - \|z_1 - z_3\|_Z^2 - \|z_2 - z_1\|_Z^2, \tag{B.30}$$

we then have that (B.29) is equivalent to:

$$0 \leq \|q[\tau] - q^*\|_W^2 - \|q[\tau+1] - q^*\|_W^2 - \|q[\tau] - q[\tau+1]\|_W^2, \tag{B.31}$$

where:

$$W = \begin{bmatrix} \tilde{V}_2(\rho,\Gamma) \, 0 \\ 0 I_{N^\intercal} \end{bmatrix} \tag{B.32}$$

$$q[\tau] = [a[\tau]; r[\tau]] \tag{B.33}$$

$$q^* = [a^*; r^*]. \tag{B.34}$$

From (B.31), we then have, for any $q[\tau+1] \neq q[\tau]$:

$$\|q[\tau] - q^*\|_W^2 \geq \|q[\tau+1] - q^*\|_W^2 + \|q[\tau] - q[\tau+1]\|_W^2 > \|q[\tau+1] - q^*\|_W^2. \tag{B.35}$$

where we used that $W$ is p.d..[2] From (B.35), we have that $\left\{\|q[\tau] - q^*\|_W^2\right\}_{\tau>0}$ is monotonically decreasing. Since this sequence has a lower bound of $0$, we can thus conclude that it converges to a fixed value:

$$\|q[\tau] - q^*\|_W^2 \underset{\tau\to\infty}{\to} \zeta \in R_+. \tag{B.36}$$

From (B.36) we have that $\{q[\tau]\}_{\tau\geq 0}$ deterministically converges to a vector $\overline{q}$ which is a distance $\zeta$ away from $q^*$. Due to the nature of PAC, we can then conclude that $\overline{q}$ is a fixed point of PAC and also satisfies the KKT Conditions of the XSO (B.25). By Assumption 5.2.1, $\hat{f}$ is convex and $\operatorname{dom}\left(\hat{f}\right) = \mathbb{R}^N$. Utilizing this along with Assumption 5.2.2, we can then conclude that Slater's Condition is satisfied. Hence, $\overline{q}$ is an optimal point of the XSO (B.25). Since our choice of choice $q^*$ was initially arbitrary in the sense that it only had to be an (not necessarily the) optimal solution of the XSO, we can then simply set $q^* = \overline{q}$ and therefore conclude that the PAC trajectory obtains asymptotic convergence to an optimal XSO solution:

$$\lim_{\tau\to\infty} \{q[\tau]\} = q^* \Rightarrow \lim_{\tau\to\infty} \{a[\tau]\} = a^*.$$

As for the rate, we can follow the approach of Lemma 2.2 of [88] and use the monotonicity of $\partial\hat{f}$ and the first-order condition of optimality to get, if $\Delta a[\tau+1] \triangleq a[\tau+1] - a[\tau]$:

$$\begin{aligned}
0 &\leq (\Delta a[\tau+1])^T \left(\frac{1}{\rho} c_{\hat{f}}^{[\tau+1]} - \frac{1}{\rho} c_{\hat{f}}^{[\tau]}\right) \\
&= (\Delta a[\tau+1])^T \left(-\left[R^\Gamma r[\tau+1] + V_2(\rho,\Gamma)\Delta a[\tau+1]\right] + \left[R^\Gamma r[\tau] + V_2(\rho,\Gamma)\Delta a[\tau]\right]\right) \\
&= -(\Delta a[\tau+1])^T V_2(\rho,\Gamma)(\Delta a[\tau+1] - \Delta a[\tau]) - (\Delta a[\tau+1])^T R^\Gamma (\Delta r[\tau+1]) \\
&= -(\Delta a[\tau+1])^T V_2(\rho,\Gamma)(\Delta a[\tau+1] - \Delta a[\tau]) - (\Delta r[\tau+1])^T (\Delta r[\tau+1] - \Delta r[\tau]) \\
&= -\|\Delta q[\tau+1]\|_W^2 + (\Delta a[\tau+1])^T V_2(\rho,\Gamma)(\Delta a[\tau]) + (\Delta r[\tau+1])^T (\Delta r[\tau]) \\
&\leq -\|\Delta q[\tau+1]\|_W^2 + \frac{1}{2}\left[\|\Delta q[\tau+1]\|_W^2 + \|\Delta q[\tau]\|_W^2\right] = -\frac{1}{2}\|\Delta q[\tau+1]\|_W^2 + \frac{1}{2}\|\Delta q[\tau]\|_W^2,
\end{aligned} \tag{B.37}$$

[2] This follows since $\tilde{V}_2(\rho,\Gamma)$ is p.d..

where we used the identity for $Z \in \mathbb{R}^{n\times n}$ and $z_1, z_2 \in \mathbb{R}^n$:

$$2\left(z_1\right)^T Z\left(z_2\right) \leq \|z_1\|_Z^2 + \|z_2\|_Z^2 .$$

From (B.37) we can thus conclude that $\|\Delta q\left[\tau\right]\|_W^2$ monotonically non-increasing in $\tau$, i.e.:

$$\|\Delta q\left[\tau+1\right]\|_W^2 \leq \|\Delta q\left[\tau\right]\|_W^2 .$$

Since $\lim\limits_{\tau\to\infty}\left\{q\left[\tau\right]\right\} = q^*$, we can conclude using (B.31) that $\sum_{\tau=1}^{\infty}\|\Delta q\left[\tau\right]\|_W^2 < 0$. Hence, we use Lemma 1.1 of [88] to determine that PAC converges with $\|\Delta q\left[\tau+1\right]\|_W^2 = o\left(\frac{1}{\tau}\right)$. Thus, a) is proven.

b): We begin by showing that we can bound both the cost and infeasibility of the ergodic cost. We then exploit the structure of the associated XSO from (B.25) to bound the ergodic cost and ergodic infeasibility even further.

