

Coordinated operation of electric power transmission and distribution systems

Dissertation presented by
Céline GERARD

for obtaining the Master's degree in
Mathematical Engineering

Supervisor
Anthony PAPAVALIOU

Readers
Emmanuel DE JAEGER, Yves SMEERS

Academic year 2016-2017

Abstract

In the modern context of seeking sustainability through the large-scale integration of renewable energy, the amount of renewable resources located in the distribution part of the electric power network has been growing and becoming a progressively important component of the electric power supply chain. These renewable resources are originating mainly from rooftop solar panels. The distribution part of electric power supply chain is composed of low-voltage transmission lines to which residential and commercial loads are connected, unlike the transmission part of the network which consists of high-voltage transmission lines, to which large generators and industrial loads are connected. For the moment, only the transmission part of the system is taken into account in the optimal scheduling of power systems. This creates an opportunity to optimize the low-voltage part of the network in order to generate savings in the consumption of energy. Indeed, the distribution network has an important impact on the scheduling of the consumption of energy. However, the operation of the network's distribution part is in some aspects more complex than the one of the transmission network, since the linear approximation of Kirchhoff's laws does not provide an adequate representation of the physical laws governing the flow of power, voltage needs to be controlled and can not be assumed to fluctuate around its nominal value, and reactive power needs to be accounted for explicitly.

The goal of this thesis is to analyze if the co-optimization of the transmission and distribution part of the network justify the required increase in computational effort. For this purpose, the first chapter of this document is devoted to the settlement of all the optimization models needed for the remaining part of this thesis. An example also proves the possibility of an improvement in the operation of the electrical network due to the co-optimization. After that, some distributed algorithms are presented in order to deal with the co-optimization problem in a distributed fashion. Among these algorithms, the ADMM algorithm, the Caramanis method and the Level method are presented. The final objective of this project is to compare various algorithms for performing co-optimization, on large-scale data, in order to determine the resulting efficiency gains and assess the scalability of the developed algorithms for networks of realistic scale.

Acknowledgements

I would like to take the opportunity to thank all the people who have contributed in any aspect of this master thesis and who have supported me during one year of work.

I must especially express my profound gratitude to my supervisor, Professor Papavasiliou, for his attention and for all the time he devoted to me. Indeed, his door was always open whenever I ran into a problem with the codes or had a question about the research. His advice and expertise were really appreciated and his guidance allowed me to conduct my work in the right direction.

My thanks also go to Professor De Jaeger and Professor Smeers for agreeing to be part of the thesis jury as readers.

Moreover, I am very grateful for the precious help and advice which Professor Papavasiliou Ph'Ds gave, especially Ignacio Aravena. Without his input, my work could not have been successfully conducted.

Furthermore, I would like to thank my friends and my boyfriend for the interesting discussions, for their interest in my work, for the mutual aid approaching deadlines and for the fun we had this last five years.

Last but not the least, I would like to thank my family, my parents and my sisters, for providing me an unfailing support and continuous encouragement throughout my years of engineering studies and this thesis. This accomplishment would not have been possible without them.

Contents

List of Symbols and Abbreviations	v
List of Figures	ix
List of Tables	xii
1 Introduction	1
2 Models and Example	4
2.1 Power Systems Representation	4
2.2 Optimal Power Flow	5
2.2.1 Alternating Current Optimal Power Flow	6
2.2.2 Direct Current Optimal Power Flow	9
2.3 Unit Commitment	10
2.3.1 Variables	10
2.3.2 Constraints	11
2.4 Models	13
2.4.1 Transmission	13
2.4.2 Distribution	14
2.4.3 Linking Constraint	15
2.4.4 Total Problem	17
2.5 Motivating Example	19
3 State of the Art	22
3.1 Optimal Power Flow Convex Relaxations	22
3.2 Unit Commitment Formulations	24
3.3 Distributed Algorithms for Large-Scale Interconnected Power Systems	25
3.4 Global Analysis of the Whole Power System	25
4 Algorithms	27
4.1 Level Method	27
4.1.1 Characteristics and Reformulation of the Original Problem	27
4.1.2 Algorithm	31
4.1.3 Stopping Criterion	32
4.2 Alternating Direction Method of Multipliers	33
4.2.1 Augmented Lagrangian	33
4.2.2 Algorithm	34
4.2.3 Convergence	35
4.2.4 Stopping Criterion	36
4.3 Caramanis Method	36

4.3.1	Algorithm	36
4.3.2	Stopping Criterion	38
5	Results	39
5.1	Description Test Cases	39
5.1.1	Transmission Data	39
5.1.2	Distribution	40
5.1.3	General Remarks	41
5.2	Computational Results	41
5.2.1	Tuning Parameters	41
5.2.2	Efficiency Algorithms and Commercial Solver Results	48
5.3	Discussion	54
5.4	Further details	55
5.4.1	Unit Commitment Added Features	55
5.4.2	Computational Observations	55
6	Conclusion	56
6.1	Main Results	56
6.2	Future Perspectives	57
	Bibliography	57
	A Second Formulation of the Global Optimization Problem	62
	B Impact of the Penalty Parameter on Primal Variables	64

List of Symbols and Abbreviations

Sets

$\mathcal{C}_i$	Set of nodes which are children of node i in the distribution network. A child of node i is a node linked to node i by a directed line going from the child node towards node i .
$\mathcal{F}_D^{n,t_n}$	Set of values that the group of variables associated with the distribution networks can take in order to satisfy the distribution set of constraints of one distribution network, indexed by (n,t_n) , connected to node n of the transmission network such that $\mathcal{T}_n \neq \emptyset$
$\mathcal{F}_D$	Set of values that the group of variables associated with a distribution network can take in order to satisfy the distribution set of constraints
$\mathcal{F}_T$	Set of values that can be taken by the group of variables associated with the transmission network in order to satisfy the transmission set of constraints
$\mathcal{G}_n$	Set of generators located at node n in the transmission network, $\bigcup_{n \in N} \mathcal{G}_n = \mathcal{G}$
$\mathcal{G}$	Set of generating units located in the transmission network
$\mathcal{I}_0$	Set of buses located in the distribution network. Since the distribution network is radial, lines in the distribution network own the same index as the corresponding child node adjacent to this line. Therefore, $\mathcal{I} \cup \{0\} = \mathcal{I}_0 = \{0, 1, \dots, I\}$.
$\mathcal{I}$	Set of lines in the distribution network, $\mathcal{I} = \{1, \dots, I\}$
$\mathcal{K}$	Set of lines which links two buses of the transmission network, $\mathcal{K} = \{1, \dots, K\}$
$\mathcal{N}$	Set of buses belonging to the transmission network
$\mathcal{T}_n$	Set of distribution networks connected to the transmission bus n
$\mathcal{T}$	Set of time periods (hours, minutes, days,...), $\mathcal{T} = \{1, \dots, T\}$. Usually T is equal to 24 and time periods represent hours.

Indices

A_i	Ancestor's index of node i in the distribution network. The ancestor of node i is a node linked to node i by a directed line going from node i to its ancestor.
τ	Alternative time period index, $\tau \in \mathcal{T}$
c_i	Children's index for node i belonging to the distribution network, $c_i \in \mathcal{C}_i$
g	Generators index for the transmission network, $g \in \mathcal{G}$
i	Bus or line index for the distribution network, $i \in \mathcal{I}_0$ and $\mathcal{I} \subset \mathcal{I}_0$
$k = (\cdot, n)$	Directed edges going from any node of the transmission network to bus n , $k = (\cdot, n) \in \mathcal{K}$
$k = (m, n)$	Directed edges going from node m to node n , $k = (m, n) \in \mathcal{K}$
$k = (n, \cdot)$	Directed edges going from node n to any bus of the transmission network, $k = (n, \cdot) \in \mathcal{K}$

k	Line index for lines belonging to the transmission network, $k \in \mathcal{K}$
m	Alternative bus index for the transmission network, $m \in \mathcal{N}$
n	Bus index for buses belonging to the transmission network, $n \in \mathcal{N}$
t_n	Index representing a distribution network connected to the transmission node n , $t_n \in \mathcal{T}_n$
t	Time periods index, $t \in \mathcal{T}$

Variables

D_{∞}^{t,n,t_n}	Real power withdrawn from the transmission node n in order to inject it in the distribution network indexed by t_n at time period t , expressed in [p.u.]
$D_{\infty}^{t,n}$	Real power withdrawn from the transmission node n in order to inject it in the set of distribution networks linked to this node at time period t , expressed in [p.u.]
D_{∞}	Real power withdrawn from the transmission network at the connection node between the two networks in order to inject this power into the distribution network, expressed in [p.u.]
P_{∞}^{t,n,t_n}	Real power injected into the distribution network t_n , at the root, coming from the transmission bus n at time period t , expressed in [p.u.]
P_{∞}	Real power injected into the distribution network, at the root, coming from the transmission network, expressed in [p.u.]
$P_{i,t}$	Real power flowing from node i to node A_i at time period t , expressed in [p.u.]
$Q_{i,t}$	Reactive power flowing from node i to node A_i at time period t , expressed in [p.u.]
$\eta_{t,n}$	Dual variable corresponding to the linking constraint indexed by (t,n) related to the first formulation of the global optimization problem
$\theta_{n,t}$	Electrical phase angle at node n at time period t , expressed in [p.u.]
$\tilde{I}_{i,j}$	Complex current flowing from node i to node j , $\tilde{I}_{i,j} = \tilde{I}_{i,j} ^2$, expressed in [A]
$\tilde{P}_{i,j}$	Real power flowing from node i to node j , expressed in [MW]
$\tilde{Q}_{i,j}$	Reactive power flowing from node i to node j , expressed in [MVAR]
$\tilde{S}_{i,j}$	Complex branch power flowing from node i to node j , $\tilde{S}_{i,j} = \tilde{P}_{i,j} + \mathbf{i}\tilde{Q}_{i,j}$, expressed in [MVA]
$\tilde{V}_i$	Complex voltage at node i , $\tilde{v}_i = \tilde{V}_i ^2$, expressed in [V]
$\tilde{\theta}_i$	Electrical phase angle at node i , expressed in [rad]
$\tilde{l}_{i,j}$	Current magnitude squared flowing from node i to node j , expressed in [A ²]
$\tilde{p}_i$	Real power injection at node i , expressed in [MW]
$\tilde{q}_i$	Reactive power injection at node i , expressed in [MVAR]
$\tilde{s}_i$	Net complex power injection at node i , $\tilde{s}_i = \tilde{p}_i + \mathbf{i}\tilde{q}_i$, expressed in [MVA]
$\tilde{v}_i$	Voltage magnitude squared at node i , expressed in [V ²]
$f_{k,t}$	Flow through line k at time period t , expressed in [p.u.]
$l_{i,t}$	Current magnitude squared flowing from node i to node A_i at time period t , expressed in [p.u.]
$p_{g,t}$	Total power generated by generator g at time period t , expressed in [p.u.]
$p_{i,t}$	Real power injection at node i at time period t expressed in [p.u.]
$q_{i,t}$	Reactive power injection at node i at time period t , expressed in [p.u.]
$sd_{g,t}$	Binary variable equal to 1 if unit g is shutdown at time t
$su_{g,t}$	Binary variable equal to 1 if unit g is started up at time t
$u_{g,t}$	Binary variable equal to 1 if unit g is online at time t
$v_{i,t}$	Voltage magnitude squared at node i at time period t , expressed in [p.u.]

Parameters

B_i	Shunt susceptance located at node i , expressed in [p.u.]
B_k	Susceptance of transmission line k , expressed in [p.u.]
DT_g	Minimum downtime for generator g , expressed in hours [h]
FC_g	Fixed cost of generator g , expressed in [\$]
G_i	Shunt conductance located at node i , expressed in [p.u.]
MC_i	Marginal cost of real power injection at node i , expressed in [\$/p.u.]
MC_g	Linear term of the variable cost of generator g , expressed in [\$/p.u.]
R_g^+	Upward ramp rate limiting the real power delivered by generator g , expressed in [p.u./h]
R_g^-	Downward ramp rate limiting the real power delivered by generator g , expressed in [p.u./h]
R_i	Resistance of the line going from node i to node A_i , expressed in [p.u.]
SD_g	Shut down cost of generator g , expressed in [\$]
SU_g	Start-up cost of generator g , expressed in [\$]
UT_g	Minimum uptime for generator g , expressed in hours [h]
X_i	Reactance of the line going from node i to node A_i , expressed in [p.u.]
α	Level Method's parameter
ϵ	Parameter used to penalize high values for current magnitude in the objective function
ρ	ADMM Mehtod's parameter related to the regularization term
$\tilde{B}_i$	Shunt susceptance located at node i , expressed in [S]
$\tilde{B}_{i,j}$	Susceptance of the line going from node i to node j , expressed in [S]
$\tilde{G}_i$	Shunt conductance located at node i , expressed in [S]
$\tilde{P}_{i,j}^+$	Real power flow upper limit of the electrical link going from node i to j , expressed in [MW]
$\tilde{P}_{i,j}^-$	Real power flow lower limit of the electrical link going from node i to j , expressed in [MW]
$\tilde{R}_{i,j}$	Resistance of the line going from node i to node j , expressed in [Ω]
$\tilde{X}_{i,j}$	Reactance of the line going from node i to node j , expressed in [Ω]
$\tilde{Y}_i$	Complex shunt admittance on node i , $\tilde{Y}_i = \tilde{G}_i + \mathbf{i}\tilde{B}_i$, expressed in [S]
$\tilde{Z}_{i,j}$	Complex impedance on line going from node i to node j , $\tilde{Z}_{i,j} = \tilde{R}_{i,j} + \mathbf{i}\tilde{X}_{i,j}$, expressed in [Ω]
$\tilde{d}p_i$	Real power consumption at node i , expressed in [MW]
$\tilde{d}q_i$	Reactive power consumption at node i , expressed in [MVAR]
$\tilde{d}s_i$	Complex power consumption at node i , $\tilde{d}s_i = \tilde{d}p_i + \mathbf{i}\tilde{d}q_i$, expressed in [MVA]
$\tilde{p}_i^+$	Maximum real power injection at node i , expressed in [MW]
$\tilde{p}_i^-$	Minimum real power injection at node i , expressed in [MW]
$\tilde{q}_i^+$	Maximum reactive power injection at node i , expressed in [MVAR]
$\tilde{q}_i^-$	Minimum reactive power injection at node i , expressed in [MVAR]
$\tilde{v}_i^+$	Maximum squared magnitude voltage at node i , expressed in [V]
$\tilde{v}_i^-$	Minimum squared magnitude voltage at node i , expressed in [V]
$d_{n,t}$	Electrical power demand at bus n in the transmission network at time t , expressed in [p.u.]

$dp_{i,t}$	Real power demand at node i at time period t , expressed in [p.u.]
$dq_{i,t}$	Reactive power demand at node i at time period t , expressed in [p.u.]
f_k^+	Maximum electrical flow through line k , expressed in [p.u.]
hub	Reference node of the transmission network, $hub \in \mathcal{N}$
p_g^+	Maximum electrical power produced by generator g , expressed in [p.u.]
p_g^-	Minimum possible power to produce for generator g while online, expressed in [p.u.]
p_i^+	Maximum real power injection at node i , expressed in [p.u.]
p_i^-	Minimum real power injection at node i , expressed in [p.u.]
q_i^+	Maximum reactive power injection at node i , expressed in [p.u.]
q_i^-	Minimum reactive power injection at node i , expressed in [p.u.]
tol	Tolerance applied for algorithms' stopping criterion
v_i^+	Maximum squared magnitude voltage at node i , expressed in [p.u.]
v_i^-	Minimum squared magnitude voltage at node i , expressed in [p.u.]

Abbreviations

ACOPF	Alternating Current Optimal Power Flow
ADMM	Alternating Direction Method of Multipliers
DCOPF	Direct Current Optimal Power Flow
DN	Distribution Network
OPF	Optimal Power Flow
SDP	Semi-definite program
SOCP	Second-order cone program
TN	Transmission Network
AMPL®	Mathematical programming language uses to describe and solve optimization problems
CPLEX®	Optimization solver that can deal with integer, linear and conic programs
MatLab®	Programming language developed by MathWorks used for numerical computations
MatPower®	MatPower® is a package of MatLab® that aims to solve power flows and optimal power flows problems. This package is intended as a simulation tool for researchers and educators.

List of Figures

1.1	Illustration of the electricity supply chain coming from [1]	1
2.1	Illustration of the electricity supply chain coming from [1]	4
2.2	Example of a graph representing a small electrical power system	5
2.3	Illustration of notations used to derive power flow equations	6
2.4	Illustration of minimum up and down times of a generator on its On/Off status over 12 hours	11
2.5	Illustration of the effect of ramp rates constraints on generator's production presented in example 2.2	12
2.6	Electrical power system representation with highlighted transmission part	13
2.7	Electrical power system representation with highlighted distribution part	15
2.8	Illustration of the convention used in order to link the transmission and the distribution networks	16
2.9	Illustration of the global power system used for the motivating example	20
5.1	Daily load variation characteristic of the test system	40
5.2	Evolution of the maximum slack of the conic constraint for some values of the penalty parameter ϵ . The tests were made for several number of distribution networks per node. The left figure presents results obtained for the first formulation of the linking constraint, while the right figure considers the second formulation of this constraint.	42
5.3	Evolution of the objective distortion in percent for some values of parameter ϵ . The tests were made for several numbers of distribution networks per node. The left figure present results obtained for the first formulation of the linking constraint, while the right figure considers the second formulation of this constraint.	43
5.4	Evolution of the primal variable $p_{g,t}$ distortion in percent for some values of parameter ϵ . the tests were made for several numbers of distribution networks per node. The left figure present results obtained for the first formulation of the linking constraint, while the right figure considers the second formulation of this constraint.	43
5.5	Illustration of the number of iterations needed by the ADMM method to converge for a certain number of distribution networks per node and several values of the regularization parameter	45
5.6	Illustration of the time needed by the ADMM method to converge for a certain number of distribution networks per node and several values of the regularization parameter	45
5.7	Illustration of the gap obtained after convergence of the ADMM method for a certain number of distribution networks per node and several values of the regularization parameter	46

5.8	Illustration of the number of iterations needed by the Level method to converge for a certain number of distribution networks per node and several values of the algorithm's parameter	47
5.9	Illustration of the time that is required by the Level method to converge for a certain number of distribution networks per node and several values of the algorithm's parameter	47
5.10	Illustration of the gap obtained after convergence by the Level method for a certain number of distribution networks per node and several values of the algorithm's parameter	47
5.11	Graphic representation of the time required by the commercial solver to compute the solution of the global optimization program for several numbers of distribution networks per transmission node using the 6 bus dataset (left) and the 24 bus dataset (right)	48
5.12	Graphic representation of the optimal objective value of the global optimization program for several numbers of distribution networks per transmission node using the 6 bus dataset (left) and the 24 bus dataset (right)	49
5.13	Graphical representation of the relative gap in percent achieved by the commercial solver while solving the global optimization problem for different numbers of distribution networks per transmission node using the 6 bus dataset (left) and the 24 bus dataset (right)	49
5.14	Graphical representation of the time needed by the Caramanis method to converge for several numbers of distribution networks per transmission node using the 6 bus dataset (left) and the 24 bus dataset (right)	50
5.15	Graphical representation of the evolution of the number of iterations required by the Caramanis method to converge using the 6 bus dataset (left) and the 24 bus dataset (right)	51
5.16	Graphical representation of the evolution of the obtained gap after convergence of the algorithm for different numbers of distribution networks per node using the 6 bus dataset (left) and the 24 bus dataset (right)	51
5.17	Graphical representation of the evolution of the time required by the ADMM algorithm to converge for different number of distribution networks using the 6 bus dataset (left) and the 24 bus dataset (right)	52
5.18	Graphical representation of the evolution of the number of iterations required by the ADMM algorithm to converge for different number of distribution networks using the 6 bus dataset (left) and the 24 bus dataset (right)	52
5.19	Graphical representation of the evolution of the resulting gap obtained after convergence of the ADMM method for different number of distribution networks using the 6 bus dataset (left) and the 24 bus dataset (right)	53
5.20	Graphical representation of the evolution of the time (left) and the number of iterations (right) required by the Level algorithm to converge for different number of distribution networks using the 6Bus dataset and the first formulation of the linking constraint	54
B.1	Evolution of the primal variable $l_{i,t}$ distortion in percent for some values of parameter ϵ for several number of distribution networks per node. The left figure presents results obtained for the first formulation of the linking constraint, the right figure for the second formulation.	64

B.2	Evolution of the primal variable $D_{\infty}^{t,n}$ distortion in percent for some values of parameter ϵ for several number of distribution networks per node. The left figure presents results obtained for the first formulation of the linking constraint, the right figure for the second formulation.	65
-----	---	----

List of Tables

2.1	Notations and variables used to derive the expression of the power flow equations	6
2.2	Detailed Branch Data	19
2.3	Detailed Load Data	20
2.4	Detailed Generators Data	20
2.5	Solution of the motivating example	21

Chapter 1

Introduction

Electricity's impact in our lives is wider than we used to think. Indeed, electricity has turned into a true political and economical issue during the last decade, compared to the beginning of the 20th century, where electricity was considered as an expensive product reserved only to a small part of the population. In fact, nowadays, electricity can be found anywhere in our day-to-day lives from the computer to the fridge by way of streetlights. To support such a huge accumulation of technologies involving electrical energy, power systems have grown in order to become some of the most complex and expensive entities in the world [2].

Traditionally, the path followed by electricity through power systems was well defined. The electrical power was coming from generators where the energy was produced. Then, the electricity was transmitted through high-voltage transmission lines. After that, electrical energy was carried by low-voltage transmission lines in order to eventually be consumed by residential and commercial individuals. The electrical supply chain, as illustrated in figure 1.1, includes the main components of an electrical power system. The generators and industrial consumption along with high-voltage lines are considered to belong to the so-called **Transmission** network (TN) while consumers and low-voltage lines are considered to be part of the **Distribution** network (DN). Producers and consumers in such power networks are represented by buses. A bus is a location where electrical power is withdrawn from the system or injected into the grid.

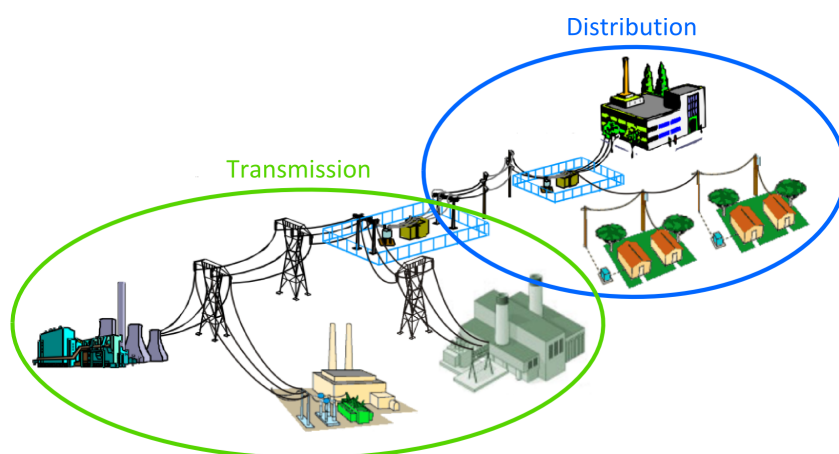


Figure 1.1: Illustration of the electricity supply chain coming from [1]

Due to the traditional unidirectional path followed by electrical energy through power grids, the transmission and distribution networks were studied as two distinct entities over the past few years. However, the large-scale integration of renewable energy sources increased the difficulty to plan and operate power systems [3]. The European Union among other organizations and protagonists encouraged this substantial renewable resources integration. Their purpose is to achieve energy policy objectives in order to decrease greenhouse gas emission and increase the efficiency of power grids' operation. In the interest of reaching these goals as, for example, reducing its dependence on fossil fuels and making its energy more sustainable, the European Union has set a target of 20 % energy coming from renewable sources in 2020 [4]. The recent rise of renewable power sources, especially rooftop solar panels, has enabled traditional consumers, located in the distribution network, to produce power and, therefore, to become a new type of producers. Consequently, part of the electrical energy generated by these new producers can now flow from the distribution to the transmission network in order to fulfill, for example, industrial demand. Since classical electricity's path has recently turned into bidirectional flows between the transmission and the distribution parts of power networks, the Europe's traditional grids have to be modernized into SmartGrids to meet objectives in the operation of power grids [5]. A SmartGrid is an electrical power system that can integrate judiciously actions of grid's users (producers and/or consumers) to deliver effectively safe, sustainable and economic electricity supplies [6].

Furthermore, apart from the large-scale integration of renewable energy sources, the development of electrical vehicles and batteries such as Tesla Powerwalls induces a significant increase in electrical demand at certain times of the day. However, these two technologies among others may bring answers to the variability and uncertainty related to some renewable sources such as wind and solar power. In fact, massive integration of renewable power has led to a system which requires a growing need for flexibility in electricity consumption and additional efforts to balance the system due to the unpredictability of several power sources. Therefore, the way the electricity is produced and distributed has to be redesigned and thought over. In this new design of electrical power grids operations, distribution system operators should have to take a deeper role than their traditional one, due, for example, to the impacts of the vast integration of renewable energy to the distribution part of the network. This may be achieved, for instance, by considering the transmission and distribution parts of the system as a whole. Ignorance of such interactions between these two networks can lead to technical issues and impact the analysis of power flows through the entire power system for dispatching, contingencies and other typical studies.

The analysis of transmission and distribution networks as a whole global power system may lead to significant benefits in the operation of power systems as said by Sun et al. in [7]. Therefore, this work will focus on the integrated analysis of both grids for computing the optimal operation of the system by taking into account physical laws governing electricity power flows in these two types of networks along with several characteristics of electricity generators such as their maximal power output. This thesis will try to answer the following questions about this optimization problem on the entire network:

1. Does the co-optimization of both networks worth the implementation?
2. Can this co-optimization even be performed?
3. Are distributed algorithms able to solve the co-optimization problem efficiently?

The second and third questions are linked to the goal of applying the co-optimization idea to an entire real network. In fact, in real networks, transmission buses number in the thousands while distribution buses number in the hundred of thousands or millions according to Caramanis et al. [8]. Therefore, the computational effort required to solve the global optimization program is expected to be huge since power networks contains a really large number of buses.

This thesis is organized in the following way. First, models, used in order to represent the physical laws governing power flows and several characteristics of generation units, are presented. These problems lead us to derive the transmission and distribution models in order to state a global optimization problem. In the same chapter, an example allows to answer the first question on the worth of the global optimization problem's implementation. Then, main literature contributions on linked subjects are explained. Subsequently, distributed algorithms are presented in order to decompose and solve efficiently the global optimization problem. Finally, a last chapter is devoted to results obtained by computing the optimal solution of the global optimization problem using distributed algorithms in terms of solving time, optimality of the obtained solution and number of iterations required by the algorithm to converge.

Chapter 2

Models and Example

Electricity is an essential element of our daily life, as emphasized in the introduction (chapter 1). In the interest of producing and providing this commodity, electrical power systems were built and optimization problems were set up to help humans schedule the production of electricity. These mathematical programs allow to formalize in a theoretical way the physical laws behind electrical power system operation. In this part of the document, some of these optimization problems such as the ***Optimal Power Flow*** and the ***Unit Commitment*** problems will be presented. After that, the mathematical program studied in this document will be stated. Finally, an example showing the necessity of this work on entire power systems will be introduced.

2.1 Power Systems Representation

An electric power system is a network consisting of components that produce electricity along with elements that enable the transfer of this electrical energy to consumers. As stated previously in the introduction (chapter 1), electric power systems can be divided into two parts: the ***Transmission*** (TN) and the ***Distribution*** (DN) networks. The ***Transmission*** network consists of three main components: generators that produce electricity, industrial consumption and high-voltage transmission lines carrying this electricity. The remaining parts of an electric power system belong, therefore, to the ***Distribution*** network: low-voltage local transmission lines, also called distribution lines, to which residential and commercial loads are connected. The entire electricity supply chain and the distinction between both parts of the network can be seen in figure 2.1.

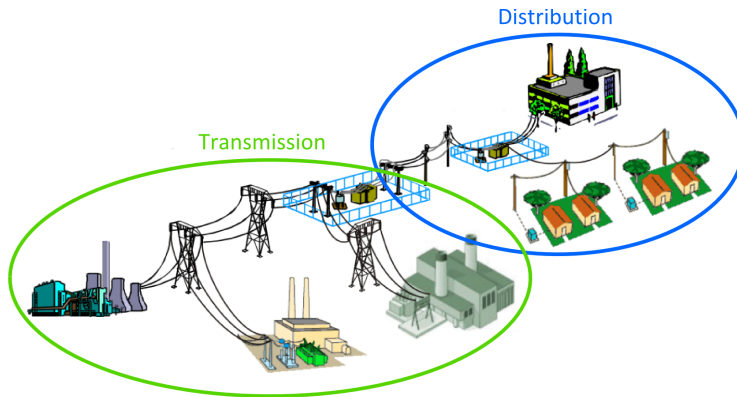


Figure 2.1: Illustration of the electricity supply chain coming from [1]

Throughout this document, electrical power systems are represented by a directed graph with a set of nodes and directed lines as in [2]. An example of such a graph representation is given in figure 2.2. Nodes or buses of the graph stand for places of the network where electrical energy is withdrawn and/or produced (generators and/or consumers). Edges or lines reproduce directed electrical links between these locations.

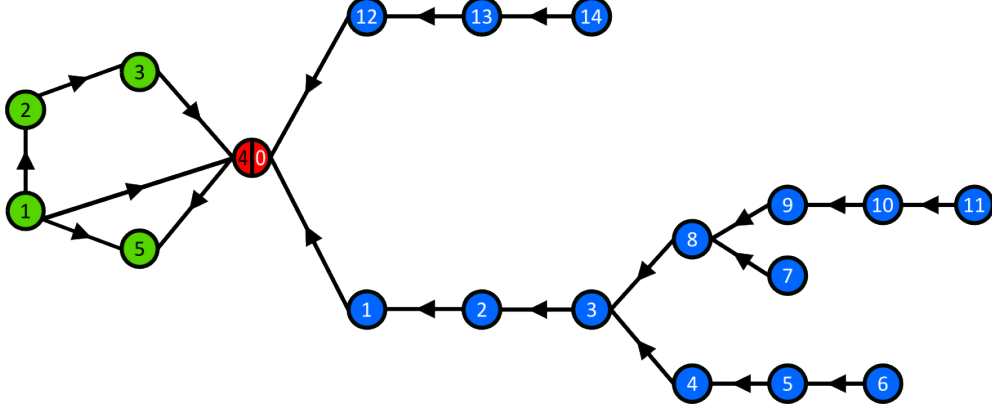


Figure 2.2: Example of a graph representing a small electrical power system

2.2 Optimal Power Flow

The simplest optimization problem allowing to dispatch efficiently the production of electricity is called ***Economic Dispatch*** problem. The production of electricity contains the use of fossil, nuclear and renewable energy sources such as wind, solar, and biomass. The mathematical program behind the ***Economic Dispatch*** problem minimizes the production cost of all generators while satisfying the total demand and the maximum output power limit of generation units. More details on this mathematical program can be found in [9]. This problem does not account for numerous physical and technical constraints behind the operation of an electrical power system such as the distance the electricity has to travel before reaching consumers and the technical constraints related to the schedule of generator's productions. Since the ***Economic Dispatch*** problem covers only a small part of the features linked to electrical power system's operations, this problem has been enriched in order to account for more characteristics of power systems.

Unfortunately, electrical lines can carry only a limited amount of electricity because of their thermal stability and voltage limits. Moreover, the flow of electricity during its transportation through the network is governed by Kirchhoff laws. By means of these laws, the ***Power Flow*** equations, a nonlinear mapping between power injections at each bus and power flows on lines, can be derived [10]. Therefore, a new optimization problem was created in order to dispatch efficiently productions and demands of electricity in a power system while taking into account technical and physical limits due to the transmission of electricity. The optimization problem providing the optimal operation of the power grid while accounting for these transmission constraints is called the ***Optimal Power Flow*** (OPF) model. There exist two ways to model these transmission constraints used to derive mathematical equations representing the behaviour of the energy flowing through the network: the ***Bus Injection*** model and the ***Branch Flow*** model [2]. These two models are equivalent and their use depends on the variables and input data chosen by the user. The first model involves the definition of the injection decision for

each node of the network which represents the amount of power shipped from this node to the reference node taken for the network. This is not the amount of power that flows over the line connecting these two nodes but the net amount of power that a node injects into the system, and which is absorbed by the reference node, fixed for the entire network. The second model focuses on the use of branch variables such as the power flowing over a line and uses susceptance of lines as input data. In this document, the focus will be on the **Branch Flow** model in order to represent the behaviour of power systems.

In this part of the document, the derivation of the **Optimal Power Flow** model is presented. For a more detailed explanation, see [2] and [11]. After that, the problem derived will be relaxed and even approximated by a linear model in order to obtain two problems that are easier to solve.

2.2.1 Alternating Current Optimal Power Flow

As mentioned in section 2.2, the goal of the OPF problem, also called Alternating Current Optimal Power Flow (ACOPF) program, is to optimize a given objective function, such as electricity generation cost or power losses along lines while accounting for Kirchhoff's laws and transmission line limits. In order to derive the **Branch Flow** model, a balance among the three phases of the system is assumed and the system is assumed to operate in steady-state. The variables and notations used to derive the **Branch Flow** equations are presented in table 2.1 and are illustrated in figure 2.3. The (*) sign represents the mathematical operation that computes the complex conjugate of a complex number.

Notations	Meaning
$\tilde{s}_i = \tilde{p}_i + \mathbf{i}\tilde{q}_i$	Complex power injected in bus i of the power network
$\tilde{S}_{i,j} = \tilde{P}_{i,j} + \mathbf{i}\tilde{Q}_{i,j}$	Complex power flowing from bus i to bus j
$\tilde{d}s_i = \tilde{d}p_i + \mathbf{i}\tilde{d}q_i$	Complex power withdrawn from bus i
$\tilde{V}_i$	Complex voltage of bus i of the power network
$\tilde{I}_{i,j}$	Complex current flowing from bus i to bus j through the power network
$\tilde{Y}_i$	Admittance present at node i
$\tilde{Z}_{i,j}$	Complex impedance of the transmission line going from bus i to bus j

Table 2.1: Notations and variables used to derive the expression of the power flow equations

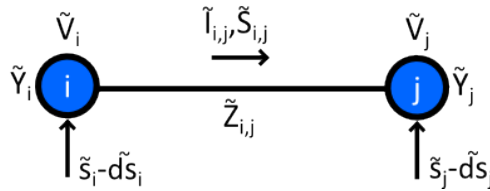


Figure 2.3: Illustration of notations used to derive power flow equations

Without loss of generality, the power system used to derive the **Branch Flow** model is represented by a directed graph with a set of nodes N and a set of edges E . A branch of the network is considered. Ohm's law can be applied to this branch and this gives the following equation:

$$\tilde{V}_i - \tilde{V}_j = \tilde{Z}_{i,j} \tilde{I}_{i,j} \quad \forall (i, j) \in E \quad (2.1)$$

Moreover, the definition of the power flowing through a branch has also to be satisfied. This gives the following expression:

$$\tilde{S}_{i,j} = \tilde{V}_i \tilde{I}_{i,j}^* \quad \forall (i,j) \in E \quad (2.2)$$

Power balance at each bus of the power system has to be verified. Indeed, the amount of power that enters the node has to be equal to the power that flows out. This principle is represented by the following equation:

$$\sum_{e:j \rightarrow e} \tilde{S}_{j,e} - \sum_{i:i \rightarrow j} (\tilde{S}_{i,j} - \tilde{Z}_{i,j} |\tilde{I}_{i,j}|^2) + \tilde{Y}_j^* |\tilde{V}_j|^2 = \tilde{s}_j - \tilde{d}s_j \quad \forall j \in N \quad (2.3)$$

Equations (2.1) to (2.3) form the **Power Flow** equations. The set of variables is therefore given by $(\tilde{S}_{i,j}, \tilde{I}_{i,j}, \tilde{V}_i, \tilde{s}_i, \tilde{d}s_i)$. In order to obtain the **Optimal Power Flow** problem, an objective function has to be defined. Any objective function that does not depend on voltage and current phase angles can be used, such as minimizing real power losses along lines $(\tilde{R}_{i,j} |\tilde{I}_{i,j}|^2)$ or the cost of real power generation or even energy savings. However, other constraints can be added to the **Branch Flow** model such as bounds on power consumption and generation at each node. Moreover, voltage magnitude can also be maintained in tight ranges and limits can also be imposed on branch currents.

The OPF model derived in this part of the document is highly nonlinear due to the **Power Flow** equations. Therefore, this problem is generally hard to solve since it is usually nonconvex and therefore NP-hard. In the aim to solve this problem using off-the-shelf solvers efficiently, this problem will be relaxed using two relaxation steps. The first step will eliminate voltage and current phase angles. The second step will relax the problem into a conic program. The goal of these relaxation steps is to replace the set of variables containing complex quantities $(\tilde{S}_{i,j}, \tilde{I}_{i,j}, \tilde{V}_i, \tilde{s}_i, \tilde{d}s_i)$ by the following set of variables $(\tilde{P}_{i,j}, \tilde{Q}_{i,j}, \tilde{l}_{i,j}, \tilde{v}_i, \tilde{p}_i, \tilde{q}_i, \tilde{d}p_i, \tilde{d}q_i)$. These two steps are detailed hereafter, for more explanations, see [11].

Angle Relaxation

The first relaxation step consists in eliminating current and voltage's phase angles from the expression of the nonlinear power flow equations. However, these angles are needed in order to have a physically meaningful optimal solution. Therefore, there are certain conditions which verify that the found relaxed solution corresponds to an actual optimal solution of the OPF problem. They are called **Angle Recovery** conditions. Fortunately, radial networks satisfy these conditions due to their particular layout. Angle recovery algorithms have been proposed in [11] and [12] in order to compute phase angles related with the optimal solution found for the relaxed problem.

By means of equations (2.1) to (2.3), a relaxed branch flow model can be obtained by eliminating current and voltage phase angles. Substituting equation (2.2) into (2.1) yields:

$$\tilde{V}_j = \tilde{V}_i - \tilde{Z}_{i,j} \frac{\tilde{S}_{i,j}^*}{\tilde{V}_i^*} \quad \forall (i,j) \in E$$

For the sake of simplicity, let $\tilde{l}_{i,j}$ stand for the squared magnitude of the current in the branch linking node i to j ($|\tilde{I}_{i,j}|^2$) and $\tilde{v}_i$ for the squared magnitude of the voltage at node i ($|\tilde{V}_i|^2$). Taking the magnitude squared of the previous equation gives:

$$\tilde{v}_j = \tilde{v}_i + |\tilde{Z}_{i,j}|^2 \tilde{l}_{i,j} - (\tilde{Z}_{i,j} \tilde{S}_{i,j}^* + \tilde{Z}_{i,j}^* \tilde{S}_{i,j}) \quad \forall (i,j) \in E$$

Expressing the previous equation in terms of real variables gives:

$$\tilde{v}_j = \tilde{v}_i - 2(\tilde{R}_{i,j}\tilde{P}_{i,j} + \tilde{X}_{i,j}\tilde{Q}_{i,j}) + (\tilde{R}_{i,j}^2 + \tilde{X}_{i,j}^2)\tilde{l}_{i,j} \quad \forall (i, j) \in E \quad (2.4)$$

After that, expressing equation (2.3) in terms of real variables gives:

$$\tilde{p}_j - \tilde{d}p_j = \sum_{e:j \rightarrow e} \tilde{P}_{j,e} - \sum_{i:i \rightarrow j} (\tilde{P}_{i,j} - \tilde{R}_{i,j}\tilde{l}_{i,j}) + \tilde{G}_j\tilde{v}_j \quad \forall (i, j) \in E \quad (2.5)$$

$$\tilde{q}_j - \tilde{d}q_j = \sum_{e:j \rightarrow e} \tilde{Q}_{j,e} - \sum_{i:i \rightarrow j} (\tilde{Q}_{i,j} - \tilde{X}_{i,j}\tilde{l}_{i,j}) - \tilde{B}_j\tilde{v}_j \quad \forall (i, j) \in E \quad (2.6)$$

Finally, taking the squared magnitude of equation (2.2) gives:

$$\tilde{l}_{i,j} = \frac{\tilde{P}_{i,j}^2 + \tilde{Q}_{i,j}^2}{\tilde{v}_i} \quad \forall (i, j) \in E \quad (2.7)$$

The relaxed OPF model obtained by eliminating voltage and current phase angles from the previous branch flow set of equations is a new optimization problem where the new power flow equations are given by equations (2.4) to (2.7). However, due to quadratic equalities in equation (2.7), this problem remains nonconvex. Another relaxation to the problem will therefore be applied to make the optimization problem convex.

As a reminder, hereafter are some notions about convexity found in [13] and [14].

Definition 2.1 (Convex Optimization Problem). *A convex optimization problem is an optimization problem minimizing a convex objective function on a convex set of feasible solutions.*

Definition 2.2 (Convex Function). *A function $f(x)$ is called convex if its domain is convex and for all $x, y \in \text{dom}_f$ and $\lambda \in [0, 1]$ the following inequality holds:*

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

We called f concave if $-f$ is convex.

Definition 2.3 (Convex Sets). *A set A is called convex if for any $x, y \in A$ and $\lambda \in [0, 1]$, we have:*

$$\lambda x + (1 - \lambda)y \in A$$

The point $\lambda x + (1 - \lambda)y$ with $\lambda \in [0, 1]$ is called the convex combination of these two points.