So, we start by letting $a^*$ be the optimal ASO solution that the PAC trajectory converges to, as proven in in part a). We continue by using the first-order property of convex functions $\hat{f}$ to get, for some dual variable $r \in \mathbb{R}^{N^{\mathrm{T}}}$ and $c_{\hat{f}}^{[\tau+1]} \in \partial\hat{f}\left(a\left[\tau+1\right]\right)$:

$$\hat{f}\left(a\left[\tau+1\right]\right) + \left(a^* - a\left[\tau+1\right]\right)^T c_{\hat{f}}^{[\tau+1]} \leq \hat{f}\left(a^*\right)$$

$$\Downarrow$$

$$\left[\hat{f}\left(a\left[\tau+1\right]\right) - \hat{f}\left(a^*\right)\right] + \rho r^T R^{\Gamma} a\left[\tau+1\right] \;\leq\; \left(a\left[\tau+1\right] - a^*\right)^T c_{\hat{f}}^{[\tau+1]} + \rho r^T R^{\Gamma} a\left[\tau+1\right]$$

$$\Downarrow$$

$$\begin{aligned}\tfrac{2}{\rho}\left[\hat{f}\left(a\left[\tau+1\right]\right) - \hat{f}\left(a^*\right)\right] + 2r^T R^{\Gamma} a\left[\tau+1\right] \;\leq\; \\ \|q\left[\tau\right] - q\|_W^2 - \|q\left[\tau+1\right] - q\|_W^2 - \|q\left[\tau+1\right] - q\left[\tau\right]\|_W^2\end{aligned} \tag{B.38}$$

$$\Downarrow$$

$$\tfrac{2}{\rho}\left[\hat{f}\left(a\left[\tau+1\right]\right) - \hat{f}\left(a^*\right)\right] + 2r^T R^{\Gamma} a\left[\tau+1\right] \;\leq\; \|q\left[\tau\right] - q\|_W^2 - \|q\left[\tau+1\right] - q\|_W^2 \;, \tag{B.39}$$

where $q = \left[a^*; r\right]$ and $W$ is p.s.d..[3] We then sum (B.39) from $0$ to $T-1$, using Jensen's Inequality on the LHS and using the non-negativeness of the $l_2$ metric for the RHS, to get:

$$\sum_{\tau=0}^{T-1}\left(\;\tfrac{2}{\rho}\left[\hat{f}\left(a\left[\tau+1\right]\right) - \hat{f}\left(a^*\right)\right] + 2r^T R^{\Gamma} a\left[\tau+1\right]\;\right) \leq \sum_{\tau=0}^{T-1}\left(\begin{array}{c}\|q\left[\tau\right] - q\|_W^2 \\ -\|q\left[\tau+1\right] - q\|_W^2\end{array}\right)$$

$$\Downarrow$$

$$\hat{f}\left(\bar{a}\left[T\right]\right) - \hat{f}\left(a^*\right) + \rho r^T R^{\Gamma}\bar{a}\left[T\right] \leq= \left(\frac{\rho}{2T}\right)\|q\|_W^2 , \tag{B.40}$$

[3]We can conclude that $W$ is p.s.d. from Assumption 5.2.4 using similar reasoning to how we concluded that $W$ is p.d. from Assumption 5.2.3 in part a).

for $\bar{a}[T] = \frac{1}{T}\sum_{s=1}^{T} a[s]$ and zero initialized $q[0] = [0_{N^\intercal}; 0_{N^\intercal}]$. Our goal now is to bound the absolute value of the first two summands in LHS of (B.40) and then bound the absolute value of the third summand in (B.40). As such, we first set $r = 0_{N^\intercal}$ in (B.40) and obtain:

$$\hat{f}(\bar{a}[T]) - \hat{f}(a^*) \leq \left(\frac{\rho}{2T}\right) \|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} . \tag{B.41}$$

Using the saddle point inequality associated with the XSO of (B.25) we also have:

$$\hat{f}(a^*) \leq \hat{f}(\bar{a}[T]) + \rho (r^*)^T R^\Gamma \bar{a}[T] \tag{B.42}$$

$$\Downarrow$$

$$\rho (r^*)^T R^\Gamma \bar{a}[T] \leq \hat{f}(\bar{a}[T]) - \hat{f}(a^*) + \rho (2r^*)^T R^\Gamma \bar{a}[T] . \tag{B.43}$$

Applying (B.40) to (B.43) with $r = 2r^*$, we then have:

$$\rho (r^*)^T R^\Gamma \bar{a}[T] \leq \left(\frac{\rho}{2T}\right) \left[\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + \|2r^*\|^2\right] . \tag{B.44}$$

Combining (B.44) with (B.42), we then have:

$$\hat{f}(a^*) - \hat{f}(\bar{a}[T]) \leq \left(\frac{\rho}{2T}\right) \left[\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + \|2r^*\|^2\right] ,$$

from which we can conclude that:

$$\left|\hat{f}(a^*) - \hat{f}(\bar{a}[T])\right| \leq \left(\frac{\rho}{2T}\right) \left[\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + 4\|r^*\|^2_W\right] . \tag{B.45}$$

We can also bound the ergodic infeasibility $R^\Gamma \bar{a}[T]$ in a similar fashion. Letting $r = r^* + \frac{R^\Gamma \bar{a}[T]}{\|R^\Gamma \bar{a}[T]\|_2}$, we have from (B.40):

$$\begin{aligned} \hat{f}(\bar{a}[T]) - \hat{f}(a^*) + \rho (r^*)^T R^\Gamma \bar{a}[T] + \rho \left\|R^\Gamma \bar{a}[T]\right\|_2 \; \leq \\ \left(\tfrac{\rho}{2T}\right) \left[\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + \left\|r^* + \tfrac{R^\Gamma \bar{a}[T]}{\|R^\Gamma \bar{a}[T]\|_2}\right\|_2^2\right] . \end{aligned} \tag{B.46}$$

Applying the saddle-point inequality of (B.25) to (B), we then have:

$$\begin{aligned} \left\|R^\Gamma \bar{a}[T]\right\|_2 \leq \left(\frac{1}{2T}\right) \left[\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + \left\|r^* + \frac{R^\Gamma \bar{a}[T]}{\|R^\Gamma \bar{a}[T]\|_2}\right\|_2^2\right] \leq \\ \left(\frac{1}{2T}\right) \left[\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} + 2\left(\|r^*\|_2^2 + 1\right)\right] . \end{aligned} \tag{B.47}$$

where we used that $\|w + z\|_2^2 \leq 2\|w\|_2^2 + 2\|z\|_2^2$ for any $w, z \in \mathbb{R}^{|T|}$. Finally, to further bound (B.45) and (B.48) we note that:

$$\|a^*\|^2_{\tilde{V}_2(\rho,\Gamma)} \leq \lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right) \|a^*\|_2^2 = \left[\frac{1}{\rho^2} - \lambda_{\min}\left(\tilde{V}_1(\Gamma)\right)\right] \|a^*\|_2^2 \leq \frac{1}{\rho^2}\|a^*\|_2^2 , \tag{B.48}$$

and, using (B.26) in conjunction with Lemma 8 of [69] and Lemma B.0.1.b:

$$\|r^*\|_2^2 \leq \frac{\Phi^2}{\rho^2 \hat{\lambda}_{\min}\left(\tilde{V}_1(\Gamma)\right)} \leq \frac{\Phi^2}{\rho^2 \gamma_{min} \hat{\lambda}_{\min}(V_1)}, \tag{B.49}$$

where $\Phi = \sup_{z \in \partial \hat{f}(a^*)} \{\|z\|_2\}$. Applying (B.48) and (B.49) to (B.45) and (B), we get:

$$\left|\hat{f}(a^*) - \hat{f}(\bar{a}[T])\right| \leq \left(\frac{1}{\rho T}\right)\left[\frac{\|a^*\|_2^2}{2} + \frac{2\Phi^2}{\gamma_{\min} \hat{\lambda}_{\min}(V_1)}\right],$$

and:

$$\left\|R^\Gamma \bar{a}[T]\right\|_2 \leq \left(\frac{1}{\rho^2 T}\right)\left[\|a^*\|^2 + 2\left(\frac{\Phi^2}{\gamma_{\min} \hat{\lambda}_{\min}(V_1)} + \rho^2\right)\right].$$

c) We shall proceed in a similar fashion to the proof of Theorems 3 of [69]. We once again start by letting $a^*$ be the optimal ASO solution that the PAC trajectory converges to, as proven in in part a). Since $\hat{f}$ is $\alpha$-strongly convex and $L$-strongly continuous, we can use the following inequality:[4]

$$\begin{gathered}
\min\left\{\alpha \|a[\tau+1] - a^*\|_2^2, \frac{1}{L}\left\|\nabla \hat{f}(a[\tau+1]) - \nabla \hat{f}(a^*)\right\|_2^2\right\} \leq \\
\left(\nabla \hat{f}(a[\tau+1]) - \nabla \hat{f}(a^*)\right)^T (a[\tau+1] - a^*) \\
\Updownarrow \\
\theta\alpha \|a[\tau+1] - a^*\|_2^2 + \frac{(1-\theta)}{L}\left\|\nabla \hat{f}(a[\tau+1]) - \nabla \hat{f}(a^*)\right\|_2^2 \leq \\
\left(\nabla \hat{f}(a[\tau+1]) - \nabla \hat{f}(a^*)\right)^T (a[\tau+1] - a^*).
\end{gathered} \tag{B.50}$$

for all $\theta \in [0,1]$. Using (B.50) in conjunction with (B.29) and (B.31), we have:

$$\begin{gathered}
\frac{2\theta\alpha}{\rho}\|a[\tau+1] - a^*\|_2^2 + \frac{2(1-\theta)}{\rho L}\left\|\nabla \hat{f}(a[\tau+1]) - \nabla \hat{f}(a^*)\right\|_2^2 \leq \\
\|q[\tau] - q^*\|_W^2 - \|q[\tau+1] - q^*\|_W^2
\end{gathered} \tag{B.51}$$

Letting $\delta \geq 0$ satisfy:

$$\delta \leq \min\left\{\frac{2\alpha\theta\gamma_{min}\hat{\lambda}_{\min}(V_1)}{\rho\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\left(\gamma_{min}\hat{\lambda}_{\min}(V_1) + 2\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\right)},\ \frac{\rho\gamma_{min}\hat{\lambda}_{\min}(V_1)(1-\theta)}{L}\right\}, \tag{B.52}$$

[4]These inequalities are depicted in Lemma 3 in [69]

we have from (B.26) and (B.29):

$$\begin{aligned}
\delta \left\|r[\tau+1]-r^*\right\|_2^2 &\leq \left(\frac{\delta}{\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right)\left\|r[\tau+1]-r^*\right\|_{\tilde{V}_1(\Gamma)}^2 \\
&= \left(\frac{\delta}{\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right)\left\|R^{\Gamma}\left(r[\tau+1]-r^*\right)\right\|_2^2 \\
&= \left(\frac{\delta}{\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right)\left\|\frac{1}{\rho}\left(\nabla\hat{f}(a[\tau+1])-\nabla\hat{f}(a^*)\right)+\tilde{V}_2(\rho,\Gamma)(a[\tau+1]-a[\tau])\right\|_2^2 \\
&\leq \left(\frac{2\delta}{\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right)\left[\frac{1}{\rho^2}\left\|\nabla\hat{f}(a[\tau+1])-\nabla\hat{f}(a^*)\right\|_2^2+\left\|\tilde{V}_2(\rho,\Gamma)(a[\tau+1]-a[t])\right\|_2^2\right] \\
&\leq \left(\frac{2\delta}{\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right)\left[\frac{1}{\rho^2}\left\|\nabla\hat{f}(a[\tau+1])-\nabla\hat{f}(a^*)\right\|_2^2\right. \\
&\quad \left.+\left(\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\right)^2\left\|a[\tau+1]-a[\tau]\right\|_2^2\right]. \qquad \text{(B.53)}
\end{aligned}$$

Using the second RHS component of (B.52) with the first RHS summand of (B.53), we have:

$$\left[\frac{2\delta}{\rho^2\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right]\left\|\nabla\hat{f}(a[\tau+1])-\nabla\hat{f}(a^*)\right\|_2^2 \leq \left[\frac{2(1-\theta)}{\rho L}\right]\left\|\nabla\hat{f}(a[\tau+1])-\nabla\hat{f}(a^*)\right\|_2^2. \qquad \text{(B.54)}$$

Similarly, using the first RHS component of (B.52) with the second summand of (B.53), we have:

$$\begin{aligned}
&\left[\frac{2\left(\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\right)^2\delta}{\gamma_{min}\hat{\lambda}_{\min}(V_1)}\right]\left\|a[\tau+1]-a[\tau]\right\|_2^2 \leq \\
&\left[\frac{4\alpha\theta\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)}{\rho\left(\gamma_{min}\hat{\lambda}_{\min}(V_1)+2\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\right)}\right]\left\|a[\tau+1]-a[\tau]\right\|_2^2 \\
&= \left[\frac{2\alpha\theta}{\rho}-\frac{2\alpha\theta\gamma_{min}\hat{\lambda}_{\min}(V_1)}{\rho\left(\gamma_{min}\hat{\lambda}_{\min}(V_1)+2\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\right)}\right]\left\|a[t+1]-a[t]\right\|_2^2 \\
&\leq \left[\frac{2\alpha\theta}{\rho}-\delta\lambda_{\max}\left(\tilde{V}_2(\rho,\Gamma)\right)\right]\left\|a[\tau+1]-a[\tau]\right\|_2^2 \leq \\
&\left(\frac{2\alpha\theta}{\rho}\right)\left\|a[\tau+1]-a[t]\right\|_2^2-\delta\left\|a[\tau+1]-a[\tau]\right\|_{\tilde{V}_2(\rho,\Gamma)}^2 \qquad \text{(B.55)}
\end{aligned}$$