SOCP Relaxation

As stated previously, the OPF problem is nonconvex and this non convexity stems from Kirchhoff's laws. Therefore, a new relaxation is applied to the problem defined by equations (2.4) to (2.7). Since the nonconvexity is due to equation (2.7) because the constraint can be represented by a circle with empty surface and a radius of $\sqrt{\tilde{v}_i\tilde{l}_{i,j}}$, this equation is relaxed to an inequality:

$$\tilde{l}_{i,j} \geq \frac{\tilde{P}_{i,j}^2 + \tilde{Q}_{i,j}^2}{\tilde{v}_i} \quad \forall (i, j) \in E \quad (2.8)$$

This relaxation transforms the quadratic equality of equation (2.7) to a convex second-order cone inequality (2.8). The new OPF problem is a **Second-Order Cone Program** (SOCP) if the objective function is linear. Second-order cone programming (SOCP) problems are convex optimization problems in which a linear function is minimized over the intersection of a polyhedron created with linear inequalities and second-order cones [15]. The definition of a second-order cone is given hereafter.

Definition 2.4 (Second-order Cone [15]). *The second-order cone, also called Lorentz cone, is defined by the convex set: $\mathbb{L}^n = \{(x_0, \dots, x_n) \in \mathbb{R}^{n+1} | \sqrt{x_1^2 + \dots + x_n^2} \leq x_0\}$.*

Using the trick $4xy = (x + y)^2 - (x - y)^2$, the previous equation 2.8 can be expressed as:

$$4\tilde{P}_{i,j}^2 + 4\tilde{Q}_{i,j}^2 + (\tilde{l}_{i,j} - \tilde{v}_i)^2 \leq (\tilde{l}_{i,j} + \tilde{v}_i)^2 \quad \forall (i, j) \in E \quad (2.9)$$

This form of the quadratic constraint will be used in the remaining of this document. The new problem obtained after the two relaxations steps will be called the **Conic Relaxation** along the document. Considering a convex objective function $h(\tilde{l}_{i,j}, \tilde{p}_i, \tilde{q}_i, \tilde{P}_{i,j}, \tilde{Q}_{i,j}, \tilde{v}_i, \tilde{d}p_i, \tilde{d}q_i)$, the Conic Relaxation program is given by:

$$\begin{aligned} & \min_{\tilde{l}_{i,j}, \tilde{p}_i, \tilde{q}_i, \tilde{P}_{i,j}, \tilde{Q}_{i,j}, \tilde{v}_i, \tilde{d}p_i, \tilde{d}q_i} h(\tilde{l}_{i,j}, \tilde{p}_i, \tilde{q}_i, \tilde{P}_{i,j}, \tilde{Q}_{i,j}, \tilde{v}_i, \tilde{d}p_i, \tilde{d}q_i) \\ & \text{subject to} \quad \tilde{v}_j = \tilde{v}_i - 2(\tilde{R}_{i,j}\tilde{P}_{i,j} + \tilde{X}_{i,j}\tilde{Q}_{i,j}) + (\tilde{R}_{i,j}^2 + \tilde{X}_{i,j}^2)\tilde{l}_{i,j} \\ & \quad 4\tilde{P}_{i,j}^2 + 4\tilde{Q}_{i,j}^2 + (\tilde{l}_{i,j} - \tilde{v}_i)^2 \leq (\tilde{l}_{i,j} + \tilde{v}_i)^2 \\ & \quad \tilde{p}_j - \tilde{d}p_j = \sum_{e:j \rightarrow e} \tilde{P}_{j,e} - \sum_{i:i \rightarrow j} (\tilde{P}_{i,j} - \tilde{R}_{i,j}\tilde{l}_{i,j}) + \tilde{G}_j\tilde{v}_j \quad \forall j \in N \\ & \quad \tilde{q}_j - \tilde{d}q_j = \sum_{e:j \rightarrow e} \tilde{Q}_{j,e} - \sum_{i:i \rightarrow j} (\tilde{Q}_{i,j} - \tilde{X}_{i,j}\tilde{l}_{i,j}) - \tilde{B}_j\tilde{v}_j \quad \forall j \in N \\ & \quad \tilde{p}_i^- \leq \tilde{p}_i \leq \tilde{p}_i^+ \quad \forall i \in N \\ & \quad \tilde{q}_i^- \leq \tilde{q}_i \leq \tilde{q}_i^+ \quad \forall i \in N \\ & \quad \tilde{v}_i^- \leq \tilde{v}_i \leq \tilde{v}_i^+ \quad \forall i \in N \end{aligned}$$

Some very interesting results about this relaxation are given in [11]. The Conic Relaxation program provides a lower bound on the Angle Relaxation model and the original problem since the initial feasible set is enlarged by the inequality constraint. However, a theorem allows us to link the optimal solution of the Conic Relaxation with the one obtained from the Angle Relaxation program.

Theorem 2.1 (Exact Conic Relaxation [11]). *Suppose $\tilde{d}p_i$ and $\tilde{d}q_i$ are variables with no upper bounds, then any optimal solution of the Conic Relaxation program is also an optimal solution of the Angle Relaxation program.*

Therefore, this relaxation is exact provided that there are no upper bounds on the real and reactive demand of the power system. However, removing upper bounds on the loads may be unrealistic [16]. Moreover, if the considered power system is radial which means that its representation as a graph is a tree, then the Angle Relaxation is exact compared to the original OPF problem [11].

2.2.2 Direct Current Optimal Power Flow

A popular approximation of the OPF is the **Direct Current Optimal Power Flow** (DCOPF) model. This approximation is a linearization of the power flow equations and is therefore easier to solve than a nonlinear OPF. Some assumptions are made in order to linearize the OPF model. For more details on the approximation see [2]. The four assumptions are:

1. $\tilde{v}_i$ is supposed to be constant in magnitude (equal to 1 [p.u.])

2. Conductances are negligible: $\tilde{G}_i \approx 0$
3. Voltage phase angle differences are very small: $\sin(\tilde{\theta}_i - \tilde{\theta}_j) \approx \tilde{\theta}_i - \tilde{\theta}_j$
4. Reactive power flows are negligible: $\tilde{Q}_{i,j}$, $\tilde{q}_i$ and $\tilde{d}q_i \approx 0$

Considering a convex objective function $h(\tilde{p}_i, \tilde{d}p_i, \tilde{P}_{i,j}, \tilde{\theta}_i)$ and these assumptions, the DCOPF optimization program can be described by:

$$\begin{aligned}
& \min_{\tilde{p}_i, \tilde{d}p_i, \tilde{P}_{i,j}, \tilde{\theta}_i} && h(\tilde{p}_i, \tilde{d}p_i, \tilde{P}_{i,j}, \tilde{\theta}_i) \\
& \text{subject to} && \tilde{p}_i + \sum_{j:j \rightarrow i} \tilde{P}_{j,i} = \tilde{d}p_i + \sum_{i:i \rightarrow j} \tilde{P}_{i,j} \quad \forall (i,j) \in E \\
& && \tilde{P}_{i,j} = \tilde{B}_{i,j}(\tilde{\theta}_i - \tilde{\theta}_j) \quad \forall (i,j) \in E \\
& && \tilde{P}_{i,j}^- \leq \tilde{P}_{i,j} \leq \tilde{P}_{i,j}^+ \quad \forall (i,j) \in E \\
& && \tilde{p}_i^- \leq \tilde{p}_i \leq \tilde{p}_i^+ \quad \forall i \in N \\
& && \tilde{\theta}_{hub} = 0
\end{aligned}$$

If the last constraint of the problem is removed then, by translating the phase angles values by some fixed constant, identical flows are obtained. In order to avoid this problem, a reference node, called hub node, has to be defined and its phase angle value has to be equal to 0 [2].

2.3 Unit Commitment

Several decisions in power system operations for one time period need to be performed in conjunction with choices that will be taken for future time periods. Indeed, some generators are governed by operating constraints that give rise to dependent operating decisions for these generators across time periods. Some examples of these operating constraints are: minimum up and down times, ramp rates limits and so on. Therefore, several decisions must be taken in advance in order to balance supply and demand for all the time periods of the scheduling horizon. The **Unit Commitment** problem allows us to account for time dependent operating constraints and costs while programming the generators' on-off status over a scheduling horizon. This problem is used in practice to schedule operations one day ahead based on forecasted demand, renewable energy and system state. Another problem is used after defining the state of units for all time periods in order to balance demand using whatever unit is available. This operation is called the OPF and is described in section 2.2 in opposition to the day-ahead scheduling. This part of the document will be dedicated to the description of the constraints used in the **Unit Commitment** problem. However, this problem is presented extensively in [17], [18] and [19].

2.3.1 Variables

In order to record the on-off status of the generation units, it is crucial to introduce binary variables $u_{g,t}$. These variables will indicate whether a unit is on if it is equal to 1 or off if the variable is equal to 0 for a certain generator g at a given time period t .

Due to start-up and shutdown limitations while operating a unit, other binary variables have to be introduced. A new binary variable $su_{g,t}$ will represent the start-up of a unit g at time period t if equal to 1. In a similar way, a new variable $sd_{g,t}$ will describe the shutdown of a unit g at time period t .

Finally, the amount of power produced by each unit is a continuous decision variable described by $p_{g,t}$.

2.3.2 Constraints

All the constraints of the *Unit Commitment* problem, used in the remainder of the document, are presented in this section. Among the constraints linked with the operation of a generation unit, some will not be presented here because they are not considered in the modelling of the transmission and the distribution networks. These constraints are the initial conditions of the unit at the beginning of the scheduling horizon, technical minimum power output, start-up and shutdown profiles, temperature segments and reserve requirements.

Minimum Up/Down Times

First, the minimum up and down time constraints are presented. Generators require a minimum amount of time to cool down once they are shut down. Equivalently, a unit can not be shut down right after it has been turned on; some time has to elapse before being able to do so. Let the minimum up time and down time be respectively given by UT_g and DT_g . These constraints are accounted for as follows:

$$\sum_{\tau \in [t-UT_g+1, t], \tau \geq 1} su_{g,\tau} \leq u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \quad (2.10)$$

$$\sum_{\tau \in [t-DT_g+1, t], \tau \geq 1} sd_{g,\tau} \leq 1 - u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \quad (2.11)$$

Equation (2.10) describes the minimum time the unit has to be on after a start-up and equation (2.11) concerns the minimum time a unit has to be shut down before it can be started up again.

Example 2.1 (Minimum Up/Down Times). *Consider a system with one generator scheduled for 12 hours. The on/off status of the generation unit is shown in figure 2.4. The generator has a minimum up time of 2 and down time of 3. Therefore, the orange marks are free choices of the operator for the status of the unit. Purple marks correspond to states forced by the generator due to minimum up and down times.*

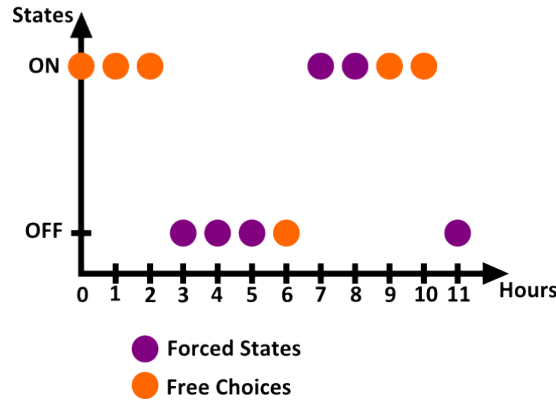


Figure 2.4: Illustration of minimum up and down times of a generator on its On/Off status over 12 hours

Transitions

In order to describe scheduling constraints, it is fundamental to express the link that exists between start-up, shutdown and unit status variables. The unit commitment decision $u_{g,t}$ is linked to the start-up and shutdown decisions in the following way:

$$u_{g,t} = u_{g,t-1} + su_{g,t} - sd_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \quad (2.12)$$

Dispatchable Production

While a generation unit is running, its production is not allowed to go beyond a certain value (p_g^+). This constraint represents the technical maximum of a running generator. However, if a unit is off then obviously it can not produce any power. This generation constraint can be expressed by the following equation:

$$p_{g,t} \leq p_g^+ u_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \quad (2.13)$$

Ramp Rates

From one period to the next, the production of a generation unit is not allowed to change too fast, either in the upward or downward direction. The maximum increase of power generation from one time period to the next is called the ramp up rate limit (R_g^+). Equivalently, the maximum decrease of production from one period to the next is the ramp down rate limit (R_g^-). The ramp constraints are therefore expressed by the following statements:

$$p_{g,t} - p_{g,t-1} \leq R_g^+ \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \quad (2.14)$$

$$p_{g,t-1} - p_{g,t} \leq R_g^- \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \quad (2.15)$$

Example 2.2 (Ramp Rates). *Consider a generator with a ramp down rate limit of 15 [MW/h] and a ramp up rate limit of 25 [MW/h]. Let this generator produce 100 [MW] at hour 0. Due to ramp rates constraints, this generator can only produce between 85 and 125 [MW] at hour 1. The orange surface of figure 2.5 represents the possible outputs of this generator at a certain time.*

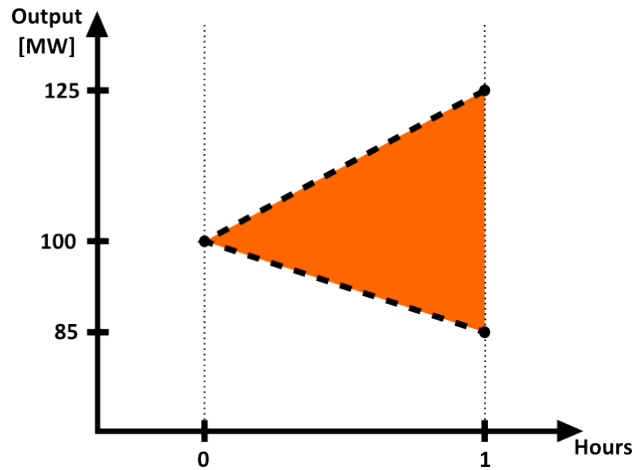


Figure 2.5: Illustration of the effect of ramp rates constraints on generator's production presented in example 2.2

Remarks

The *Unit Commitment* problem is therefore a model consisting of expressions (2.10) to (2.15). In practice, this is a large-scale mixed integer program and is quite challenging to solve because of its nonconvex nature due to binary variables.

2.4 Models

The previous sections of this chapter have described two essential optimization problems in order to outline the operation of electrical power systems. These two approaches will be used in the following section in order to detail the working optimization problem of this document.

2.4.1 Transmission

As explained previously, the transmission part of an electrical power system is composed of generators along with high voltage transmission lines. A representation of this part of the network is shown in figure 2.6.

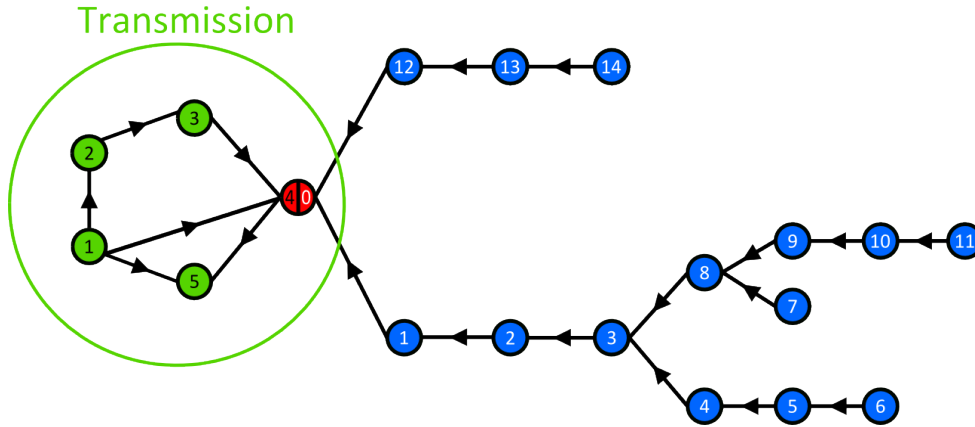


Figure 2.6: Electrical power system representation with highlighted transmission part

This part of the network can be modelled using the linear approximation of the OPF problem called the DCOPF problem. This approximation is accurate for this part of the power grid due to the high-voltage nature of the transmission lines. Indeed, for this type of transmission lines, the assumptions made for the DCOPF problem are verified. The DCOPF problem allows taking into account the effects on the system due to the transmission of electricity. However, the transmission part of the network is also the one containing the generators. Therefore, the unit commitment model will be added to the DCOPF model in order to represent accurately the behaviour of the transmission part of the electrical power system. The entire optimization problem describing

the behaviour of the transmission network is therefore given by:

$$\begin{aligned}
& \min_{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}} \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} \\
& \text{subject to} \quad \sum_{g \in \mathcal{G}_n} p_{g,t} + \sum_{k=(\cdot, n) \in \mathcal{K}} f_{k,t} = d_{n,t} + \sum_{k=(n, \cdot) \in \mathcal{K}} f_{k,t} \quad \forall n \in \mathcal{N}, t \in \mathcal{T} \\
& \quad f_{k,t} = B_k(\theta_{m,t} - \theta_{n,t}) \quad \forall k = (m, n) \in \mathcal{K}, t \in \mathcal{T} \\
& \quad u_{g,t} = u_{g,t-1} + su_{g,t} - sd_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \quad \sum_{\tau \in [t-UT_g+1, t], \tau \geq 1} su_{g,\tau} \leq u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \\
& \quad \sum_{\tau \in [t-DT_g+1, t], \tau \geq 1} sd_{g,\tau} \leq 1 - u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \\
& \quad p_{g,t} - p_{g,t-1} \leq R_g^+ \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \quad p_{g,t-1} - p_{g,t} \leq R_g^- \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \quad \theta_{hub,t} = 0 \quad \forall t \in \mathcal{T} \\
& \quad p_{g,t} \leq p_g^+ u_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \\
& \quad -f_k^+ \leq f_{k,t} \leq f_k^+ \quad \forall k \in \mathcal{K}, t \in \mathcal{T} \\
& \quad u_{g,t}, su_{g,t}, sd_{g,t} \in \{0, 1\} \quad \forall g \in \mathcal{G}, t \in \mathcal{T}
\end{aligned}$$

For the sake of simplicity, this optimization program will be expressed in a compact form using the following equations:

$$\begin{aligned}
& \min_{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}} \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} \\
& \text{subject to} \quad (su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}) \in \mathcal{F}_T
\end{aligned}$$

where $\mathcal{F}_T$ is the set of variables values verifying the constraints of the transmission model. A feasible solution of the transmission optimization program is going to belong to this set.

2.4.2 Distribution

As described previously, the distribution part of an electrical power system contains the low-voltage transmission lines which are connected to industrial and residential loads. In practice, distribution networks are radial as this can be observed in figure 2.7. Radial means that the network does not contain any cycle and therefore, this representation as a graph is a tree. Flow in lines is assumed to point towards the root of the graph (node 0).

In order to represent the behaviour of the network's distribution part, the Conic Relaxation of the OPF problem will be used. This choice is due to the fact that the angle relaxation of the OPF problem is exact for radial networks [11]. However, due to the radial nature of the distribution network, lines can be numbered as nodes. Indeed, a node in the distribution network will have a unique ancestor and a set of children, therefore, lines going from a node i to its ancestor A_i can be numbered as i . Moreover, without loss of generality, the voltage of node 0 is considered to be fixed at 1 [p.u.] and the injection of power at node 0 is supposed to be equal to 0. Therefore, $v_i^+ = v_i^- = 1$ for node i equal to 0. The behaviour of the distribution network can

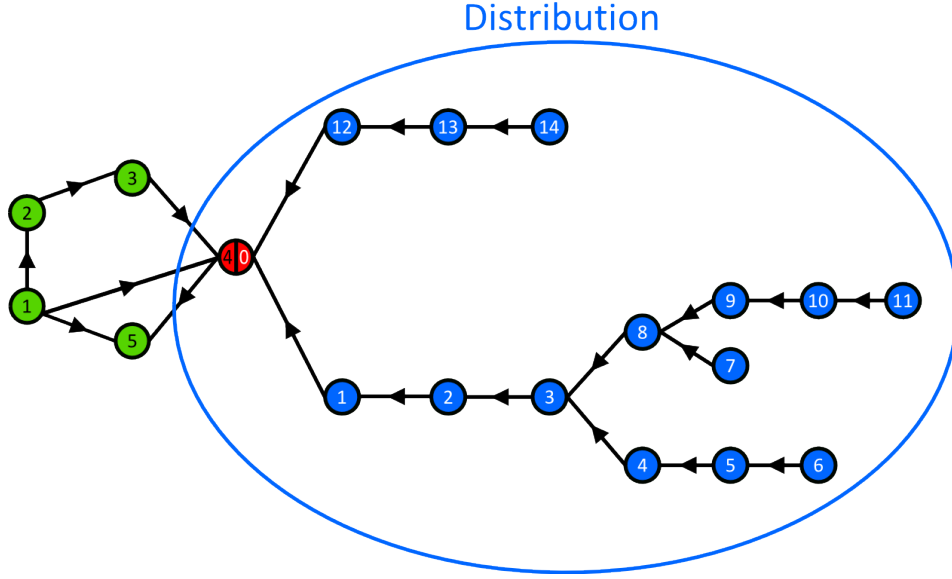


Figure 2.7: Electrical power system representation with highlighted distribution part

be expressed by the following optimization problem:

$$\begin{aligned}
& \min_{P_{i,t}, Q_{i,t}, v_{i,t}, l_{i,t}, p_{i,t}, q_{i,t}} \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}} MC_i p_{i,t} \\
& \text{subject to} \quad v_{i,t} - 2(R_i P_{i,t} + X_i Q_{i,t}) + l_{i,t}(R_i^2 + X_i^2) = v_{A_i,t}, \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& \quad 4P_{i,t}^2 + 4Q_{i,t}^2 + (l_{i,t} - v_{i,t})^2 \leq (l_{i,t} + v_{i,t})^2, \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& \quad - \sum_{c_0 \in \mathcal{C}_0} (P_{c_0,t} - l_{c_0,t} R_{c_0}) + dp_{0,t} + G_0 v_{0,t} = 0 \quad \forall t \in \mathcal{T} \\
& \quad P_{i,t} - \sum_{c_i \in \mathcal{C}_i} (P_{c_i,t} - l_{c_i,t} R_{c_i}) - p_{i,t} + dp_{i,t} + G_i v_{i,t} = 0 \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& \quad - \sum_{c_0 \in \mathcal{C}_0} (Q_{c_0,t} - l_{c_0,t} X_{c_0}) - q_{0,t} + dq_{0,t} - B_0 v_{0,t} = 0 \quad \forall t \in \mathcal{T} \\
& \quad Q_{i,t} - \sum_{c_i \in \mathcal{C}_i} (Q_{c_i,t} - l_{c_i,t} X_{c_i}) - q_{i,t} + dq_{i,t} - B_i v_{i,t} = 0, \quad \forall i \in \mathcal{I}, t \in \mathcal{T} \\
& \quad p_i^- \leq p_{i,t} \leq p_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T} \\
& \quad q_i^- \leq q_{i,t} \leq q_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T} \\
& \quad v_i^- \leq v_{i,t} \leq v_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}
\end{aligned}$$

The distribution model will also be written in a concise way. The optimization program is given by the following set of equations:

$$\begin{aligned}
& \min_{P_{i,t}, Q_{i,t}, v_{i,t}, l_{i,t}, p_{i,t}, q_{i,t}} \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}} MC_i p_{i,t} \\
& \text{subject to} \quad (P_{i,t}, Q_{i,t}, v_{i,t}, l_{i,t}, p_{i,t}, q_{i,t}) \in \mathcal{F}_D
\end{aligned}$$

where the previous set of distribution constraints is accounted for in set $\mathcal{F}_D$.

2.4.3 Linking Constraint

Since the distribution network starts from a node in the transmission network, one needs to add a linking constraint between the two networks in order to link both problems. The convention

used for this linking constraint is explained in figure 2.8.

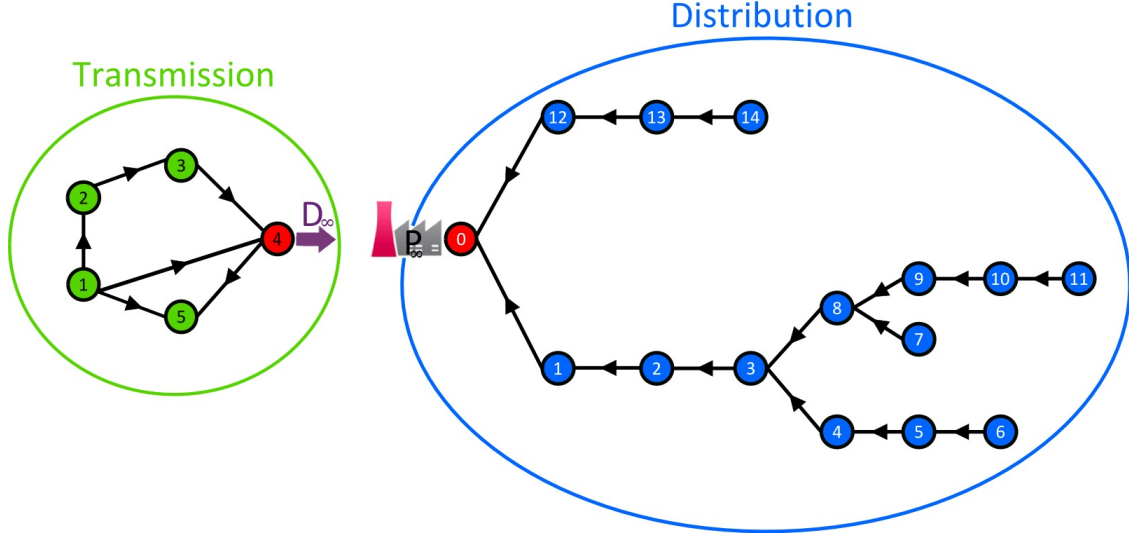


Figure 2.8: Illustration of the convention used in order to link the transmission and the distribution networks

The linking constraint expresses the fact that, at the transmission side, a load D_∞ is withdrawn in order to transmit it as a power generated by a virtual generator for the distribution network given by P_∞ . The linking constraint is then given by :

$$P_\infty - D_\infty = 0$$

However, there will not be only one distribution network connected to only one node of the transmission network. In fact, there will be several distribution networks attached to each node of the transmission network. Therefore, the optimization problem presented in section 2.4.2 is computed for each distribution network present in the power system. For the sake of simplicity, the distribution network will be indexed by the number of the bus in the transmission network it is linked to and an index going from 1 to the number of distribution network attached to this transmission node. For example, if node 4 of the transmission network is linked to 3 different distribution networks, these distribution networks will be indexed by (4,1), (4,2) and (4,3) respectively. The set of distribution networks connected to a node n of the transmission network will be given by the set $\mathcal{T}_n$. The linking constraint can therefore be expressed in two different ways:

$$\sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n} - D_\infty^{t,n} = 0 \quad \forall t \in \mathcal{T}, n \in \mathcal{N} \quad (2.16)$$

$$P_\infty^{t,n,t_n} - D_\infty^{t,n,t_n} = 0 \quad \forall t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n \quad (2.17)$$

The expression of the constraint presented in equation (2.16) will be referred to as the first formulation of the linking constraint. Equation 2.17 will be called the second formulation of the linking constraint for the remaining of this document. The two formulations of the linking constraints will be compared in terms of performances (solving time, precision, etc.) in chapter 5.

2.4.4 Total Problem

The global optimization problem considering transmission and distribution parts of the power network as a whole is given by the following set of equations for the first formulation of the linking constraint. The global optimization program using the second formulation of the linking constraint can be obtained similarly by replacing equation (2.16) by equation (2.17) in the following optimization program. The global program containing the second formulation of the linking constraint is stated in appendix A.

$$\begin{aligned}
& \min_{\substack{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, \\ f_{k,t}, \theta_{n,t}, P_{i,t,n,t_n}, \\ Q_{i,t,n,t_n}, v_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, \\ p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ P_{\infty}^{t,n,t_n}, D_{\infty}^{t,n}}} \\
& \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n} + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} \\
& \text{subject to} \\
& \sum_{g \in \mathcal{G}_n} p_{g,t} + \sum_{k=(\cdot, n) \in \mathcal{K}} f_{k,t} = d_{n,t} + \sum_{k=(n, \cdot) \in \mathcal{K}} f_{k,t} \quad \forall n \in \mathcal{N} : \mathcal{T}_n = \emptyset, t \in \mathcal{T} \\
& \sum_{g \in \mathcal{G}_n} p_{g,t} + \sum_{k=(\cdot, n) \in \mathcal{K}} f_{k,t} = d_{n,t} + D_{\infty}^{t,n} + \sum_{k=(n, \cdot) \in \mathcal{K}} f_{k,t} \quad \forall t \in \mathcal{T}, n \in \mathcal{N} : \mathcal{T}_n \neq \emptyset \\
& f_{k,t} = B_k(\theta_{m,t} - \theta_{n,t}) \quad \forall k = (m, n) \in \mathcal{K}, t \in \mathcal{T} \\
& u_{g,t} = u_{g,t-1} + su_{g,t} - sd_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \sum_{\tau \in [t-UT_g+1, t], \tau \geq 1} su_{g,\tau} \leq u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \\
& \sum_{\tau \in [t-DT_g+1, t], \tau \geq 1} sd_{g,\tau} \leq 1 - u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \\
& p_{g,t} - p_{g,t-1} \leq R_g^+ \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& p_{g,t-1} - p_{g,t} \leq R_g^- \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \theta_{hub,t} = 0 \quad \forall t \in \mathcal{T} \\
& v_{i,t,n,t_n} - 2(R_i P_{i,t,n,t_n} + X_i Q_{i,t,n,t_n}) + l_{i,t,n,t_n}(R_i^2 + X_i^2) = v_{A_i,t,n,t_n}, \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& 4(P_{i,t,n,t_n})^2 + 4(Q_{i,t,n,t_n})^2 + (l_{i,t,n,t_n} - v_{i,t,n,t_n})^2 \leq (l_{i,t,n,t_n} + v_{i,t,n,t_n})^2, \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& - \sum_{c_0 \in \mathcal{C}_0} (P_{c_0,t,n,t_n} - l_{c_0,t,n,t_n} R_{c_0}) - p_{0,t,n,t_n} - P_{\infty}^{t,n,t_n} + dp_{0,t,n,t_n} + G_0 v_{0,t,n,t_n} = 0 \\
& \quad \forall t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n \\
& P_{i,t,n,t_n} - \sum_{c_i \in \mathcal{C}_i} (P_{c_i,t,n,t_n} - l_{c_i,t,n,t_n} R_{c_i}) - p_{i,t,n,t_n} + dp_{i,t,n,t_n} + G_i v_{i,t,n,t_n} = 0 \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& - \sum_{c_0 \in \mathcal{C}_0} (Q_{c_0,t,n,t_n} - l_{c_0,t,n,t_n} X_{c_0}) - q_{0,t,n,t_n} + dq_{0,t,n,t_n} - B_0 v_{0,t,n,t_n} = 0 \\
& \quad \forall t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& Q_{i,t,n,t_n} - \sum_{c_i \in \mathcal{C}_i} (Q_{c_i,t,n,t_n} - l_{c_i,t,n,t_n} X_{c_i}) - q_{i,t,n,t_n} + dq_{i,t,n,t_n} - B_i v_{i,t,n,t_n} = 0, \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& \sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} = 0 \quad \forall t \in \mathcal{T}, n \in \mathcal{N} : \mathcal{T}_n \neq \emptyset \\
& u_{g,t}, su_{g,t}, sd_{g,t} \in \{0, 1\} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \\
& -f_k^+ \leq f_{k,t} \leq f_k^+ \quad \forall k \in \mathcal{K}, t \in \mathcal{T} \\
& p_{g,t} \leq p_g^+ u_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \\
& p_i^- \leq p_{i,t,n,t_n} \leq p_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& q_i^- \leq q_{i,t,n,t_n} \leq q_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& v_i^- \leq v_{i,t,n,t_n} \leq v_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset
\end{aligned}$$

By means of the compact formulations of the distribution and the transmission models, the global optimization program can be reduced to the following expressions:

$$\begin{aligned}
& \min_{\substack{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \\ \theta_{n,t}, P_{i,t,n,t_n}, Q_{i,t,n,t_n}, \\ v_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}, D_{\infty}^{t,n}}} \\
& \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n} + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} \\
& \text{subject to} \quad \sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} = 0 \quad \forall t \in \mathcal{T}, n \in \mathcal{N} : \mathcal{T}_n \neq \emptyset \\
& \quad (su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}, D_{\infty}^{t,n}) \in \mathcal{F}_T \\
& \quad (P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}) \in \mathcal{F}_D^{n,t_n}
\end{aligned}$$

2.5 Motivating Example

This section presents an example of a power system which leads to a situation where considering the network as a whole furnishes better results in terms of system's operation than a traditional approach. Over the last decade, the distribution grid was treated as a known equivalent load when modelling the transmission grid [20]. The equivalent load is obtained by adding up all the electricity power withdrawn from the distribution network. In order to use a more accurate model, the set of generators located in the distribution network will be aggregated to the transmission node where the distribution grid is attached. This model is called the traditional approach. For the sake of simplicity, only the technical maximum output power constraint was considered as unit commitment features. The other characteristics of a unit commitment problem were not taken into account in this example (ramp rates limits, minimum up and down times, etc.).

The total power system used for this example is represented in figure 2.9. All distribution nodes have a generator assumed to be solar panels and there are 5 generators in the transmission network. The data sets used for this example are a 5 bus test network given by the MatPower® test case PJM 5-bus system [21] for the transmission network and a power system used in [22] for the radial distribution network. In order to obtain a distribution network that deals with smaller power flows than the transmission system, the maximum and minimum power outputs were divided by 25 and resistances along with reactances were multiplied by the same number. For more details about parameters of the distribution network, see [22].

The data representing power systems are given in [p.u.] (per unit) with a 100[MVA] base. This means that the data were made dimensionless with respect to a power base 100[MW] and a voltage base specific to the type of network considered. The data representing the branches of the network are given in table 2.2, the power flows on branches is not considered to be limited.

Branch ID	Bus From	Bus To	Susceptance [p.u.]
1To2	1	2	0.00712
1To4	1	4	0.00658
1To5	1	5	0.03126
2To3	2	3	0.01852
3To4	3	4	0.00674
4To5	4	5	0.00674

Table 2.2: Detailed Branch Data

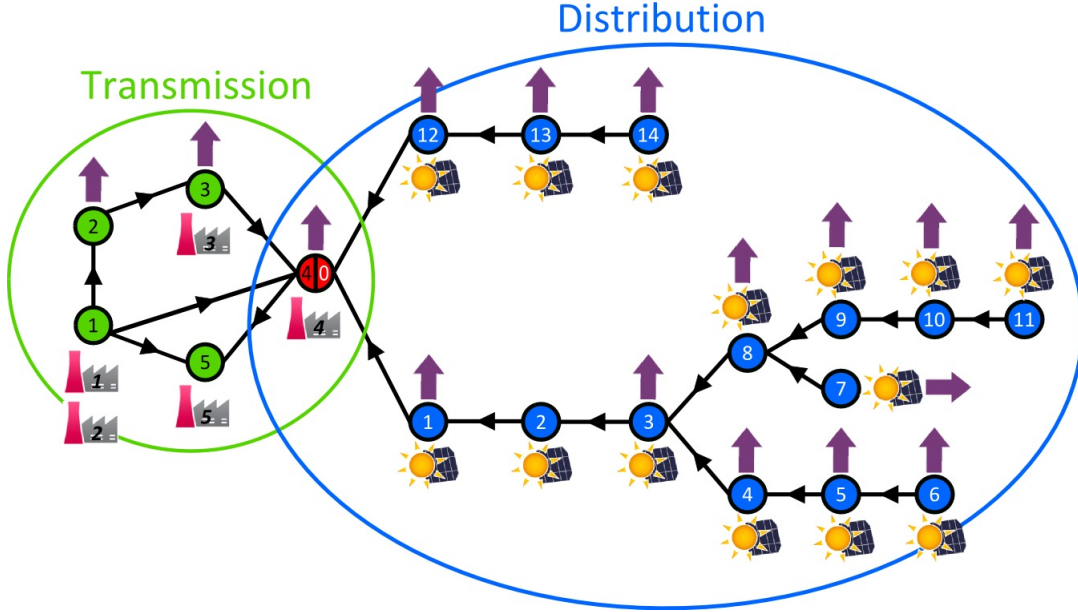


Figure 2.9: Illustration of the global power system used for the motivating example

The information on bus loads are given in table 2.3. Moreover, information about generators present in the transmission network are given in table 2.4.

Bus ID	Load [p.u.]
1	0
2	3
3	3
4	4
5	0

Table 2.3: Detailed Load Data

Generators ID	Bus	Marginal Cost [\$/p.u.]	Maximum Power [p.u.]
Gen1	1	14	0.4
Gen2	1	15	1.7
Gen3	3	30	1.75
Gen4	4	40	2
Gen5	5	10	6

Table 2.4: Detailed Generators Data

We are considering solar panels that can deliver 0.016 [p.u.] as maximum power in each node except the root (node 0) with a zero cost. By solving the unit commitment problem using the co-optimization method and the traditional method, we obtain very different results which are shown in table 2.5.

By considering the results presented in table 2.5, we observe that the traditional method will not start up generator 4 because it will consider using all the available solar power from the

Generators ID	Traditional	Co-optimization
Gen1	0.4	0.4
Gen2	1.7	1.7
Gen3	1.73	1.75
Gen4	0	0.15
Gen5	6	6

Table 2.5: Solution of the motivating example

distribution network. Indeed, all generators from the distribution network aggregated at node 4 of the transmission network produce each 0.016 [p.u.] of electrical power. However, due to physical constraints governing the behaviour of the distribution network, the co-optimization method is forced to start up the fourth generator to fulfill the entire demand of the network. This leads to the conclusion that not all solar panels can be used in order to fulfill the demand of the network because of constraints related with the physical behaviour of the power system and voltage limits in the distribution network. We can note the counter-intuitiveness of the results about renewable energy. Therefore, while scheduling these generators, the operators would have been wrong to turn off generator 4. This example demonstrates that the co-optimization method is worth being implemented since it leads to different and more efficient conclusions than the traditional method.

Chapter 3

State of the Art

3.1 Optimal Power Flow Convex Relaxations

The Optimal Power Flow (OPF) model, introduced in section 2.2, is an essential optimization problem in electrical power systems analysis which allows to account for physical constraints governing the behaviour of power systems. This mathematical program aims to be solved every few minutes for industry purposes. Moreover, as seen in section 2.2, this problem is a nonconvex optimization problem which in practice is hard to solve. Consequently, due to these two characteristics, the scientific and industrial worlds are looking for efficient and fast solutions to deal with this fundamental program. First, researches concentrated on the enhancement of nonlinear methods. However, nonlinear optimization algorithms do not provide any quantification for the quality of the obtained solution [23]. Therefore, researchers have focused on simplifying the original problem rather than improving the performances of nonlinear algorithms. Over the last decade, they have consequently shown a recent interest in convex relaxations of this nonconvex problem.

In the optimization field, linear programs are the most practicable problems to solve thanks to efficient algorithms. Consequently, the OPF problem was first approximated by a popular linear problem called the Direct Current Optimal Power Flow (DCOPF) approximation which uses a linearization of the power flow equations. This approximation of the Alternating Current Optimal Power Flow (ACOPF) problem, other designation for the OPF program, exploits some physical properties of the power flows in typical power systems to state assumptions on the topology of the network to derive a linear program [2]. These assumptions consider properties of power systems as tight bounds on voltage magnitudes at buses and small voltage angle differences between buses and others [23]. A lot of publications were released on the subject such as, for example, [2] and [24]. In 2006, Van Hertem et al. investigated the usefulness of the DCOPF formulation for active power flow analysis with monitoring flows devices. They also compared the validity of the DCOPF formulation to the ACOPF approach.

The DCOPF problem is widespread in the planning and control of power systems, however it ignores important physical aspects governing power flows in networks such as voltage magnitudes and reactive powers. Unfortunately, these features are necessary in many fields to certify voltage stability and the feasibility of the approximated solution. Therefore, Coffrin et al. proposed a linearized version of the OPF problem that includes reactive powers and voltage magnitudes called the LPAC model [25]. The ignorance of reactive powers and voltage magnitudes in the DCOPF formulation is a drawback that can be avoided by solving a set of power flow equations using the DCOPF optimal solution to compute feasible reactive powers and voltage magnitudes

corresponding to the acquired optimal set of power flows. Nonetheless, this approach can not guarantee any optimality on the obtained solution [23].

The DCOPF formulation has a lot of inconveniences, as inferred from the previous paragraphs, presented in [26]. First, even if this approximation is useful in some applications, it can not be applied to all types of problems. Moreover, sometimes, the solution obtained by the DCOPF program may not satisfy nonlinear power flow equations. Therefore, the user will have to tighten the problem and solve it again. This process may reduce the efficiency of the computation. Finally, it does not take into account voltage magnitudes and reactive powers. Compared to the disadvantages of such an approximation, convex relaxations of the OPF problem manifest several advantages. Indeed, a convex relaxation provides the possibility to check the quality of the acquired solution. Furthermore, if a convex relaxation of the problem is infeasible, then this is a certificate for the infeasibility of the original problem. Two types of convex relaxations were explored over the last decade: semi-definite programming formulations SDP and second-order cone programs SOCP. A survey [26] summarizes the particularities of both types of convex relaxations while other publications preferred to specialize in one type.

A first important observation was made about the composition of the OPF program. In fact, it was noticed that the OPF can be formulated as a quadratically constrained quadratic problem. These types of programs are generally nonconvex and NP-hard. However, such a formulation can be relaxed to a semi-definite program ([26], [27]). This involves reformulating the problem as a semi-definite program and using the interior point algorithm to solve it. One semi-definite formulation, based on the bus injection model for power flow equations, is presented in [27]. The convex relaxation based on semi-definite programming was popular at the beginning of this research field since it is empirically exact for a restricted class of applications. Nonetheless, there exists some counterexamples that have been pointed out in the literature [28]. In this case, it is really difficult to put a physical meaning on the acquired optimal solution of the semi-definite program [23]. Furthermore, the semi-definite formulation could not solve extremely large systems due to limitations of memory and time. In fact, these relaxations can be computationally prohibitive as the problem size increases [19]. As a result to this drawback, some publications ensued developing algorithms to solve these semi-definite relaxations by bearing in mind that power systems are sparse. The exploitation of graph sparsity of power systems is, for example, adopted in [29] and [30]. An other type of semi-definite relaxation of the OPF problem is presented in [31]. This paper proposes a semi-definite formulation for the dual of the OPF problem. Two conditions for strong duality, one necessary and one sufficient, are also stated in this publication in order to be able to recover a global optimum solution of the relaxed OPF problem.