Combining (B.54) and (B.55) with (B.53), we have:

$$\delta \left\| q\left[\tau+1\right] - q^* \right\|_W^2 \leq \left(\frac{2\alpha\theta}{\rho}\right) \left\| a\left[\tau+1\right] - a\left[t\right] \right\|_2^2 + \left(\frac{2\left(1-\theta\right)}{\rho L}\right) \left\| \nabla \hat{f}\left(a\left[\tau+1\right]\right) - \nabla \hat{f}\left(a^*\right) \right\|_2^2. \tag{B.56}$$

Using (B.51) on (B.56), we get:

$$\delta \left\| q\left[\tau+1\right] - q^* \right\|_W^2 \leq \left\| q\left[\tau\right] - q^* \right\|_W^2 - \left\| q\left[\tau+1\right] - q^* \right\|_W^2$$

$$\Updownarrow$$

$$\left\| q\left[\tau+1\right] - q^* \right\|_W^2 \leq \left(\frac{1}{1+\delta}\right) \left\| q\left[\tau\right] - q^* \right\|_W^2 \tag{B.57}$$

Iterating on (B.57) and using the zero-initialization assumption, we then obtain:

$$\left\| q\left[\tau+1\right] - q^* \right\|_W^2 \leq \left(\frac{1}{1+\delta}\right)^{\tau+1} \left\| q^* \right\|_W^2. \tag{B.58}$$

Bounding the LHS of (B.58) further, we obtain the desired:

$$\left\| a\left[\tau+1\right] - a^* \right\|_{\tilde{V}_2(\rho,\Gamma)}^2 \leq \left(\frac{1}{1+\delta}\right)^{\tau+1} \left\| q^* \right\|_W^2, \tag{B.59}$$

$$\left\| r\left[\tau+1\right] - r^* \right\|_2^2 \leq \left(\frac{1}{1+\delta}\right)^{\tau+1} \left\| q^* \right\|_W^2. \tag{B.60}$$

We can actually further bound the RHS of (B.58) using (B.48) and (B.49):

$$\begin{aligned} \left(\frac{1}{1+\delta}\right)^{\tau+1} \left\| q^* \right\|_W^2 &= \left(\frac{1}{1+\delta}\right)^{\tau+1} \left[ \left\| a^* \right\|_{\tilde{V}_2(\rho,\Gamma)}^2 + \left\| r^* \right\|_2^2 \right] \\ &\leq \left(\frac{1}{1+\delta}\right)^{\tau+1} \left[ \frac{1}{\rho^2} \left\| a^* \right\|_2^2 + \frac{\left\| \nabla \hat{f}\left(a^*\right) \right\|_2^2}{\rho^2 \gamma_{min} \hat{\lambda}_{\min}\left(V_1\right)} \right] \end{aligned} \tag{B.61}$$

Finally, we can determine the optimal $\theta^*$ to use so as to maximize $\delta$. This is accomplished easily enough by setting both terms of (B.52) equal to one another and solving for $\theta^*$:

$$\theta^* = \frac{\gamma_{min} \hat{\lambda}_{\min}\left(V_1\right) + 2\rho^{-2}}{2\alpha L + \gamma_{min} \hat{\lambda}_{\min}\left(V_1\right) + 2\rho^{-2}}. \tag{B.62}$$

Plugging $\theta^*$ into (B.52), we get the following $\delta^*$ rate:

$$\delta^* = \left(\frac{\rho \gamma_{min} \hat{\lambda}_{\min}\left(V_1\right)}{L}\right) \left(1 - \theta^*\right) = \frac{2\alpha\rho\gamma_{min} \hat{\lambda}_{\min}\left(V_1\right)}{2\alpha L + \gamma_{min} \hat{\lambda}_{\min}\left(V_1\right) + 2\rho^{-2}}.$$

□

*Proof of Theorem 5.2.8.*
a) In this part, we want to choose PAC parameters $\rho$ and $\gamma_{min}$ to maximize the ergodic convergence rate (5.44):

$$\xi_{\text{E1}}\left(\rho, \gamma_{min}\right) = \frac{\rho}{2}\left[\left\|a^*\right\|_2^2 + \frac{4\Phi^2}{\gamma_{min}\hat{\lambda}_{\min}\left(V_1\right)}\right]^{-1}, \tag{B.63}$$

subject to Assumption 5.2.4:

$$1 \geq \rho^2\gamma_{min}\lambda_{\max}\left(V_1\right). \tag{B.64}$$

Noting that (B.63) is directly proportional to $\rho$, we have from (B.64):

$$\rho\left(\gamma_{min}\right) = \left(\sqrt{\gamma_{min}\lambda_{\max}\left(V_1\right)}\right)^{-1} \tag{B.65}$$

Plugging (B.65) into (B.63) and then optimizing over $\gamma_{min}$ by setting the partial of (B.63) w.r.t. $\gamma_{min}$ to zero, we get:

$$\xi_{\text{E1}}^* = \frac{\sqrt{\psi\left(\mathcal{D}\right)}}{2\Phi\left\|a^*\right\|_2}.$$

where:

$$\gamma_{\text{E1}}^* = \frac{4\Phi^2}{\hat{\lambda}_{\min}\left(V_1\right)\left\|a^*\right\|_2^2} \tag{B.66}$$

$$\rho_{\text{E1}}^* = \left(\sqrt{\gamma_{\text{E1}}^*\lambda_{\max}\left(V_1\right)}\right)^{-1}. \tag{B.67}$$

b) In this part, we want to choose PAC parameters $\rho$ and $\gamma_{min}$ to maximize the ergodic convergence rate (5.45):

$$\xi_{\text{P}}\left(\rho, \gamma_{min}\right) = \frac{2\alpha\rho\gamma_{min}\hat{\lambda}_{\min}\left(V_1\right)}{2\alpha L + 2\rho^{-2} + \gamma_{min}\hat{\lambda}_{\min}\left(V_1\right)}, \tag{B.68}$$

subject to Assumption 5.2.4:

$$1 \geq \rho^2\gamma_{min}\lambda_{\max}\left(V_1\right) \tag{B.69}$$

Noting that (B.68) is directly proportional to $\rho$, we have from (B.64):

$$\rho\left(\gamma_{min}\right) = \left(\sqrt{\gamma_{min}\lambda_{\max}\left(V_1\right)}\right)^{-1}. \tag{B.70}$$