To the best of my knowledge, another convex relaxation was proposed first by Jabr et al. in 2006 to compute the OPF optimal solution [32]. The OPF problem is formulated as a second-order cone program. This relaxation was studied extensively by [11] and [12]. It consists in ignoring phase angles for current and voltage along with a relaxed quadratic inequality constraint instead of an equality, as introduced in the SOCP relaxation part of section 2.2. This model is also used in a lot of publications ([16], [33], [34] and [22]). Farivar et al. proves that this relaxation is exact for radial networks provided unbounded reactive and real power consumption. For meshed networks, the angle relaxation is not exact, therefore the authors create angle recovery algorithms. Kocuk et al., to alleviate this issue, proposed three other SOCP relaxations for the OPF problem. The nonconvexity in the OPF problem are generated by the quadratic equality constraint but also due to the use of the arctangent function on voltage angles. Instead of forgetting the angle constraint as in the classic SOCP relaxation, the idea in this publication is to

add convex constraint involving the arctangent function to this classic relaxation in order to strengthen it. Three approaches are proposed:

1. The arctangent is reformulated using a polynomial whose degree is proportional to the length of the cycle.
2. A polyhedral envelope for the arctangent function is constructed in three dimensions.
3. The classic **SOCP** relaxation is strengthened by dynamically generating valid linear inequalities that separates the **SOCP** solution from the **SDP** cone constraints over a cycle.

The usefulness of these three techniques is only observed for meshed networks since the angle relaxation is exact for radial networks.

Finally, some researchers focused on the comparison of all these relaxations techniques ([29], [31], [26] and [23]). From their common conclusions, the high computational time required for large size networks needed by the **SDP** relaxations can be highlighted. The **SOCP** relaxations offer therefore a computationally appealing option to the classic **SDP** relaxations for large-scale systems [23]. This is a reason why we choose to use the classic **SOCP** relaxation in this document along with the fact that the angle relaxation used for the classic **SOCP** relaxation is exact for radial networks.

3.2 Unit Commitment Formulations

The unit commitment problem, introduced in section 2.3, depicts one important part of this work and is also an essential application from the electric power industry. This model allows to schedule a set of electric power generators in response to forecasted demand while minimizing the production cost of each generator. Over the last decade, researchers have conducted extensive research on modeling and solving unit commitment problems leading to several formulations. Some formulations of the unit commitment problem are given in [35], [36], [17], [37].

Even if there exists a large number of formulations of the unit commitment problem, these models can be divided into two groups: a single-binary formulation (**1BIN**) and a three-binary formulation (**3BIN**). The **1BIN** formulation only considers one type of binary variable that describes the on-off status of a unit. On the contrary, the **3BIN** formulation needs 3 types of binary variables: one is given by the on-off status of the unit as for the **1BIN** formulation. The other two variables characterize respectively if a unit starts up at a certain time period and the other one if a unit shuts down at this time period [19].

Several surveys allow to have a broad overview of all the formulations used in the literature. They also compared the efficiency of each formulation in terms of computing time and closeness of the feasible set compared to its convex hull. Among these surveys, Pandzic et al. describes 4 ways to express the unit commitment problem along with a comparison of the **1BIN** and the **3BIN** formulations [38]. An other survey [37] compares formulations of several constraints linked to the unit commitment problem. One important conclusion has been brought out by the majority of publications on the subject: Commercial solvers are more efficient on **3BIN** formulation despite the fact that it is composed by more variables and constraints ([36], [19] and [38]). Due to this conclusion, I choose to use a **3BIN** formulation as unit commitment model for this work. Several publications focused on the available formulations of only a small set of constraints belonging to the unit commitment problem. These formulations are also compared

in terms of feasible set and convex hull. For instance, Rajan et al. published conclusions about several formulations of the minimum up and down constraints taking in consideration start up costs in their formulation [39].

3.3 Distributed Algorithms for Large-Scale Interconnected Power Systems

Due to the scale of power systems and optimization problems used to operate efficiently these electrical networks, it has become of main interest to develop efficient algorithms. Their objective is to capture the complexity of contemporary problems while being sufficiently scalable in order to process large scale power networks by using, for example, parallelization. Therefore, a series of algorithms have been proposed to support distributed or parallel computation for solving some optimization problems in large-scale power systems. Only a small number of these publications will be presented but there are a lot of works on distributed algorithms for power systems.

Due to numerous publications in the domain, the focus of this state-of-the-art will be on researches using a distributed algorithm in order to compute the solution of a relaxation of the OPF problem. For example, Peng et al. used a distributed algorithm in order to compute the solution of the classical SOCP relaxation of the OPF problem on radial networks [34]. They divided the problem in terms of nodes variables and applied the ADMM algorithm which will be presented in section 4.2 of this document. Using a similar idea but based on the Predictor Corrector Proximal Multipliers (PCPM) algorithm, Li et al. employed a distributed algorithm to coordinate user's demand response decisions by means of local communication with neighbors to achieve a global optimum decision pattern [16]. They used the same relaxation as the previous publication. Another researcher applied on radial networks a distributed algorithm based on primal and dual decomposition on a SDP relaxation of the OPF problem [40]. Finally, a linearized augmented Lagrangian approach is used in another publication to compute the solution of the OPF program by dividing the problem into node variables as in all the publications mentioned above [41].

Instead of considering only a part of the power system (transmission or distribution), the following publications ([8] and [42]) take advantages of a distributed algorithm in order to solve a relaxation of the OPF problem for the entire power system (transmission and distribution) by breaking the network into small pieces which are, in this case, nodes.

3.4 Global Analysis of the Whole Power System

As explained in the introduction (chapter 1), the literature has focused during the last decade on transmission and distribution parts of the power system as two disjoint problems. However, since the recent large-scale penetration of renewable energy such as rooftop solar panels and also electrical vehicle, there is a growing need to consider these two parts together as a global network. Recently, a small number of publications have developed that idea of taking into account both parts of the network together. Kraning et al. proposed a distributed algorithm by decomposing the network into nodes in order to compute the solution of the OPF problem by considering transmission and distribution as a global entity [42]. The same kind of purpose was reached in another publication [8] by adding reserve requirements to the OPF problem. Furthermore, Sun et al. described a distributed algorithm where distribution and transmission networks are studied as a whole in order to solve power flow equations [20]. The algorithm decomposed

the problem into subproblems: one given by the transmission network, the other one given by the distribution networks. These gave an iterative method to balance the boundary mismatch between the transmission and distribution parts of a power system in terms of voltage magnitude and power. This study was extended to state some convergence results of the algorithm [7]. Finally, a contingency analysis was made using the OPF problem by also considering the distribution and transmission parts of the network as a whole [43]. Contingency analysis allows the power system's operator to maintain the secure operation of that system when a failure occurs.

The purpose of this thesis is to use a distributed algorithm by considering the entire power system as a whole. However, this work differs from the literature since the goal is to solve a SOCP relaxation of the OPF problem with added unit commitment constraints. Moreover, the decomposition approach will divide the problem into subproblems containing networks and not only one bus such as in [20].

Chapter 4

Algorithms

The necessity of solving the global optimization problem was highlighted with an example in section 2.5. We have seen in the introduction (chapter 1) that this global model will be applied to large power systems. However, we can observe from the formulation of the global program, introduced in section 2.4, that the size of this problem will increase quickly with the number of distribution networks connected to the nodes of the transmission networks. Indeed, the number of variables and constraints in the problem will expand significantly by adding more distribution networks. Therefore, valuable tools have to be developed in order to solve this problem efficiently for large scale power system. Consequently, several algorithmic schemes are presented in this chapter in order to address the global optimization problem in a way that requires less computational efforts. First, the *Level* method will be presented, followed by the *Alternating Direction Method of Multipliers* (ADMM) algorithm and the chapter will end with the introduction of the *Caramanis* method.

4.1 Level Method

The first scheme proposed to deal with the global optimization problem for large power systems is the *Level* method. This algorithm is presented extensively in [14]. Before applying this algorithm, the original problem has to be reformulated in such a way that enables the *Level* method to address it efficiently. Concepts linked with the characteristics of the original problem (such as duality, Lagrangian relaxation and dual approximation) and the new formulation of this program are presented first. After that, the algorithm is discussed in details along with related optimization program of the method. Finally, the stopping criterion of the algorithm will be described accurately.

4.1.1 Characteristics and Reformulation of the Original Problem

In order to use the Level method to deal with the original problem, the notions of Lagrangian relaxation and duality will be used to reformulate the global program which needs to be solved. Therefore, these concepts will be presented in this section. Moreover, the Level algorithm requires a linear approximation of the function to optimize. Consequently, the linear approximation of the dual function will be introduced in the second part of this section.

Lagrangian Relaxation

The Lagrangian function is an essential concept of optimization theory. The following section will therefore introduce this concept and other related ones. Most of the results appearing here

come from [44] and [45].

Consider an optimization problem in a standard form not necessarily convex:

$$\begin{aligned} \min_{x \in X} \quad & f_0(x) \\ \text{subject to} \quad & (\mu_i) : f_i(x) \leq 0, \quad i = 1, \dots, m \\ & (\nu_j) : h_j(x) = 0, \quad j = 1, \dots, p \end{aligned}$$

which is called the primal problem. Among the constraints of this optimization problem, the equality ($h_j(x) = 0$) and inequality ($f_i(x) \leq 0$) constraints make the optimization problem hard to solve while constraints due to the set $\mathcal{X}$ do not complicate the problem. The **Lagrangian** function is then introduced to remove these hard constraints:

$$\mathcal{X} \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R} : (x, \mu, \nu) \rightarrow L(x, \mu, \nu) := f_0(x) + \sum_{i=1}^m \mu_i f_i(x) + \sum_{j=1}^p \nu_j h_j(x)$$

Primal variables in this function are given by x and dual variables by $\mu \in \mathbb{R}^m$ and $\nu \in \mathbb{R}^p$. The dual variables are also called **Dual Multipliers** and they are associated with each complicating constraints. The Lagrangian function is obtained by removing complicating constraints and penalizing their violations. It may be challenging to solve the original problem with the complicating constraints, however, it may need less computational effort to minimize the Lagrangian function for a given choice of μ and ν since the complicating constraints are present as penalties in the objective function.

Then, the dual function¹ associated with the original problem can be defined for a certain choice of (μ, ν) :

$$(\mu, \nu) \in \mathbb{R}^{m \times p} \rightarrow g(\mu, \nu) := \inf_{x \in X} L(x, \mu, \nu)$$

It can be observed that the dual function provides a lower bound to the original problem if $\nu \geq 0$. For more details about the proof of this result, see [45]. Therefore, it is legitimate to maximize the dual function in order to obtain the tightest possible lower bound on the original problem and so to find appropriate values for (μ, ν) . Moreover, the dual function is always concave whatever the choice is of the objective function to optimize and the constraint set (for a detailed proof, see [45]). Consequently, the dual problem is a convex optimization problem. This process generates the Lagrange dual problem:

$$\begin{aligned} \max_{(\mu, \nu) \in \mathbb{R}^{m \times p}} \quad & g(\mu, \nu) \\ \text{subject to} \quad & \mu \geq 0 \end{aligned}$$

As described, the Lagrangian relaxation is the obtained dual program that replaces the original problem. Since the dual function provides a lower bound on the original problem, the optimal objective of the dual problem d^* and the optimal objective of the primal problem p^* are such that $d^* \leq p^*$. This relationship is called **Weak duality** [44]. Strong duality holds when both objective values are equal ($d^* = p^*$) but unlike weak duality this only happens for special types of optimization problems.

An extremely well-known utilization of the Lagrangian dual function and relaxation is for problems containing two or more sets of independent variables linked only by a few constraints

¹The inf operator can be replaced by the minimization operator if set $\mathcal{X}$ is a closed set

which make the optimization program hard to solve. Indeed, the dual function obtained after the relaxation of these constraints will generate an optimization program that can be decoupled if the objective function of the original problem is separable [46]. Our global optimization problem can be written as (see section 2.4):

$$\begin{aligned}
& \min_{\substack{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, \\ f_{k,t}, \theta_{n,t}, P_{i,t,n,t_n}, \\ Q_{i,t,n,t_n}, v_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, \\ p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ P_{\infty}^{t,n,t_n}, D_{\infty}^{t,n}}} \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n} + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} \\
& \text{subject to} \quad \sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} = 0 \quad \forall t \in \mathcal{T}, n \in \mathcal{N} : \mathcal{T}_n \neq \emptyset \\
& \quad (su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}, D_{\infty}^{t,n}) \in \mathcal{F}_T \\
& \quad (P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}) \in \mathcal{F}_D^{n,t_n}
\end{aligned}$$

By means of this formulation of the optimization problem, we can observe that this program is composed of several independent sets of variables given by the transmission and the distribution parts of the entire network. However, these sets of variables are combined by the linking constraints. Therefore, if the linking constraints are relaxed then the resulting Lagrangian dual problem is composed of smaller decoupled subproblems. This leads to the conclusion that the linking constraints are in fact constraints that make the problem hard to solve because they link the transmission and the distribution subproblems together. Therefore, the Lagrangian function is given by:

$$\begin{aligned}
L(su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}, P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}, D_{\infty}^{t,n}) = \\
\sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n} + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} + \sum_{t \in \mathcal{T}, n \in \mathcal{N}} \eta_{t,n} \left(\sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} \right) \quad (4.1)
\end{aligned}$$

We can see that the dual function which consists in minimizing the Lagrangian function can be represented as a sum of several minimization problems: one for the transmission network and one for each distribution networks connected to transmission nodes. Optimizing all this problems separately is a reasonable task compared to the global optimization problem. By doing so, we obtain the following Lagrangian dual subproblems:

Transmission Subproblem:

$$\begin{aligned}
& \min_{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}} \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} - \sum_{t \in \mathcal{T}, n \in \mathcal{N}} \eta_{t,n} D_{\infty}^{t,n} \\
& \text{subject to} \quad (su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}, D_{\infty}^{t,n}) \in \mathcal{F}_T
\end{aligned}$$

Distribution Subproblems:

$$\begin{aligned}
& \min_{\substack{P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, \\ p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}}} \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}} MC_i p_{i,t,n,t_n} + \sum_{t \in \mathcal{T}} \eta_{t,n} P_{\infty}^{t,n,t_n} \\
& \text{subject to} \quad (P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}) \in \mathcal{F}_D^{n,t_n}
\end{aligned}$$

Approximation of the Dual Function

The Lagrangian relaxation problem seen in the previous section occurred in a certain context:

1. The dual function of the problem is relatively easy to evaluate.
2. Complicating constraints make the problem difficult to solve.

This context, as seen previously, gives a Lagrangian dual function which is always concave regardless of the primal problem. Therefore, a lot of methods are based on the maximization of this lower bound in order to get an objective value close to the one of the original problem. We observe that this dual function may be non-differentiable. That's why, there are methods such as the Level method for which the idea is to maximize this lower bound by relying on linear approximations of this dual function ([47], [14]).

These linear approximations rely on the concept of subgradients. The subgradient generalizes the gradient notion for non-differentiable functions.

Definition 4.1 (Subgradient [14]). *Let f be a convex function, a vector π is called the subgradient of f at point $x_0 \in \text{Dom} f$, if for any $x \in \text{Dom} f$ we have:*

$$f(x) \geq f(x_0) + \langle \pi, x - x_0 \rangle$$

The set of all subgradients of f at x_0 , $\partial f(x_0)$ is called the subdifferential of function f at point x_0 .

The subgradient can be interpreted geometrically as a vector that defines a linear under-estimator of a function f . Similarly, if $f(x) \leq f(x_0) + \langle \pi, x - x_0 \rangle$ for all x then π is a supergradient [48].

In order to manage to construct a linear approximation of this dual function, one property is essential. This property will be proved for our global optimization program.

Proposition 4.1 (Dual Function Property [47]). *Let $(p_{g,t}^*, f_{k,t}^*, \theta_{n,t}^*, u_{g,t}^*, su_{g,t}^*, sd_{g,t}^*, D_\infty^{t,n*}, P_\infty^{t,n,t_n*}, P_{i,t,n,t_n}^*, Q_{i,t,n,t_n}^*, v_{i,t,n,t_n}^*, l_{i,t,n,t_n}^*, p_{i,t,n,t_n}^*, q_{i,t,n,t_n}^*)$ be optimal primal values for the dual function evaluated at $\eta_{t,n}$, then $\pi_t = (\sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n*} - D_\infty^{t,n*})$ is a supergradient of the dual function at $\eta_{t,n}$.*

Proof.

$$\begin{aligned}
g(\tilde{\eta}_{t,n}) &= \min_{\substack{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, \\ f_{k,t}, \theta_{n,t}, P_{i,t,n,t_n}, \\ Q_{i,t,n,t_n}, v_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, \\ p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ P_\infty^{t,n,t_n}, D_\infty^{t,n}}} \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n} + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} + \sum_{t \in \mathcal{T}, n \in \mathcal{N}} \tilde{\eta}_{t,n} \left(\sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n} - D_\infty^{t,n} \right) \\
&\leq \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n}^* + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t}^* + \sum_{t \in \mathcal{T}, n \in \mathcal{N}} \tilde{\eta}_{t,n} \left(\sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n*} - D_\infty^{t,n*} \right) \\
&= \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n}^* + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t}^* + \sum_{t \in \mathcal{T}, n \in \mathcal{N}} (\tilde{\eta}_{t,n} + \eta_{t,n} - \eta_{t,n}) \left(\sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n*} - D_\infty^{t,n*} \right) \\
&= g(\eta_{t,n}) + \sum_{t \in \mathcal{T}, n \in \mathcal{N}} (\tilde{\eta}_{t,n} - \eta_{t,n}) \left(\sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n*} - D_\infty^{t,n*} \right)
\end{aligned}$$

This sequence of inequalities is valid for all $\tilde{\eta}_{t,n}$ different from $\eta_{t,n}$. By using the definition of the supergradient, this ends the proof. $\square$

Consequently, the dual function can be approximated by a growing numbers of hyperplanes as this is given in the definition of an optimization problem's model (or approximation) [14].

Definition 4.2 (Model of a Concave Function). *Let $\Gamma = \{\gamma\}_{k=0}^{\infty}$ be a sequence in Q . Denote*

$$\hat{g}_k(\Gamma; \gamma) = \min_{0 \leq i \leq k} g(\gamma_i) + \langle \pi(\gamma_i), (\gamma - \gamma_i) \rangle \quad (4.2)$$

where $\pi(\gamma_k)$ are some supergradients of g at γ_i . The function $\hat{g}_k(\Gamma; \gamma)$ is called the model of the concave function.

From this definition, we can observe that $\hat{g}_k(\Gamma; \gamma)$ is a piecewise linear function of γ . Moreover, the next model is always better than the previous one: $\hat{g}_{k+1}(\Gamma; \gamma) \geq \hat{g}_k(\Gamma; \gamma)$.

4.1.2 Algorithm

The Level method is an algorithm used in order to minimize (or maximize) a non-differentiable function by using a linear approximation of this function. This algorithm is presented extensively in the work of Nesterov [14]. This Level method is applied in order to compute the closest lower bound on the objective of our original problem. In order to use this algorithm, two values must be recorded at each iteration of the method:

$$g_k^* = \max_{0 \leq i \leq k} g(\gamma_i), \quad \hat{g}_k^* = \max_{\gamma \in D} \hat{g}_k(\Gamma; \gamma) \quad (4.3)$$

The first problem is the maximal dual function value found so far and the second problem is the recorded value of the model. Undoubtedly, $g_k^* \leq g^* \leq \hat{g}_k^*$. Let $\alpha \in]0, 1[$ and define:

$$l_k(\alpha) = (1 - \alpha)g_k^* + \alpha\hat{g}_k^* \quad (4.4)$$

Consider the sublevel set:

$$\mathcal{L}_k(\alpha) = \{\gamma \in \mathcal{D} \mid g_k(\gamma) \geq l_k(\alpha)\} \quad (4.5)$$

This set has some interest for the optimization scheme; the set is stable with respect of small variations of the data. Moreover, projections on this sublevel sets are computationally cheap [49]. The Level method requires the use of these sublevel sets; the minimization scheme is therefore given in algorithm 1 where k is the number of the current iteration, k^{max} the maximum number of iterations allowed, LB the lower bound and UB the upper bound.

Algorithm 1: Level Method

Input: α, tol, k^{max} **Output:** γ_k and LB

```
1  $k := 0$ 
2  $\gamma_0 := 0$ 
3 while  $\frac{UB-LB}{UB} < tol$  and  $k < k^{max}$  do
4   Compute  $g_k \leftarrow g(\gamma_k)$ 
5   if  $g_k \geq g_{k-1}$  then
6     | Let  $g_k^* \leftarrow g_k$ 
7   end
8   else
9     | Let  $g_k^* \leftarrow g_{k-1}$ 
10  end
11  Let  $LB \leftarrow g_k^*$ 
12  Compute  $\hat{g}_k^*$ 
13  Let  $UB \leftarrow \hat{g}_k^*$ 
14  Compute  $l_k(\alpha) \leftarrow (1 - \alpha)UB + \alpha LB$ 
15  Compute  $\gamma_{k+1} \leftarrow \text{ProjectionLevelSet}(\gamma_k, l_k(\alpha))$ 
16   $k \leftarrow k + 1$ 
17 end
18 return  $\gamma_{k-1}$ 
```

The last operation in the loop of the algorithm given by the **ProjectionLevelSet** function projects the current iterate on the current sublevel set. The optimization program used to achieve this goal is given by:

$$\begin{aligned} \min_{\gamma} \quad & ||\gamma - \gamma_k||_2^2 \\ \text{subject to} \quad & g(\gamma_i) + \langle \pi(\gamma_i), (\gamma - \gamma_i) \rangle \geq l_k(\alpha) \quad \forall i = 0, \dots, k \end{aligned}$$

The convergence of this algorithm is stated and proven in [14].

Specific comments can now be made on how to apply this algorithm to our global optimization problem. First, the dual function computed at line 4 of algorithm 1 will correspond to the computation of the dual function presented in section 4.1.1. Since this dual function can be decoupled into subproblems (one for the transmission network and another one for each distribution network), then the computation of this function will be splitted accordingly in order to speed up the algorithm. Then, the model's value computed at line 12 is given using the definition 4.2 on our problem by considering supergradients presented in proposition 4.1. The projection problems also used these supergradients.

4.1.3 Stopping Criterion

Rather than using the stopping criterion proposed by the algorithm, in order to have a standardized criterion for all the proposed algorithms, we will use a criterion called the gap. The gap measures the difference between the best lower bound found on the original problem and the objective linked with the best feasible solution found so far. A lower bound on the original optimization program is already given in the algorithm thanks to the dual function, see the details of algorithm 1. However, in order to obtain a feasible solution at each iteration, a recovery

problem has to be implemented that allows to retrieve a feasible solution from the relaxed solution obtained from the Lagrangian dual function.

Recovery Program

The recovery optimization program will be applied after each iteration on the transmission part of the network in order to retrieve a feasible solution from the relaxed solution found by the Lagrangian function. This recovery problem uses the current values computed for P_∞^{t,n,t_n} variables by the Lagrangian dual function. $D_\infty^{t,n}$ variables are fixed to these values in the following way:

$$D_\infty^{t,n} = \sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n} \quad \forall t \in \mathcal{T}, n \in \mathcal{N} : \mathcal{T}_n \neq \emptyset \quad (4.6)$$

After that, the optimization model devoted to the transmission of the network is solved with $D_\infty^{t,n}$ variables fixed. Using this technique, a feasible solution is recovered from the relaxed solution obtained by the Lagrangian dual function. If the transmission problem is infeasible for this particular value of P_∞^{t,n,t_n} then the cost of the feasible solution is put to $+\infty$. The recovery problem is given by the following set of equations.

$$\begin{aligned} & \min_{p_{g,t}, f_{k,t}, \theta_{n,t}, u_{g,t}, su_{g,t}, sd_{g,t}} \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} \\ & \text{subject to} \quad (p_{g,t}, f_{k,t}, \theta_{n,t}, u_{g,t}, su_{g,t}, sd_{g,t}, D_\infty^{t,n}) \in \mathcal{F}_T \end{aligned}$$

4.2 Alternating Direction Method of Multipliers

The alternating direction method of multipliers (ADMM) is a powerful algorithm well suited to distributed convex optimization for separable problems [46]. It was first invented by Gabay and Mercier [50] along with Glowinski and Marrocco [51] in the middle of the seventies even though some similar ideas appeared around the fifties. Since the ADMM method was already proposed by Peng et al. in [34] in order to compute the optimal value of the **Optimal Power Flow** problem relaxed to a SOCP program, this algorithm was added in this thesis to observe its benefits. The ADMM method will be described in the following section, for a more detailed exposition about this algorithm, see [49], [34] and a review on the subject [46].

First, the augmented Lagrangian concept will be presented since it is a core concept of this algorithm. After that, the algorithm will be explained. Then, some convergence results for the algorithm will be stated. Finally, a stopping criterion for this algorithm will be discussed.

4.2.1 Augmented Lagrangian

Augmented Lagrangian methods expanded in order to add robustness to dual methods based on Lagrangian relaxation [46]. This concept is presented in [44] and [46].

Consider an optimization problem of the type:

$$\begin{aligned} & \min_{x \in X} f(x) \\ & \text{subject to} \quad Ax = b \end{aligned}$$

The augmented Lagrangian function of this problem is given by relaxing the constraint:

$$L_\rho(x, \mu) := f(x) + \mu^T(Ax - b) + \frac{\rho}{2} \|Ax - b\|_2^2$$

where $\rho > 0$ is called the **Penalty Parameter**. The augmented Lagrangian function can be seen as the (unaugmented) Lagrangian function of the optimization problem [46]:

$$\begin{aligned} \min_{x \in X} \quad & f(x) + \frac{\rho}{2} \|Ax - b\|_2^2 \\ \text{subject to} \quad & Ax = b \end{aligned}$$

A corresponding augmented dual function can be defined according to this augmented Lagrangian function:

$$g_\rho(\mu) := \min_{x \in X} L_\rho(x, \mu)$$

4.2.2 Algorithm

The alternating direction method of multipliers is a procedure in which the solutions to small local subproblems are coordinated to find a solution to a larger global problem. Since this algorithm needs the augmented Lagrangian function, it can be viewed as an attempt to use the benefits of dual decomposition algorithms and the superior convergence properties of augmented Lagrangian methods for constrained problems.

These algorithms solve problems that are separable as explained in section 4.1. The optimization problem addressed by this algorithm has the following form:

$$\begin{aligned} \min_{x, z} \quad & f(x) + \phi(z) \\ \text{subject to} \quad & (\mu)Ax + Bz = c \\ & x \in \mathcal{X} \\ & z \in \mathcal{Z} \end{aligned}$$

The augmented Lagrangian of the following problem is given by:

$$L_\rho(x, z, \rho) = f(x) + \phi(z) + \mu^T(Ax + Bz - c) + \frac{\rho}{2} \|Ax + Bz - c\|_2^2$$

The ADMM algorithm is presented in Algorithm 2. This method will decoupled the problem by first solving the optimization program with only one set of variables considered. Then, the optimization problem will be solved another time by minimizing over the other set of variables. The solution to these two optimization problems will be used later to update the dual variable linked to the difficult constraint.

Algorithm 2: ADMM Method

Input: ρ, k^{max}, tol

Output: $x_k, z_k, \mu_k, f(x_k) + \phi(z_k)$

1 $k := 0$

2 $\mu_0 := 0$

3 **while** $\|Ax_{k+1} + Bz_{k+1} - c\|_1 > tol$ **and** $k < k^{max}$ **do**

4 Compute $x_{k+1} \leftarrow \underset{x \in X}{\operatorname{argmin}} L_\rho(x, z_k, \mu_k)$ (x-minimization step)

5 Compute $z_{k+1} \leftarrow \underset{z \in Z}{\operatorname{argmin}} L_\rho(x_{k+1}, z, \mu_k)$ (z-minimization step)

6 Compute $\mu_{k+1} \leftarrow \mu_k + \rho(Ax_{k+1} + Bz_{k+1} - c)$ (dual update step)

7 **end**

8 **return** $x_{k+1}, z_{k+1}, \mu_{k+1}, f(x_{k+1}) + \phi(z_{k+1})$

We can observe from the iterations of algorithm 2 that using the ADMM method allows us to separate two problems where variables x and z are updated in an alternating way. As for the Level method, the relaxed constraint in our case will be the linking constraint in order to obtain separable objectives and problems. Therefore, the following augmented dual function of our global optimization problem can be decoupled in several dual functions as for the Level method: one for the transmission network, the others for each distribution networks attached to the transmission grid.

$$\begin{aligned}
& \min_{\substack{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, \\ f_{k,t}, \theta_{n,t}, P_{i,t,n,t_n}, \\ Q_{i,t,n,t_n}, v_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, \\ p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ P_{\infty}^{t,n,t_n}, D_{\infty}^{t,n}}} \sum_{\substack{i \in \mathcal{I}_0, t \in \mathcal{T}, \\ n \in \mathcal{N}, t_n \in \mathcal{T}_n}} MC_i p_{i,t,n,t_n} + \sum_{g \in \mathcal{G}, t \in \mathcal{T}} MC_g p_{g,t} + \sum_{t \in \mathcal{T}, n \in \mathcal{N}} \eta_{t,n} \left(\sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} \right) \\
& \quad + \frac{\rho}{2} \sum_{t \in \mathcal{T}, n \in \mathcal{N}} \left\| \sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} \right\|_2^2 \\
& \text{subject to} \quad (su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, f_{k,t}, \theta_{n,t}, D_{\infty}^{t,n}) \in \mathcal{F}_T \\
& \quad (P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_{\infty}^{t,n,t_n}) \in \mathcal{F}_D^{n,t_n}
\end{aligned}$$

4.2.3 Convergence

As mentioned earlier, augmented Lagrangian methods lead to superior convergence results than classical methods using (unaugmented) Lagrangian algorithms. Two assumptions about the problem are presented in order to state these convergence results.

Assumption 4.1. *The (extended-real-valued) functions $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ and $\phi : \mathbb{R}^m \rightarrow \mathbb{R} \cup \{+\infty\}$ are closed, proper and convex.*

The first assumption is equivalent to the assumption that the epigraph of functions f and ϕ :

$$\text{epi}_f = \{(x, t) \in \mathbb{R}^n \times \mathbb{R} \mid f(x) \leq t\}, \text{epi}_\phi = \{(z, t) \in \mathbb{R}^m \times \mathbb{R} \mid \phi(z) \leq t\}$$

are closed nonempty convex sets. This assumption implies that subproblems arising in the x-minimization and z-minimization problems contains at least a solution that is not necessarily unique.

Assumption 4.2. *The unaugmented Lagrangian (L_0) has a saddle point. Explicitly, there exists (x^*, z^*, μ^*) , not necessarily unique, for which $L_0(x^*, z^*, \mu) \leq L_0(x^*, z^*, \mu^*) \leq L_0(x, z, \mu^*)$ holds for all (x, z, μ) .*

Using these two assumptions 4.1 and 4.2, a theorem allows to state the convergence of the algorithm.

Theorem 4.1 (ADMM Method Convergence [46]). *Under assumptions 4.1 and 4.2, the ADMM iterates satisfy the following:*

- *Residual convergence:* $(Ax_k + Bz_k - c)$ tends to 0 as k tends to infinity, i.e. the iterates approach feasibility.
- *Objective convergence:* $f(x_k) + \phi(z_k) \rightarrow p^*$ as $k \rightarrow +\infty$, i.e. the objective function of the iterates approaches the optimal value.

- *Dual variable convergence: μ_k tends to μ^* as k tends to infinity, where μ^* is a dual optimal point.*

Remark. If the problem is nonconvex then the algorithm may or may not converge and if it converges, it can converge to a local optimal point. There are no results concerning the convergence of this algorithm for nonconvex problems.

4.2.4 Stopping Criterion

The stopping criterion presented in algorithm 2 is a reasonable termination criterion regarding Theorem 4.1. However, in order to have a standardized criterion to stop each algorithms of this chapter, the gap criterion is also going to be used for this method as for the Level algorithm. The process used to obtained a lower bound and a feasible solution at each iteration of the problem will be described in this section.

Lower Bound

Unlike the Level algorithm, the ADMM algorithm do not produce a lower bound on the problem at each iteration. Indeed, the augmented Lagrangian dual function is not especially a lower bound on the original problem. However, the unaugmented Lagrangian dual function constructs a lower bound on the problem. Therefore, the Lagrangian dual function computed in each iteration of the Level method will also be computed in each iteration of the ADMM algorithm. The Lagrangian dual function was presented in section 4.1.1.

Best Feasible Solution

In order to compute a feasible solution from the relaxed solution computed at each iteration of the algorithm, the process applied for the Level algorithm will be used. Therefore, the feasible solution will be retrieved from the relaxed solution by using the same recovery program than the Level method, see section 4.1 for the explicit of the recovery program.

4.3 Caramanis Method

The third presented algorithm was inspired by the decomposition found in the publication of Caramanis et al. in 2016 [8]. As explained in the state of the art, this paper uses an iterative approach that decoupled the problem into node's variables and considered the power entering the node as a generator. This algorithm focuses on each participating agents which makes converging iterative decisions about dual and primal variables in order to agree on the way to operate the system.

4.3.1 Algorithm

For this distributed algorithm, two types of subproblems are needed. First, a subproblem containing only the transmission network model is used by fixing the linking variable $D_\infty^{t,n}$ to some value. Another type of subproblem is given by the optimization program describing the distribution network where a term is added on the objective containing the price of a virtual generator multiplied by the linking variable P_∞^{t,n,t_n} . The two types of subproblem are linked with the signification of introduced variables P_∞^{t,n,t_n} and $D_\infty^{t,n}$ which was illustrated in Figure 2.8

in section 2.4.

The execution of the distributed algorithm is therefore described by the following two steps. Current iteration is numbered k .

- Phase 1: The transmission subproblem is solved. Then the binary variables of the transmission subproblem are fixed and the program is solved again in order to obtain the dual variables corresponding to the linking constraints. These dual variables are registered as the prices for the electricity generated by virtual generators present at the root of each distribution network.
- Phase 2: All distribution subproblems are solved by using the registered dual variable as the price of the electricity generated by the virtual generator. The power stored in the linking variable P_{∞}^{t,n,t_n} is registered and passed to the transmission subproblem by fixing variables $D_{\infty}^{t,n}$ according to these values in the following way:

$$\sum_{t_n \in \mathcal{T}_n} P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n} = 0$$

This corresponds to the demand requested by the transmission network which will be provided to the distribution networks. Then, the algorithm goes back to phase 1.

Therefore, the Caramanis algorithm is described by the mathematical steps presented in Algorithm 3.

Algorithm 3: Caramanis Method

Input: k^{max}, tol
Output: $P_{\infty,k+1}^{t,n,t_n}$

```

1  $k := 0$ 
2  $Price_0^{t,n} := 0$ 
3  $P_{\infty,0}^{t,n,t_n} := 0$ 
4 while  $\| \frac{P_{\infty,k+1}^{t,n,t_n} - P_{\infty,k}^{t,n,t_n}}{P_{\infty,k}^{t,n,t_n}} \|_1 < tol$  and  $k < k^{max}$  do
5   Fix  $D_{\infty}^{t,n} = \sum_{t_n \in \mathcal{T}_n} P_{\infty,k}^{t,n,t_n}$ 
6   Solve Transmission Subproblem
7   Fix  $(u_{g,t}, su_{g,t}, sd_{g,t})$ 
8   Solve Transmission Subproblem
9   Let  $Price_{k+1}^{t,n} \leftarrow \eta_{t,n}$ 
10  for  $n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset$  do
11    Solve Distribution Subproblem  $(n, t_n)$ 
12    Let  $P_{\infty,k+1}^{t,n,t_n} \leftarrow P_{\infty}^{t,n,t_n}$ 
13  end
14  Let  $k := k + 1$ 
15 end
```

Remark. Note that the Caramanis algorithm has no convergence results known.

4.3.2 Stopping Criterion

In order to have a stopping criterion standardized compared to the other methods, this one will be based on the gap. The process used to obtain a lower bound on the original problem and a feasible solution is explained in this section.

Lower Bound

The prices updated at each iteration of the algorithm can be seen as the dual variables corresponding to the linking constraint. Therefore, the Lagrangian dual function of the problem can be computed at each iteration using the current prices in order to obtain a lower bound on the original problem. The Lagrangian dual function just needs to be computed for the transmission network since the distribution part of the Lagrangian dual function is already computed during the second phase of the algorithm.

Best feasible Solution

Recovery subproblems can be implemented in order to retrieve a feasible solution at each iteration of the algorithm. A feasible solution for the transmission network is already provided at the first phase of the algorithm. Therefore, only recovery problems for each distribution network has to be computed based on the values obtained for $D_\infty^{t,n}$ during the first phase of the algorithm. The recovery problem for the set of distribution networks attached at transmission node n is given by the following optimization problem.

$$\begin{aligned}
& \min_{\substack{P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ P_\infty^{t,n,t_n}, D_\infty^{t,n}}} \sum_{i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n} MC_i p_{i,t,n,t_n} \\
& \text{subject to} \quad \sum_{t_n \in \mathcal{T}_n} P_\infty^{t,n,t_n} - D_\infty^{t,n} = 0 \\
& \quad (P_{i,t,n,t_n}, Q_{i,t,n,t_n}, v_{i,t,n,t_n}, l_{i,t,n,t_n}, p_{i,t,n,t_n}, q_{i,t,n,t_n}, P_\infty^{t,n,t_n}) \in \mathcal{F}_D^{n,t_n}
\end{aligned}$$

Chapter 5

Results

This chapter is devoted to the numerical results obtained by experiments made on several aspects which are presented in this thesis. In order to provide these numerical results, several sets of data representing power systems were built. These data sets were used in order to tune some parameters of the algorithms introduced in chapter 4 but also to compare performances of the algorithms. First, the test cases used will be introduced. After that, some computational results related to the algorithms will be presented.

5.1 Description Test Cases

Data sets were created in order to observe the behaviour of the algorithms presented in chapter 4. Three data sets were used to realize the experiments. The first two data sets represent two different power systems used to reproduce transmission networks. The last data set contains the description of the distribution network.

5.1.1 Transmission Data

The data sets used in order to describe two different transmission networks are presented in this section. Both data sets were found in Castillo's publication [17].

6 Bus Unit Commitment

The **6 bus** power system is a network that contains 6 nodes linked by 7 lines. It also comprises three generators located at different nodes and three other buses only withdrawing power from the network. The data provided by [17] were first used by Li et al. in [52], more details about this power system test case can be found in these two publications. The test case takes also into account information about unit commitment features such as minimum down time, minimum up time, ramp rates limits, etc.. The maximum power that can be withdrawn from the system is 340 [MW].

24 Bus Unit Commitment

The **24 bus** power system is the IEEE RTS-96 Test system [53] which is an improved version of the IEEE RTS-79 Test case published in [54]. The enhancement of the first set of data was made by adding unit commitment information to the power system in 1996. More details about this data set can be found in [53] and [17]. This power network contains 24 buses and 38 lines. Among all the network's buses, only 17 consumed electricity. There are 32 generators, with

maximum capacity ranging from 12 to 400[MW], located in the power grid. Therefore, the maximum total power that can be generated by the system is 3405[MW].

5.1.2 Distribution

The power system used in the distribution network of the example presented in section 2.5 is also employed in this chapter. However, since the consumption of real and reactive power is given for only one time period, some real and reactive varying load data were constructed in order to cover an operation horizon of 24hours with fluctuating consumption over the day. To build the consumption data for real and reactive power, we have used the curves represented at figure 5.1 given in [55] that illustrates the total real and reactive consumption profile of their system from 6:30am to 10:45pm.

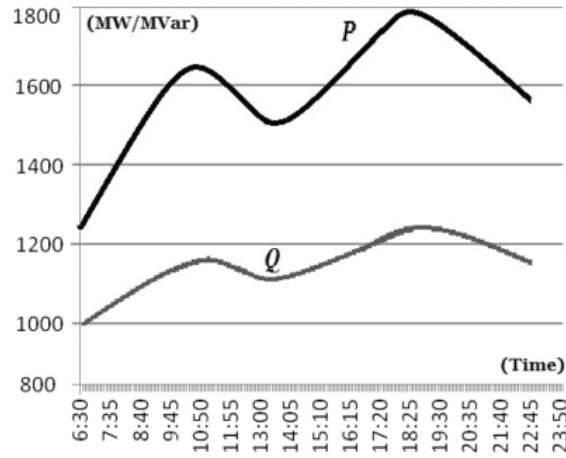


Figure 5.1: Daily load variation characteristic of the test system

The process used in order to construct the consumption data is extensively presented in [55]. However, before applying the technique of the previous publication to construct our data set, the curves were scaled in order to cover an entire day from 1am to midnight. The scaling process do not modify the consistency of the curves since the peak of power consumption is still situated around 6:30pm. The following steps describe the mechanism used in order to generate the set of consumption data:

1. First, an hour was selected as the initial point in the consumption profile. Then, the ratios between power withdrawn at this initial hour and power consumed at other hours were computed. The chosen hour for initial value was 6pm.
2. The total consumption given in the distribution data was then assumed to be the power consumption at the initial time chosen at previous step. After that, this value was scaled in order to obtain the total consumption for other hours of the day by using the ratios computed at step 1.
3. After that, the total load was allocated for each hour among each node according to the ratios between each bus' load to the system's total load.

This technique provides us consumption of real and reactive powers for each hour of the day for the distribution network.

5.1.3 General Remarks

Since this thesis is based on an optimization program considering transmission and distribution networks together, some features linked with the combination of both types of networks have to be discussed before introducing the results of our experiments.

First, the same set of data network is going to be used for each distribution networks of the entire power grid. Therefore, if three distribution networks are attached to one transmission node of the system, then it will be three times the same distribution network that will be connected to this transmission node.