Plugging (B.70) into (B.68) and then optimizing over $\gamma_{min}$ by setting the partial of (B.68) w.r.t. $\gamma_{min}$ to zero, we get:

$$\xi_{\text{P}}^* = \left(\frac{1}{2\sqrt{\alpha^{-1}L}}\right)\sqrt{\frac{2\hat{\lambda}_{\min}\left(V_1\right)\psi\left(\mathcal{D}\right)}{2\lambda_{\max}\left(V_1\right) + \hat{\lambda}_{\min}\left(V_1\right)}},$$

where:

$$\gamma_{\mathrm{P}}^{*} = \frac{2\alpha L}{2\lambda_{\max}(V_1) + \hat{\lambda}_{\min}(V_1)} \tag{B.71}$$

$$\rho_{\mathrm{P}}^{*} = \left(\sqrt{\gamma_{\mathrm{P}}^{*}\lambda_{\max}(V_1)}\right)^{-1}. \tag{B.72}$$

□

*Proof of Proposition 5.3.1.* In order to prove (5.62), we rewrite (27) as $y = Ax$ where $y = \bar{\nu}_j[\tau]$, $x = \begin{bmatrix} \sum_{s=0}^{\tau} a[s] \\ a[\tau] \end{bmatrix}$ and $A = \begin{bmatrix} \rho\gamma_j [B]^j & \rho\hat{\gamma}_j[\tau] [B]^j \end{bmatrix}$. It is easy to see that $x$ and $y$ are known where as $A$ is unknown. Noting that $A \in \mathbb{R}^{|O_j| \times 2|T|}$ and $y \in \mathbb{R}^{|O_j|}$, and that $2|T| \gg |O_j|$, it is not possible to uniquely identify $A$ using $x$ and $y$. Therefore, $\|\hat{\nu}_{j\to k}[\tau] - \bar{\nu}_j[\tau]\|_2 \not\to 0$ Therefore (5.62) follows, proving Proposition 5.3.1 . □

# Appendix C

# Simulations data

In this Appendix all the data regarding the test networks and the market simulations are reported

## C.1 Modified IEEE 13 Bus system

| Node | Phase 1 [kVA] | Phase 2 [kVA] | Phase 3 [kVA] |
|---|---|---|---|
| 632 | 0+j0 | 0+j0 | 0+j0 |
| 633 | 0+j0 | 0+j0 | 0+j0 |
| 645 | 0+j0 | 170+j125 | 0+j0 |
| 634 | 0+j0 | -(230+j132) | 0+j0 |
| 652 | 0+j0 | 0+j0 | 0+j0 |
| 646 | -(120+j90) | -(120+j90) | -(120+j90) |
| 671 | 385+j220 | 385+j220 | 385+j220 |
| 692 | 485+j190 | 68+j60 | 290+j212 |
| 684 | -(100+j10) | (0+j0) | -(100+j10) |
| 675 | 0+j0 | 0+j0 | 170+j80 |
| 611 | 128+j86 | 0+j0 | 0+j0 |
| 680 | 0+j0 | 0+j0 | 0+j0 |

Table C.1: Active and Reactive power data for the modified IEEE13 nodes test system

## C.2 Modified IEEE 123 Bus system

| Node | Phase 1 [kVA] | Phase 2 [kVA] | Phase 3 [kVA] |
|---|---|---|---|
| 1 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 2 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 3 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 4 | 0+1i*0 | 0+1i*0 | -(35+1i*20) |
| 5 | 0+1i*0 | 0+1i*0 | 50+1i*20 |
| 6 | 0+1i*0 | 0+1i*0 | 60+1i*10 |
| 7 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 8 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 9 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 10 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 11 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 12 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 13 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 14 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 15 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 16 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 17 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 18 | 0+1i*0 | 0+1i*0 | 20+1i*10 |
| 19 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 20 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 21 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 22 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 23 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 24 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 25 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 26 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 27 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 28 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 29 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 30 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 31 | 0+1i*0 | 0+1i*0 | 50+1i*20 |
| 32 | 0+1i*0 | 0+1i*0 | -(20+1i*10) |
| 33 | 0+1i*0 | 0+1i*0 | 65+1i*10 |
| 34 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 35 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 36 | 40+1i*20 | 0+1i*0 | 0+1i*0 |

| Node | Phase 1 [kVA] | Phase 2 [kVA] | Phase 3 [kVA] |
|---|---|---|---|
| 37 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 38 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 39 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 40 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 41 | -(30+1i*10) | -(30+1i*10) | -(30+1i*10) |
| 42 | 0+1i*0 | 0+1i*0 | (50+1i*10) |
| 43 | (-20+1i*10) | 0+1i*0 | 0+1i*0 |
| 44 | 0+1i*0 | -(40+1i*20) | 0+1i*0 |
| 45 | 10+1i*5 | 0+1i*0 | 0+1i*0 |
| 46 | 30+1i*10 | 0+1i*0 | 0+1i*0 |
| 47 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 48 | 35+1i*25 | 35+1i*25 | 35+1i*25 |
| 49 | 70+1i*50 | 70+1i*50 | 70+1i*50 |
| 50 | 35+1i*25 | 70+1i*50 | 35+1i*20 |
| 51 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 52 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 53 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 54 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 55 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 56 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 57 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 58 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 59 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 60 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 61 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 62 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 63 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 64 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 65 | 0+1i*0 | 75+1i*35 | 0+1i*0 |
| 66 | 35+1i*25 | 35+1i*25 | 70+1i*50 |
| 67 | 0+1i*0 | 0+1i*0 | 75+1i*35 |
| 68 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 69 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 70 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 71 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 72 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 73 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 74 | 0+1i*0 | 0+1i*0 | 40+1i*20 |

| Node | Phase 1 [kVA] | Phase 2 [kVA] | Phase 3 [kVA] |
|---|---|---|---|
| 75 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 76 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 77 | 105+1i*80 | 70+1i*50 | 70+1i*50 |
| 78 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 79 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 80 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 81 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 82 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 83 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 84 | 0+1i*0 | 0+1i*0 | 20+1i*10 |
| 85 | 0+1i*0 | 0+1i*0 | 20+1i*10 |
| 86 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 87 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 88 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 89 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 90 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 91 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 92 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 93 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 94 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 95 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 96 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 97 | 0+1i*0 | 20+1i*10 | 0+1i*0 |
| 98 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 99 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 100 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 101 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 102 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 103 | 0+1i*0 | 0+1i*0 | 20+1i*10 |
| 104 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 105 | 0+1i*0 | 0+1i*0 | 40+1i*20 |
| 106 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 107 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 108 | 0+1i*0 | 40+1i*20 | 0+1i*0 |
| 109 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 110 | 40+1i*20 | 0+1i*0 | 0+1i*0 |
| 111 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 112 | 20+1i*10 | 0+1i*0 | 0+1i*0 |

| Node | Phase 1 [kVA] | Phase 2 [kVA] | Phase 3 [kVA] |
|---|---|---|---|
| 113 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 114 | -(70+1i*20) | 0+1i*0 | 0+1i*0 |
| 115 | 20+1i*10 | 0+1i*0 | 0+1i*0 |
| 116 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 117 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 118 | 0+1i*0 | 0+1i*0 | 0+1i*0 |
| 119 | 0+1i*0 | 0+1i*0 | 0+1i*0 |

Table C.2: Active and Reactive power data for the modified IEEE123 nodes test system

# Glossary

**AAC** Active Adaptive Control.

**AC** Alternate current.

**ADMM** Alternating direction method of multipliers.

**ASO** Atomized standard optimization.

**BF** Branch flow.

**BIM** Bus injection model.

**BM** Balancing market.

**CCP** Closed, convex and prope.

**CI** Current injection.

**dADMM** Distributed alternating direction method of multipliers.

**DERs** Distributed energy resources.

**DG** Distributed generation.

**DMM** Dynamic market mechanism.

**DR** Demand response.

**DSO** Distribution system operator.

**GNEP** Generalized nash equilibrium problem.

**GSO** Global standard optimization.

**HV** High voltage.

**KKT** Karush-Kuhn-Tucker.

**MoM** Method of multipliers.

**NEP** Nash equilibrium problem.

**OPF** Optimal power flow.

**PAC** Proximal atomic coordination.

**REM** Retail energy market.

**SDP** Semidefinite programming.

**SOCP** Second order cone programming.

**TSO** Transmission system operator.

# Bibliography


[1] G. Ferro, R. Minciardi, L. Parodi, M. Robba, M. Rossi, *et al.*, “Optimal control of multiple microgrids and buildings by an aggregator,” *Energies*, vol. 13, no. 5, pp. 1–26, 2020.

[2] W. H. Kersting, “Radial distribution test feeders,” *IEEE Transactions on Power Systems*, vol. 6, no. 3, pp. 975–985, 1991.

[3] R. M. R. M. Ferro, G. and A. Annaswamy, “A new convex relaxation of the optimal power-flow problem for distribution grids,” *IEEE Transactions on Power systems*, Under review.

[4] H. W. Dommel and W. F. Tinney, “Optimal power flow solutions,” *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-87, pp. 1866–1876, Oct 1968.

[5] M. Farivar, C. R. Clarke, S. H. Low, and K. M. Chandy, “Inverter var control for distribution systems with renewables,” *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2011.

[6] J. Lavaei and S. H. Low, “Zero duality gap in optimal power flow problem,” *IEEE Transactions on Power Systems*, vol. 27, no. 1, p. 92–107, 2012.

[7] M. Farivar and S. Low, “Branch flow model: Relaxations and convexification,” *2014 IEEE PES TD Conference and Exposition*, 2014.

[8] K. Christakou, D.-C. Tomozei, J.-Y. L. Boudec, and M. Paolone, “Ac opf in radial distribution networks part 1: On the limits of the branch flow convexification and the alternating direction method of multipliers,” *Electric Power Systems Research*, vol. 143, p. 438–450, 2017.

[9] W. H. Kersting, *Distribution system modeling and analysis*. CRC press, 2006.

[10] T.-H. Chen, M.-S. Chen, K.-J. Hwang, P. Kotas, and E. Chebli, “Distribution system power flow analysis-a rigid approach,” *IEEE Transactions on Power Delivery*, vol. 6, no. 3, p. 1146–1152, 1991.

[11] V. D. Costa, N. Martins, and J. Pereira, "Developments in the newton raphson power flow formulation based on current injections," *IEEE Transactions on Power Systems*, vol. 14, no. 4, p. 1320–1326, 1999.

[12] G. P. Mccormick, "Computability of global solutions to factorable nonconvex programs: Part i — convex underestimating problems," *Mathematical Programming*, vol. 10, no. 1, p. 147–175, 1976.

[13] H. Hijazi, C. Coffrin, and P. V. Hentenryck, "Polynomial sdp cuts for optimal power flow," *2016 Power Systems Computation Conference (PSCC)*, 2016.

[14] M. R. Narimani, D. K. Molzahn, and M. L. Crow, "Improving qc relaxations of opf problems via voltage magnitude difference constraints and envelopes for trilinear monomials," *2018 Power Systems Computation Conference (PSCC)*, 2018.

[15] B. Park and C. L. Demarco, "Convex relaxation of sparse tableau formulation for the ac optimal power flow," *Electric Power Systems Research*, vol. 171, p. 209–218, 2019.

[16] Molzahn, Lesieutre, and Demarco, "A sufficient condition for power flow insolvability with applications to voltage stability margins," *IEEE Transactions on Power Systems*, vol. 28, no. 3, p. 2592–2601, 2013.

[17] B. Kocuk, S. S. Dey, and X. A. Sun, "Inexactness of sdp relaxation and valid inequalities for optimal power flow," *IEEE Transactions on Power Systems*, vol. 31, no. 1, p. 642–651, 2016.

[18] S. H. Low, "Convex relaxation of optimal power flow: A tutorial," *2013 IREP Symposium Bulk Power System Dynamics and Control - IX Optimization, Security and Control of the Emerging Power Grid*, 2013.

[19] S. H. Low, "Convex relaxation of optimal power flow—part i: Formulations and equivalence," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, p. 15–27, 2014.

[20] E. Dallanese, H. Zhu, and G. B. Giannakis, "Distributed optimal power flow for smart microgrids," *IEEE Transactions on Smart Grid*, vol. 4, no. 3, p. 1464–1475, 2013.

[21] B. A. Robbins and A. D. Dominguez-Garcia, "Optimal reactive power dispatch for voltage regulation in unbalanced distribution systems," *IEEE Transactions on Power Systems*, vol. 31, no. 4, p. 2903–2913, 2016.

[22] M. Nick, R. Cherkaoui, J.-Y. L. Boudec, and M. Paolone, "An exact convex formulation of the optimal power flow in radial distribution networks including transverse components," *IEEE Transactions on Automatic Control*, vol. 63, no. 3, p. 682–697, 2018.