Moreover, the power withdrawn directly from the transmission network is ignored. This allows us to put a maximum of distribution networks per transmission nodes. Furthermore, without loss of generality, the distribution networks can only be attached to nodes of the transmission network from which power is consumed. Therefore, the **6 bus** network can accept up to 26 distribution networks per demand node at most and the **24 bus** network can only accept 20 distribution networks per load node.

Finally, physical flow limits on the transmission lines are ignored in order to be able to add the maximum number of distribution networks to each transmission node. Indeed, the goal of this thesis is to observe the behaviour of distributed algorithms for networks of growing size. Consequently, the focus is on the number of distribution networks attached to the transmission network and not on physical flow limits on lines of the transmission network.

5.2 Computational Results

This section is devoted to the computational results obtained on the introduced data sets for the different algorithms presented in chapter 4. First, the tuning of several parameters belonging to the distributed algorithms are performed. Then, the efficiency of the distributed algorithms is compared to the computational effort required by commercial solvers to address the global optimization problem. All these tests were realized on personal computer 2.9GHz Intel Core i5 with 8GB RAM. All the codes were created in AMPL®, the version of the program is 20170207. The commercial solver used to generate is version 12.7 of CPLEX®.

5.2.1 Tuning Parameters

In this section, the tuning of several algorithm's parameters will be explained. The tests made in order to tune algorithm's parameters were only conducted for the **6 bus** transmission dataset.

Penalty Parameter on Current Tuning

The quadratic constraint of the OPF problem has been relaxed in order to give a conic constraint, for details see section 2.2. However, for the optimal solution of the global optimization program to have a physical meaning, this constraint is required to be tighten. Therefore, in order to help the solver to minimize the difference between the two sides of the conic inequality constraint, we added in the objective a term that minimizes the amount of current flowing on distribution lines multiplied by a certain coefficient. The added penalty term on the objective is:

$$Penalty = \epsilon \sum_{\substack{i \in \mathcal{I}, t \in \mathcal{T}, \\ n \in \mathcal{N}, t_n \in \mathcal{T}_n: \bar{\mathcal{T}}_n \neq \emptyset}} R_i l_{i,t,n,t_n} \quad (5.1)$$

This penalty has actually a physical meaning because it minimizes the real power losses along each distribution lines.

In order to select the best value for parameter ϵ introduced in the formulation of the objective penalty's equation (5.1), several values for this parameter were tested for a given number of distribution networks attached per node of the transmission grid. The maximum differences between each part of the conic constraint called the slack is given on figures 5.2 . The tests were made for both formulations of the linking constraint, values of the parameter in the set $\{0, 0.0001, 0.001, 0.01, 0.1, 1, 10, 100\}$ and number of distribution networks per node in the set of values $\{1, 5, 10, 15, 20, 25\}$.

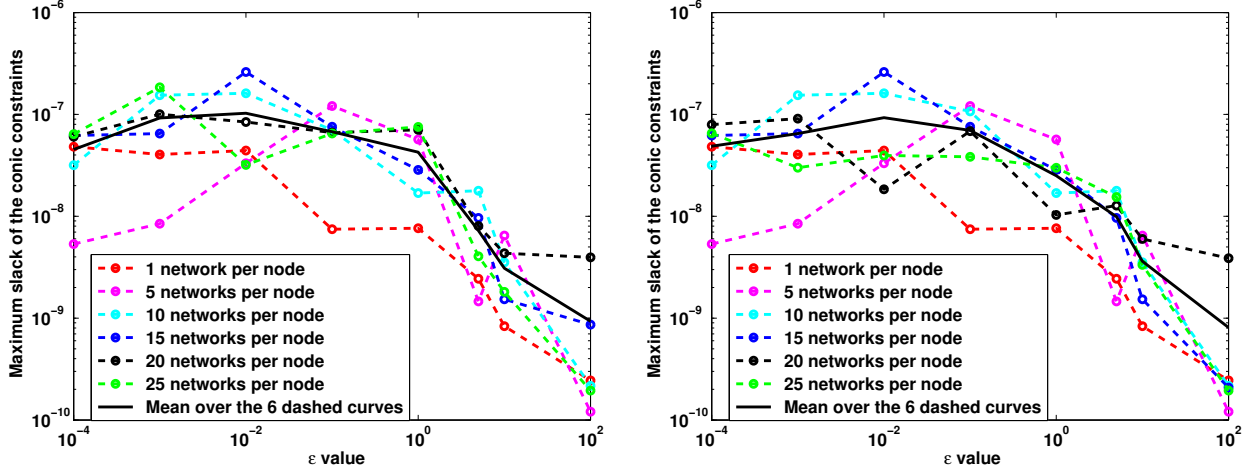


Figure 5.2: Evolution of the maximum slack of the conic constraint for some values of the penalty parameter ϵ . The tests were made for several number of distribution networks per node. The left figure presents results obtained for the first formulation of the linking constraint, while the right figure considers the second formulation of this constraint.

In order to choose the best value for the searched parameter, a trade-off is necessary. Indeed, the best parameter will be the one that reduces the most the maximum value of the slack of the conic constraint. However, the values of the primal variables of the problem along with the objective must not be modified within a high percentage from their original values. Indeed, if the optimal values of the primal variables are really modified by using one value of the parameter, then it means that the optimization problem focused more on the minimization of the real power losses instead of the wanted minimization of the production cost. This trade-off prevents us from taking the value 100 for the penalty parameter. In fact, even if it seems to be the most efficient parameter to reduce the value of the maximum slack of the conic constraint as seen in figure 5.2, this value distorts the true value of the objective function in a significant measure. The effect of the penalty parameter on the objective value of the global optimization program is shown in figures 5.3.

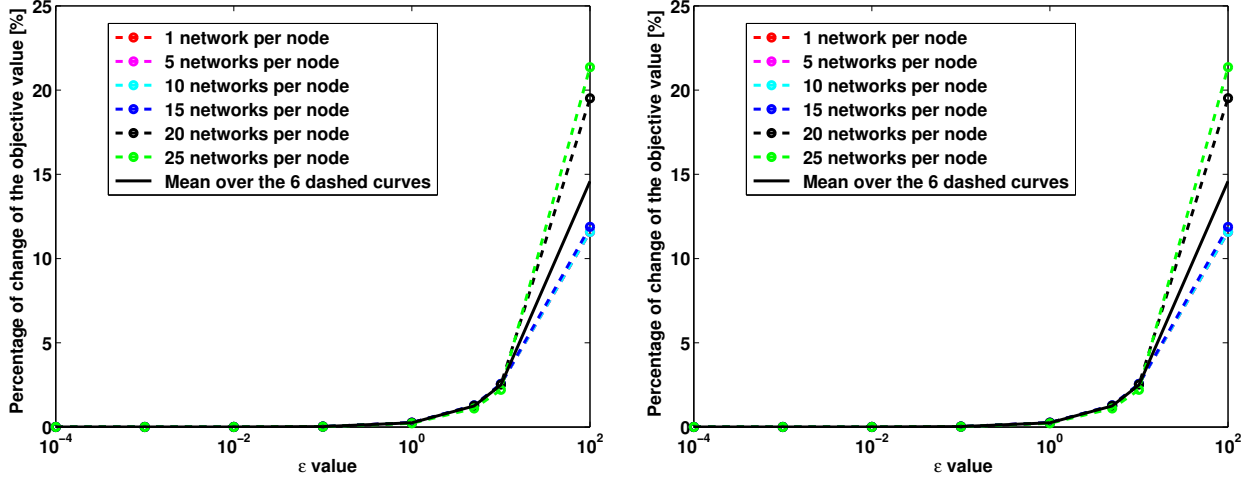


Figure 5.3: Evolution of the objective distortion in percent for some values of parameter ϵ . The tests were made for several numbers of distribution networks per node. The left figure present results obtained for the first formulation of the linking constraint, while the right figure considers the second formulation of this constraint.

Moreover, the evolution of the distortion in percent of primal variables $p_{g,t}$ is given in figures 5.4. The impact on some other primal variables are given in appendix B.

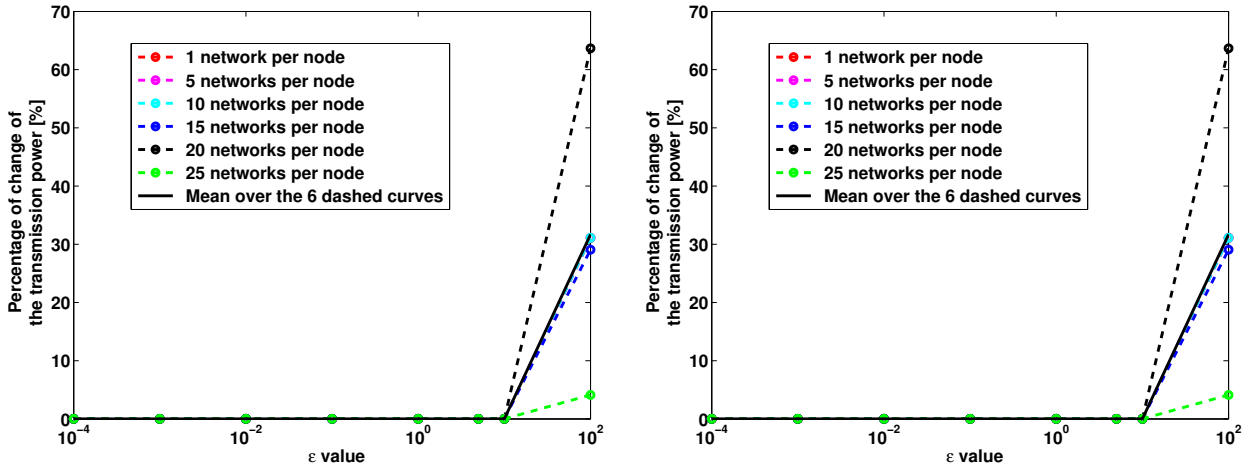


Figure 5.4: Evolution of the primal variable $p_{g,t}$ distortion in percent for some values of parameter ϵ . the tests were made for several numbers of distribution networks per node. The left figure present results obtained for the first formulation of the linking constraint, while the right figure considers the second formulation of this constraint.

Using figures 5.2, 5.3 and 5.4, we chose to use the value 5 for the penalty parameter ϵ is 5. Indeed, a distortion of the original optimal objective value by about one percent is acceptable. Moreover, we observe that the primal variable is not affected in a significant measure for this value of the parameter. Finally, the maximum value for the slack is reduced compared to the original slack.

Discussion Gap Tolerance

In order to stop the three algorithms, a tolerance value for the uniformed stopping criterion has to be set. However, this parameter is not going to be tuned since the literature about unit commitment problems has already stated commonly acceptable values and solving time requirements. In fact, concerning gap values, an algorithm used to solve a unit commitment program that computes the solution within a remaining gap superior to 5% will be considered as unacceptable from the point of view of power system operators. Since 5% is actually the minimum level of performance required for an algorithm, the value of the tolerance parameter will be arbitrarily fixed to 0.01% for the remaining part of the experiments driven on the algorithms. As an example, stochastic unit commitment models usually achieve a gap of 3% or less [3].

Furthermore, the algorithms, presented in chapter 4, will also be qualified on the time they require to solve a desired problem. However, day-ahead problems (e.g. problems involving unit commitment constraints) that will require more than a few hours to be solved (e.g. 4 hours) are judged as unsatisfactory from the point of view of power system operations firms. Therefore, the algorithms are required to solve unit commitment problems in less than a few hours.

Tuning Regularization parameter

The ADMM algorithm exposed in section 4.2 is composed of one regularization parameter. This parameter needs to be fixed to the most appropriate value for the remaining part of this document. Some experiments were made in order to determine the value of this parameter. The number of iterations required by the ADMM algorithm to converge along with the gap obtained after convergence were observed for several values of the regularization parameter belonging to $(\{0.01, 0.1, 1, 10, 20, 50, 100\})$ and for a certain number of distribution networks per node $(\{1, 5, 10, 15, 20, 25\})$. The tests were only computed for the first formulation of the global optimization program. The number of iterations needed by the algorithm is presented in figure 5.5. In figure 5.6, the time required by the algorithm to stop is shown. The obtained gap after convergence of the algorithm is presented in figure 5.7.

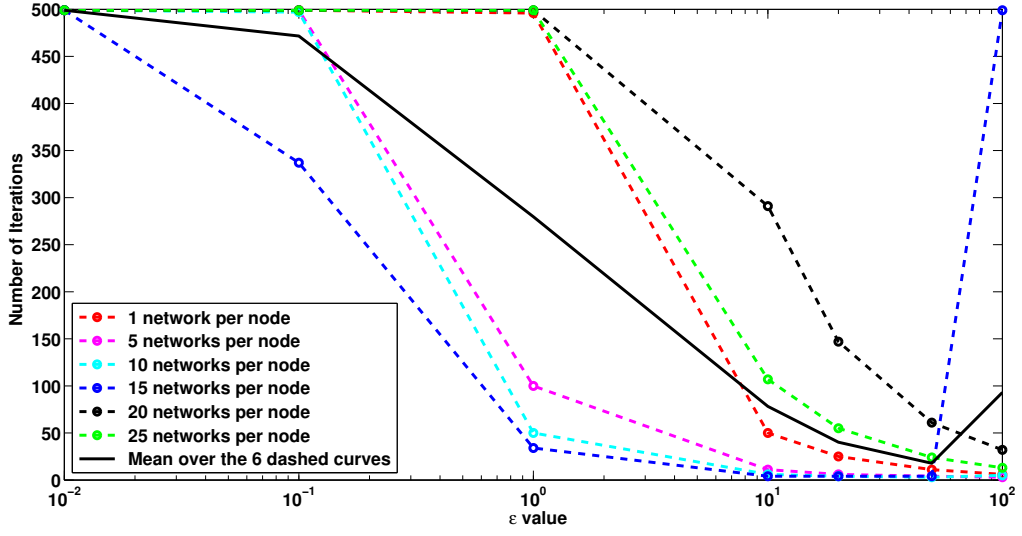


Figure 5.5: Illustration of the number of iterations needed by the ADMM method to converge for a certain number of distribution networks per node and several values of the regularization parameter

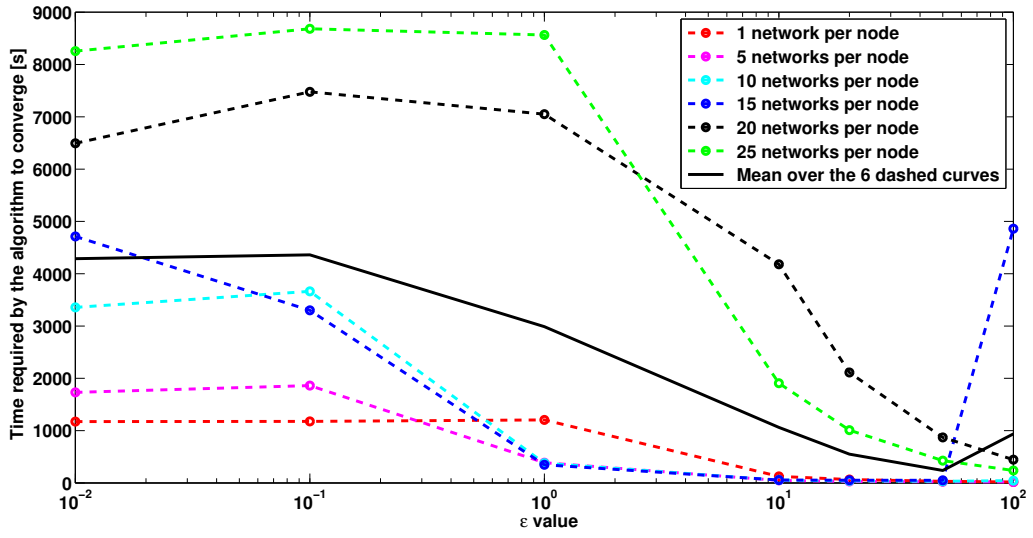


Figure 5.6: Illustration of the time needed by the ADMM method to converge for a certain number of distribution networks per node and several values of the regularization parameter

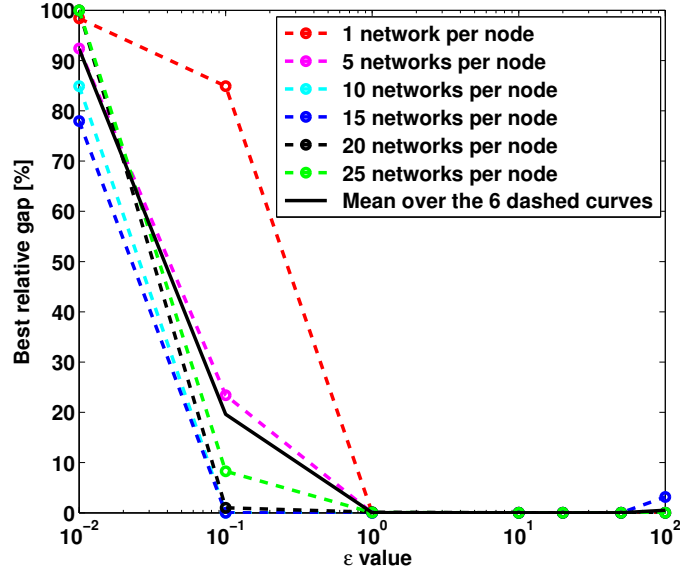


Figure 5.7: Illustration of the gap obtained after convergence of the ADMM method for a certain number of distribution networks per node and several values of the regularization parameter

By observing the results presented in figures 5.5, 5.6 and 5.7, the chosen value for the regularization parameter ρ will be given by 20. This decision results from a trade off. Indeed, by considering figure 5.5, the value 100 seems to be the more efficient in terms of number of iterations required by the algorithm to converge. However, for one case, this parameter value of 100 does not allow the ADMM algorithm to converge to a gap smaller than 0.01%. This is due to the fact that this value is too high because it will change the value of the dual variable too significantly at each iteration and the algorithm is therefore going to be stuck before reaching the tolerated gap. To solve this problem, there exists a trade off: a too high value for the regularization parameter may lead to an algorithm that do not manage to go beyond the required gap tolerance. However, a too small value for the regularization parameter prevents the algorithm to be competitive in terms of computational time regarding commercial solvers.

Tuning Parameter Level Method

Similarly, the Level Method, described in section 4.1, also requires the use of a parameter α . The value of this parameter is comprised between 0 and 1 not included. In fact, if this parameter is equal to 0 then the algorithm is not going to make any progress. However, if α is set to 1, then the algorithm is reduced to the classic cutting-plane (Kelley's) algorithm [49]. Therefore, a suitable intermediate value is desired for this parameter. Similar experiments were consequently conducted on this algorithm in order to determine a suitable value for this parameter. The number of iterations needed by the algorithm to converge along with the resulting gap after convergence of this algorithm were recorded for several values of the parameter ($\{0.2, 0.4, 0.6, 0.8\}$) and different numbers of distribution networks per node ($\{1, 5, 10, 15, 20, 25\}$). These tests were only conducted for the first formulation of the linking constraint. The number of iterations obtained and the time required by the algorithm to converge are given respectively in figure 5.8 and 5.9. The results concerning the gap obtained after convergence are presented in figure 5.10.

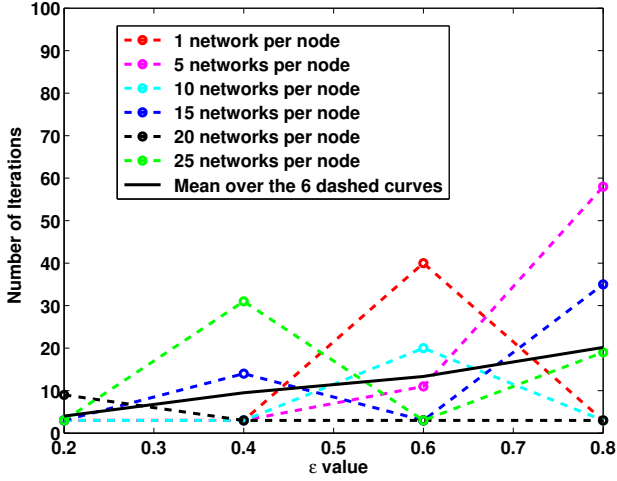


Figure 5.8: Illustration of the number of iterations needed by the Level method to converge for a certain number of distribution networks per node and several values of the algorithm's parameter

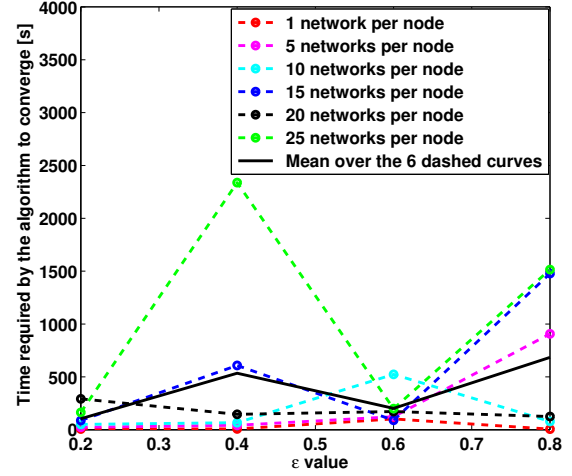


Figure 5.9: Illustration of the time that is required by the Level method to converge for a certain number of distribution networks per node and several values of the algorithm's parameter

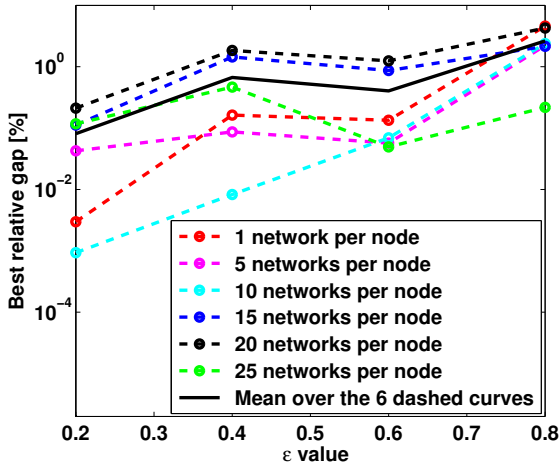


Figure 5.10: Illustration of the gap obtained after convergence by the Level method for a certain number of distribution networks per node and several values of the algorithm's parameter

After some tests, I observe that the commercial solver CPLEX® is very sensitive to conic programs involving small numbers. Indeed, due to the sensitivity of the interior point method of the solver to the scale of the problem, the solver may converge to a non optimal solution of the problem. However, this disturbs the convergence of algorithms and the Level method is the most affected by this sensitivity. Therefore, the tolerance of 0.01 % could not be reached by applying the Level method to tests cases. It remains that the residual gap after convergence of the algorithm is still smaller than the minimum required performance for unit commitment problems as this can be observed in figure 5.10.