[23] L. Gan, N. Li, U. Topcu, and S. H. Low, "Exact convex relaxation of optimal power flow in radial networks," *IEEE Transactions on Automatic Control*, vol. 60, pp. 72–87, Jan 2015.

[24] J. Nocedal, A. Wächter, and R. Waltz, "Adaptive barrier update strategies for nonlinear interior methods," *SIAM Journal on Optimization*, vol. 19, no. 4, pp. 1674–1693, 2009.

[25] "http://sites.ieee.org/pes-testfeeders/resources/."

[26] K. P. Schneider, B. A. Mather, B. C. Pal, C.-W. Ten, G. J. Shirek, H. Zhu, J. C. Fuller, J. L. R. Pereira, L. F. Ochoa, L. R. D. Araujo, and et al., "Analytic considerations and design basis for the ieee distribution test feeders," *IEEE Transactions on Power Systems*, vol. 33, no. 3, p. 3181–3188, 2018.

[27] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2019.

[28] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in matlab," in *In Proceedings of the CACSD Conference*, (Taipei, Taiwan), 2004.

[29] R. A. van der Veen and R. A. Hakvoort, "The electricity balancing market: Exploring the design challenge," *Utilities Policy*, vol. 43, pp. 186–194, 2016.

[30] S. Bahrami, M. H. Amini, M. Shafie-khah, and J. P. Catalao, "A decentralized electricity market scheme enabling demand response deployment," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 4218–4227, 2017.

[31] H. Fontenot and B. Dong, "Modeling and control of building-integrated microgrids for optimal energy management–a review," *Applied Energy*, vol. 254, p. 113689, 2019.

[32] D. Li, W.-Y. Chiu, H. Sun, and H. V. Poor, "Multiobjective optimization for demand side management program in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1482–1490, 2017.

[33] F. Verrilli, G. Gambino, S. Srinivasan, G. Palmieri, C. Del Vecchio, and L. Glielmo, "Demand side management for heating controls in microgrids," *IFAC-PapersOnLine*, vol. 49, no. 1, pp. 611–616, 2016.

[34] P. G. Da Silva, D. Ilić, and S. Karnouskos, "The impact of smart grid prosumer grouping on forecasting accuracy and its benefits for local electricity market trading," *IEEE Transactions on Smart Grid*, vol. 5, no. 1, pp. 402–410, 2013.

[35] S. Grijalva and M. U. Tariq, "Prosumer-based smart grid architecture enables a flat, sustainable electricity industry," in *ISGT 2011*, pp. 1–6, IEEE, 2011.

[36] B. Wang, M. Sechilariu, and F. Locment, "Intelligent dc microgrid with smart grid communications: Control strategy consideration and design," *IEEE transactions on smart grid*, vol. 3, no. 4, pp. 2148–2156, 2012.

[37] A. M. Carreiro, H. M. Jorge, and C. H. Antunes, "Energy management systems aggregators: A literature survey," *Renewable and sustainable energy reviews*, vol. 73, pp. 1160–1172, 2017.

[38] K. Ma, C. Wang, J. Yang, C. Hua, and X. Guan, "Pricing mechanism with noncooperative game and revenue sharing contract in electricity market," *IEEE transactions on cybernetics*, vol. 49, no. 1, pp. 97–106, 2017.

[39] C. Olk, D. U. Sauer, and M. Merten, "Bidding strategy for a battery storage in the german secondary balancing power market," *Journal of Energy Storage*, vol. 21, pp. 787–800, 2019.

[40] A. Kiani and A. Annaswamy, "Perturbation analysis of market equilibrium in the presence of renewable energy resources and demand response," in *2010 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT Europe)*, pp. 1–8, IEEE, 2010.

[41] M. J. Garcia, T. R. Nudell, and A. M. Annaswamy, "A dynamic regulation market mechanism for improved financial settlements in wholesale electricity markets," in *2017 American Control Conference (ACC)*, pp. 1425–1430, IEEE, 2017.

[42] S. Huang, Q. Wu, S. S. Oren, R. Li, and Z. Liu, "Distribution locational marginal pricing through quadratic programming for congestion management in distribution networks," *IEEE Transactions on Power Systems*, vol. 30, no. 4, pp. 2170–2178, 2014.

[43] J. Saez-Gallego, M. Kohansal, A. Sadeghi-Mobarakeh, and J. M. Morales, "Optimal price-energy demand bids for aggregate price-responsive loads," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5005–5013, 2018.

[44] J. Soares, B. Canizes, M. A. F. Ghazvini, Z. Vale, and G. K. Venayagamoorthy, "Two-stage stochastic model using benders' decomposition for large-scale energy resource management in smart grids," *IEEE Transactions on Industry Applications*, vol. 53, no. 6, pp. 5905–5914, 2017.

[45] C. Shao, Y. Ding, J. Wang, and Y. Song, "Modeling and integration of flexible demand in heat and electricity integrated energy system," *IEEE Transactions on Sustainable Energy*, vol. 9, no. 1, pp. 361–370, 2017.

[46] M. Asensio, P. M. de Quevedo, G. Muñoz-Delgado, and J. Contreras, "Joint distribution network and renewable energy expansion planning considering demand response and energy storage—part i: Stochastic programming model," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 655–666, 2016.

[47] M. Asensio, P. M. de Quevedo, G. Muñoz-Delgado, and J. Contreras, "Joint distribution network and renewable energy expansion planning considering demand response and energy storage—part ii: Numerical results," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 667–675, 2016.

[48] S. Chen and R. S. Cheng, "Operating reserves provision from residential users through load aggregators in smart grid: a game theoretic approach," *IEEE Transactions on Smart Grid*, vol. 10, no. 2, pp. 1588–1598, 2017.

[49] P.-H. Cheng, T.-H. Huang, Y.-W. Chien, C.-L. Wu, C.-S. Tai, and L.-C. Fu, "Demand-side management in residential community realizing sharing economy with bidirectional pev while additionally considering commercial area," *International Journal of Electrical Power & Energy Systems*, vol. 116, p. 105512, 2020.

[50] G. De Zotti, S. A. Pourmousavi, J. M. Morales, H. Madsen, and N. K. Poulsen, "Consumers' flexibility estimation at the tso level for balancing services," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 1918–1930, 2018.

[51] G. Aghajani, H. Shayanfar, and H. Shayeghi, "Demand side management in a smart micro-grid in the presence of renewable generation and demand response," *Energy*, vol. 126, pp. 622–637, 2017.

[52] F. Salah, R. Henríquez, G. Wenzel, D. E. Olivares, M. Negrete-Pincetic, and C. Weinhardt, "Portfolio design of a demand response aggregator with satisficing consumers," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2475–2484, 2018.

[53] S. A. Gabriel, A. J. Conejo, J. D. Fuller, B. F. Hobbs, and C. Ruiz, *Complementarity modeling in energy markets*, vol. 180. Springer Science & Business Media, 2012.

[54] T. Li and M. Shahidehpour, "Strategic bidding of transmission-constrained gencos with incomplete information," *IEEE Transactions on power Systems*, vol. 20, no. 1, pp. 437–447, 2005.

[55] A. Ortner and G. Totschnig, "The future relevance of electricity balancing markets in europe-a 2030 case study," *Energy Strategy Reviews*, vol. 24, pp. 111–120, 2019.

[56] J. Löfberg, "Yalmip: A toolbox for modeling and optimization in matlab," in *Proceedings of the CACSD Conference*, vol. 3, Taipei, Taiwan, 2004.

[57] A. Wächter and L. T. Biegler, "On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming," *Mathematical programming*, vol. 106, no. 1, pp. 25–57, 2006.

[58] J. Romvary, G. Ferro, R. Haider, and A. Annaswamy, "A proximal atomic coordination algorithm for distributed optimization," *IEEE Transactions on Automatic Control*, Under review.

[59] S. Sra, S. Nowozin, and S. J. Wright, *Optimization for machine learning*. Mit Press, 2012.

[60] B. Everitt, *Introduction to optimization methods and their application in statistics*. Springer science & business media, 2012.

[61] A. Aljanaby, E. Abuelrub, and M. Odeh, "A survey of distributed query optimization.," *Int. Arab J. Inf. Technol.*, vol. 2, no. 1, pp. 48–57, 2005.

[62] H. W. Dommel and W. F. Tinney, "Optimal power flow solutions," *IEEE Transactions on power apparatus and systems*, no. 10, pp. 1866–1876, 1968.

[63] M. T. Thai and P. M. Pardalos, *Handbook of optimization in complex networks: theory and applications*, vol. 57. Springer science business media, 2011.

[64] L. Wang, P. K. Basu, and J. P. Leiva, "Automobile body reinforcement by finite element optimization," *Finite Elements in Analysis and Design*, vol. 40, no. 8, pp. 879–893, 2004.

[65] D. K. Molzahn, F. Dörfler, H. Sandberg, S. H. Low, S. Chakrabarti, R. Baldick, and J. Lavaei, "A survey of distributed optimization and control algorithms for electric power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 6, pp. 2941–2962, 2017.

[66] G. Cohen, "Auxiliary problem principle and decomposition of optimization problems," *Journal of optimization Theory and Applications*, vol. 32, no. 3, pp. 277–305, 1980.

[67] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.

[68] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*, vol. 23. Prentice hall Englewood Cliffs, NJ, 1989.

[69] A. Makhdoumi and A. Ozdaglar, "Convergence rate of distributed admm over networks," *IEEE Transactions on Automatic Control*, vol. 62, no. 10, pp. 5082–5095, 2017.

[70] G. Chen and M. Teboulle, "A proximal-based decomposition method for convex minimization problems," *Mathematical Programming*, vol. 64, no. 1-3, pp. 81–101, 1994.

[71] J. N. Tsitsiklis, "Problems in decentralized decision making and computation.," tech. rep., Massachusetts Inst of Tech Cambridge Lab for Information and Decision Systems, 1984.

[72] J. Tsitsiklis, D. Bertsekas, and M. Athans, "Distributed asynchronous deterministic and stochastic gradient optimization algorithms," *IEEE transactions on automatic control*, vol. 31, no. 9, pp. 803–812, 1986.

[73] F. Iutzeler, P. Bianchi, P. Ciblat, and W. Hachem, "Explicit convergence rate of a distributed alternating direction method of multipliers," *IEEE Transactions on Automatic Control*, vol. 61, no. 4, pp. 892–904, 2015.

[74] A. Mokhtari, W. Shi, Q. Ling, and A. Ribeiro, "Dqm: Decentralized quadratically approximated alternating direction method of multipliers," *IEEE Transactions on Signal Processing*, vol. 64, no. 19, pp. 5158–5173, 2016.

[75] W. Shi, Q. Ling, K. Yuan, G. Wu, and W. Yin, "On the linear convergence of the admm in decentralized consensus optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 7, pp. 1750–1761, 2014.

[76] A. Falsone, K. Margellos, S. Garatti, and M. Prandini, "Dual decomposition for multi-agent distributed optimization with coupling constraints," *Automatica*, vol. 84, pp. 149–158, 2017.

[77] K. Margellos, A. Falsone, S. Garatti, and M. Prandini, "Distributed constrained optimization and consensus in uncertain networks via proximal minimization," *IEEE Transactions on Automatic Control*, vol. 63, no. 5, pp. 1372–1387, 2017.

[78] E. Kushilevitz, "Communication complexity," in *Advances in Computers*, vol. 44, pp. 331–360, Elsevier, 1997.

[79] D. P. Bertsekas, "Multiplier methods: a survey," *Automatica*, vol. 12, no. 2, pp. 133–145, 1976.

[80] N. Parikh, S. Boyd, *et al.*, "Proximal algorithms," *Foundations and Trends in Optimization*, vol. 1, no. 3, pp. 127–239, 2014.

[81] D. P. Bertsekas, "Incremental gradient, subgradient, and proximal methods for convex optimization: A survey," *Optimization for Machine Learning*, vol. 2010, no. 1-38, p. 3, 2011.

[82] W. Deng and W. Yin, "On the global and linear convergence of the generalized alternating direction method of multipliers," *Journal of Scientific Computing*, vol. 66, no. 3, pp. 889–916, 2016.

[83] G. Qu and N. Li, "Accelerated distributed nesterov gradient descent," *IEEE Transactions on Automatic Control*, 2019.

[84] Y. Nesterov, "Introductory lectures on convex programming volume i: Basic course," *Lecture notes*, vol. 3, no. 4, p. 5, 1998.

[85] E. K.Ryu and S. Boyd, "Primer on monotone operator models," *Appl. Comput. Math*, vol. 15, pp. 3–43, 2016.

[86] The Mathworks, Inc., Natick, Massachusetts, *MATLAB version 9.5.0.713579 (R2018b)*, 2018.

[87] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[88] W. Deng, M.-J. Lai, Z. Peng, and W. Yin, “Parallel multi-block admm with o (1/k) convergence,” *Journal of Scientific Computing*, vol. 71, no. 2, pp. 712–736, 2017.