Using results presented in figures 5.8, 5.9 and 5.10, a suitable value for the parameter α can be chosen. In fact, we can observe that the smallest number of iterations is achieved for parameter value 0.2. Moreover, the best feasible gaps obtained for this value of the parameter are also the lowest possible values. Therefore, the value 0.2 is chosen for the α parameter.

5.2.2 Efficiency Algorithms and Commercial Solver Results

In order to prove the efficiency of the algorithms presented in Chapter 4 compared to the optimization of the entire network as a whole, solving times needed by the algorithms and the commercial solver in order for the global problem to converge, are presented in the following section. Other results corresponding to the gap obtained after convergence are shown.

Total Problem

In order to observe the effect on the global optimization problem taking into account growing number of distribution networks attached at each transmission node, experiments were made to record the time needed by the commercial solver to compute the solution of the global optimization problem until a certain accuracy. The time needed by the commercial solver to converge for different numbers of distribution networks attached to the transmission network is represented in figure 5.11 for both datasets and both formulations of the linking constraint. The evolution of the objective function compared to the number of distribution networks per node is reproduced in figure 5.12.

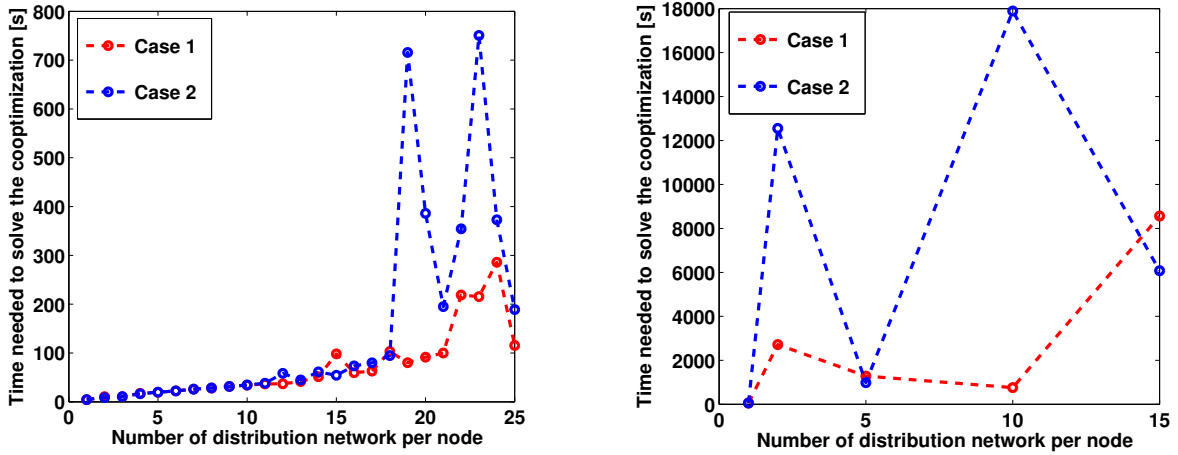


Figure 5.11: Graphic representation of the time required by the commercial solver to compute the solution of the global optimization program for several numbers of distribution networks per transmission node using the **6 bus** dataset (left) and the **24 bus** dataset (right)

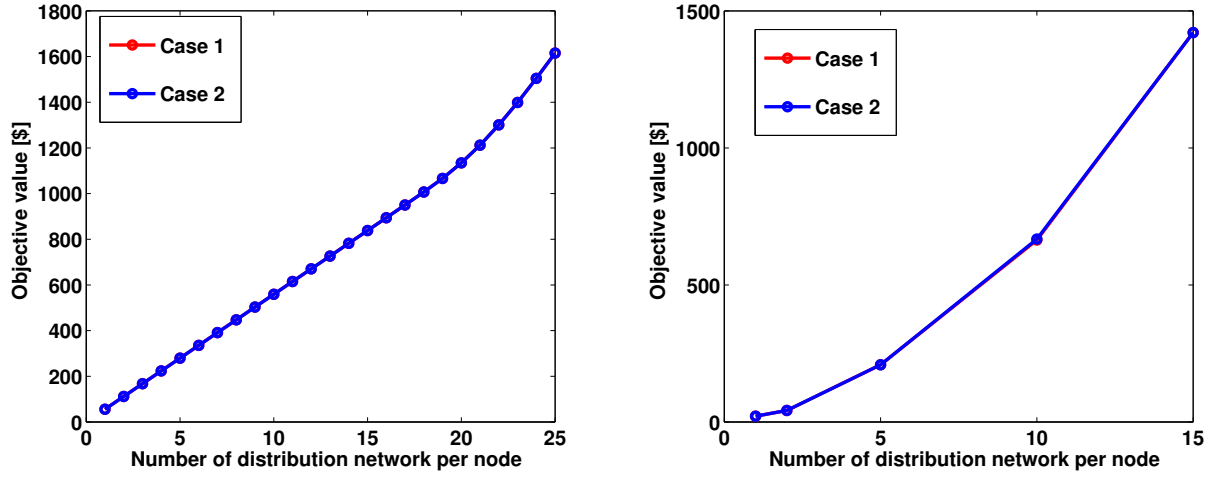


Figure 5.12: Graphic representation of the optimal objective value of the global optimization program for several numbers of distribution networks per transmission node using the **6 bus** dataset (left) and the **24 bus** dataset (right)

The observation of results presented in figures 5.11 suggests that solving times are affected by the formulation of the linking constraint considered. It seems that the first formulation is more efficient in terms of solving time. Graphical representations of relative gaps achieved by the commercial solver for this problem are respectively given in figures 5.13. It can be noted that all relative gaps achieved are under the one proposed as tolerance for the remaining part of this document.

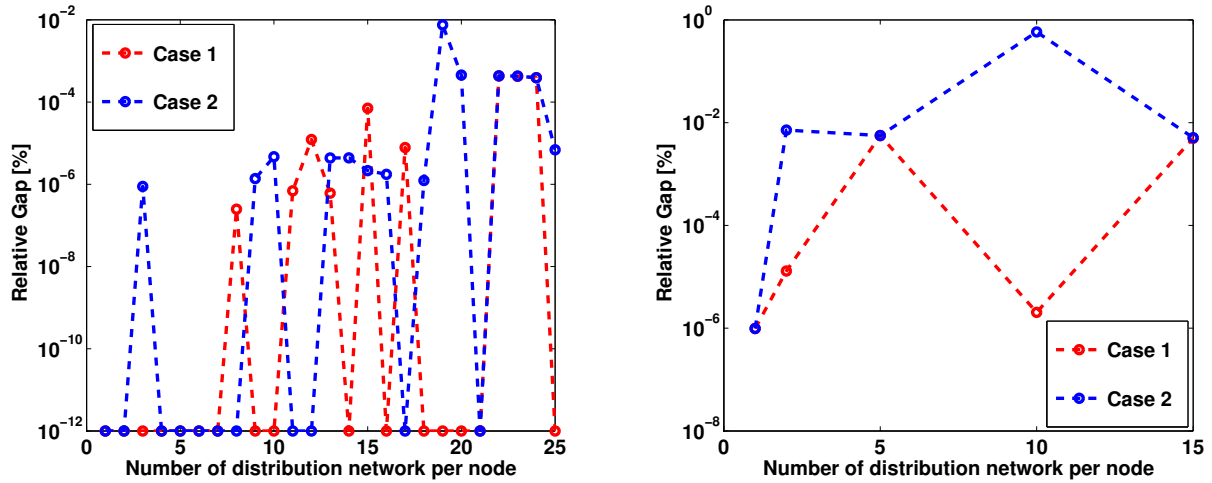


Figure 5.13: Graphical representation of the relative gap in percent achieved by the commercial solver while solving the global optimization problem for different numbers of distribution networks per transmission node using the **6 bus** dataset (left) and the **24 bus** dataset (right)

For the **24 bus** dataset with 20 distribution networks per node, an interesting result was obtained concerning the computational effort required by the commercial solver to try to solve the global optimization problem. In fact, after a certain computational time, the solvers stopped because the commercial solver was running out of memory. Therefore, it was not able to solve

the optimization program in this case. This proves our intuition about the computational effort needed by commercial solvers to handle these problems for large-scale power systems.

Caramanis Method

In order to compare the efficiency of the commercial solver with the Caramanis algorithm, the results of experiments conducted for this method are presented in this section. The time needed by the algorithm to converge for several numbers of distribution networks per node for two different sets of data is given in figures 5.14.

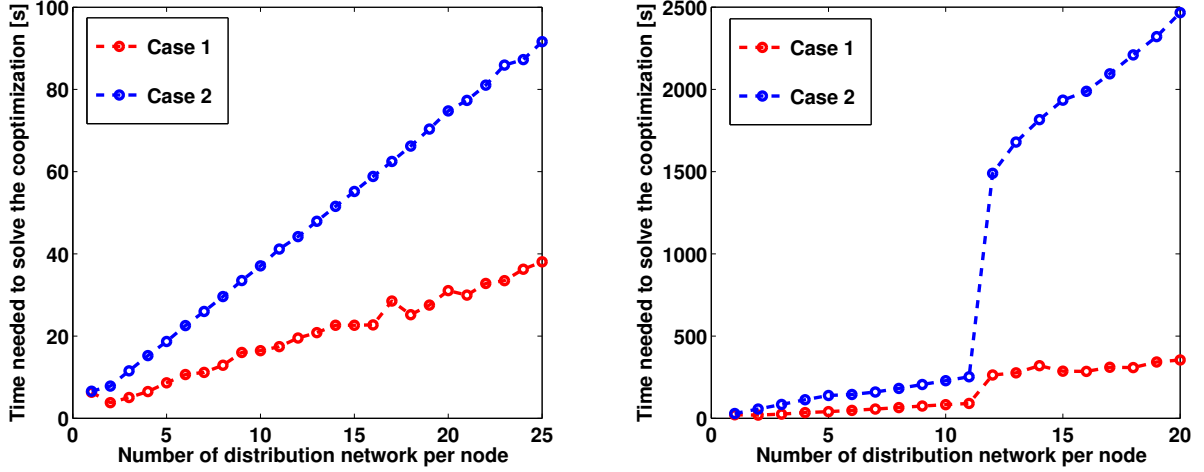


Figure 5.14: Graphical representation of the time needed by the Caramanis method to converge for several numbers of distribution networks per transmission node using the **6 bus** dataset (left) and the **24 bus** dataset (right)

An interesting result can be observed from figures 5.14. As this observation was inferred from the graphics presented in the previous section, the first formulation of the linking constraint seems to require less computational time than the other formulation. However, the number of iterations required by each formulation does not significantly differ as this is shown in figures 5.15.

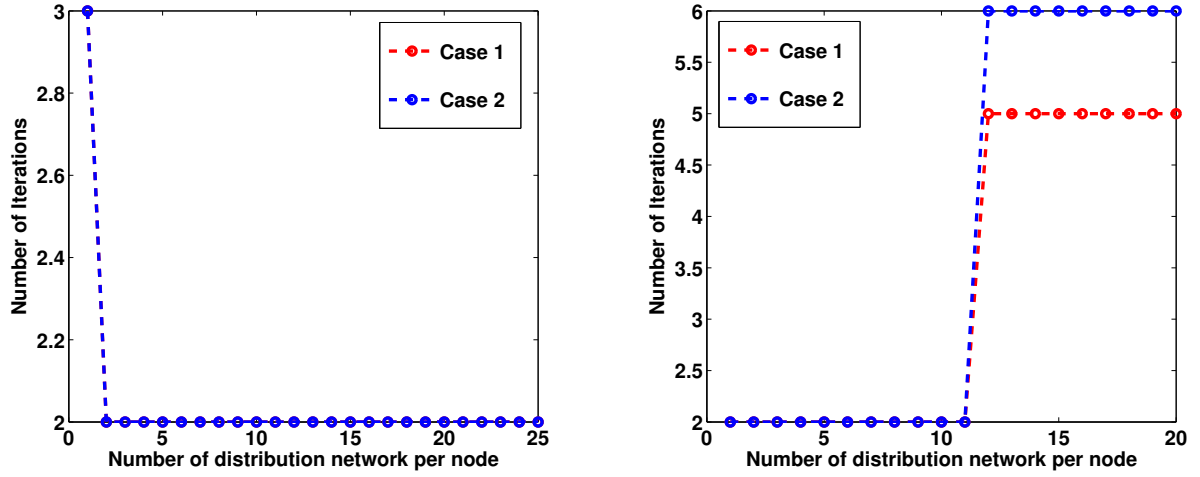


Figure 5.15: Graphical representation of the evolution of the number of iterations required by the Caramanis method to converge using the **6 bus** dataset (left) and the **24 bus** dataset (right)

In figures 5.15, a number of cases is observed to take 5 iterations to converge. In fact, these cases are stuck after 5 iterations to the same gap value before reaching the wanted tolerance. The gap obtained is given in figures 5.16. Even if the gap obtained after convergence of the algorithm is not the required one, it is still below the wanted minimal level of performance required by the algorithm.

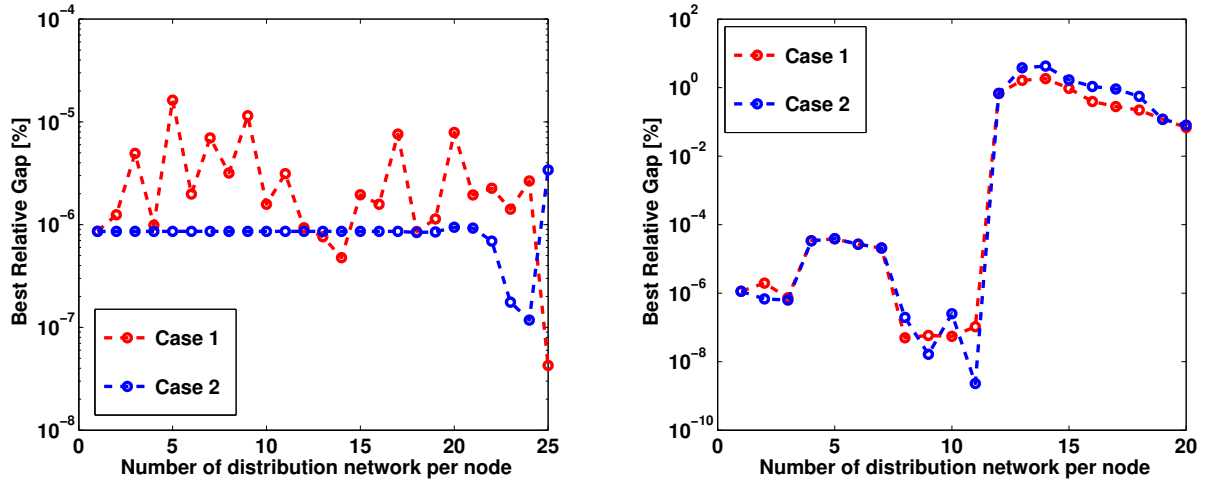


Figure 5.16: Graphical representation of the evolution of the obtained gap after convergence of the algorithm for different numbers of distribution networks per node using the **6 bus** dataset (left) and the **24 bus** dataset (right)

ADMM Algorithm

The performance of the ADMM algorithm was also experimented on different test cases. The solving time required by the ADMM method to converge is given in figures 5.17.

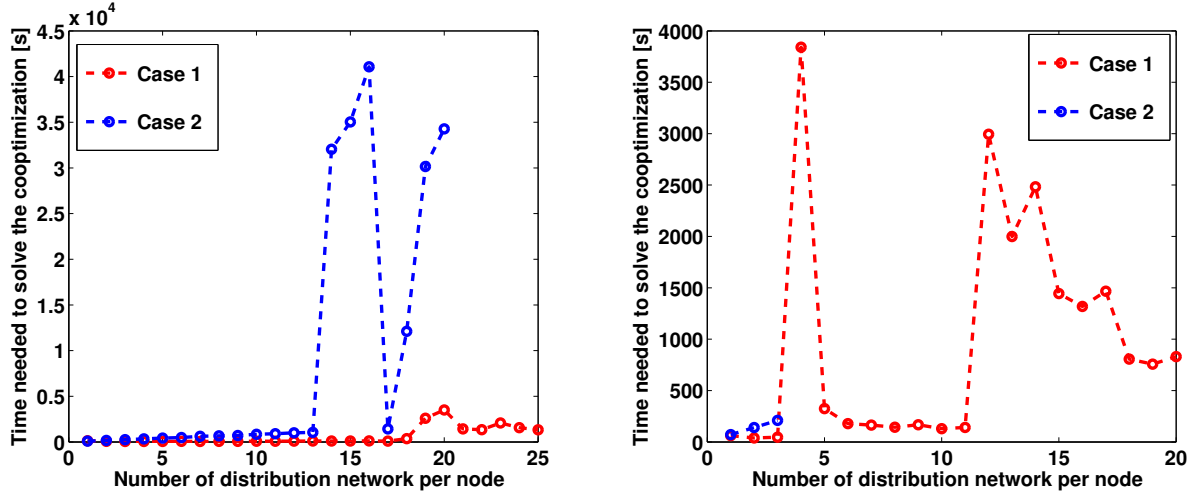


Figure 5.17: Graphical representation of the evolution of the time required by the ADMM algorithm to converge for different number of distribution networks using the **6 bus** dataset (left) and the **24 bus** dataset (right)

Results concerning the number of iterations required by the ADMM algorithm to converge are represented in figures 5.18. The resulting gap obtained after convergence of the algorithm is presented in figures 5.19.

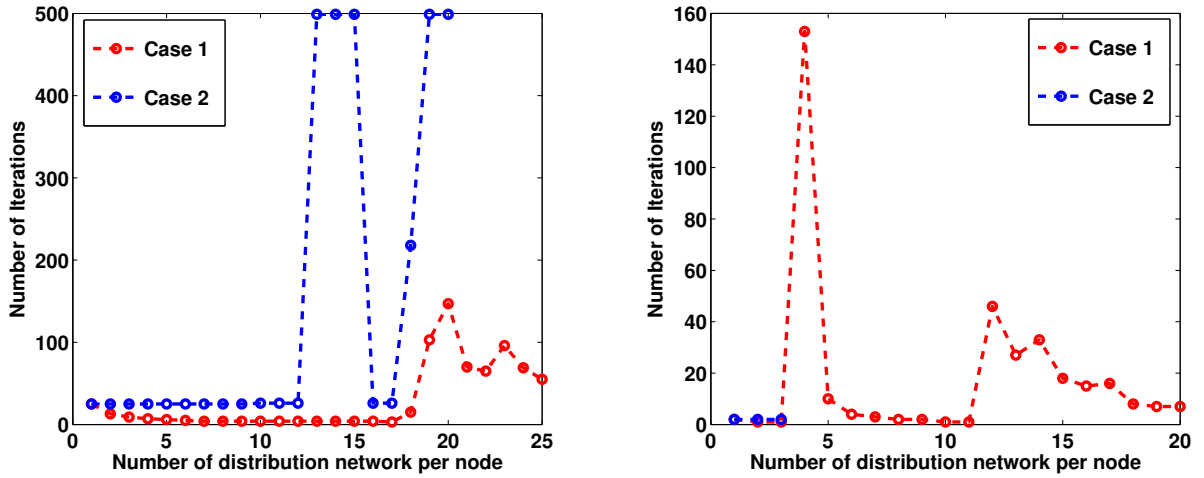


Figure 5.18: Graphical representation of the evolution of the number of iterations required by the ADMM algorithm to converge for different number of distribution networks using the **6 bus** dataset (left) and the **24 bus** dataset (right)

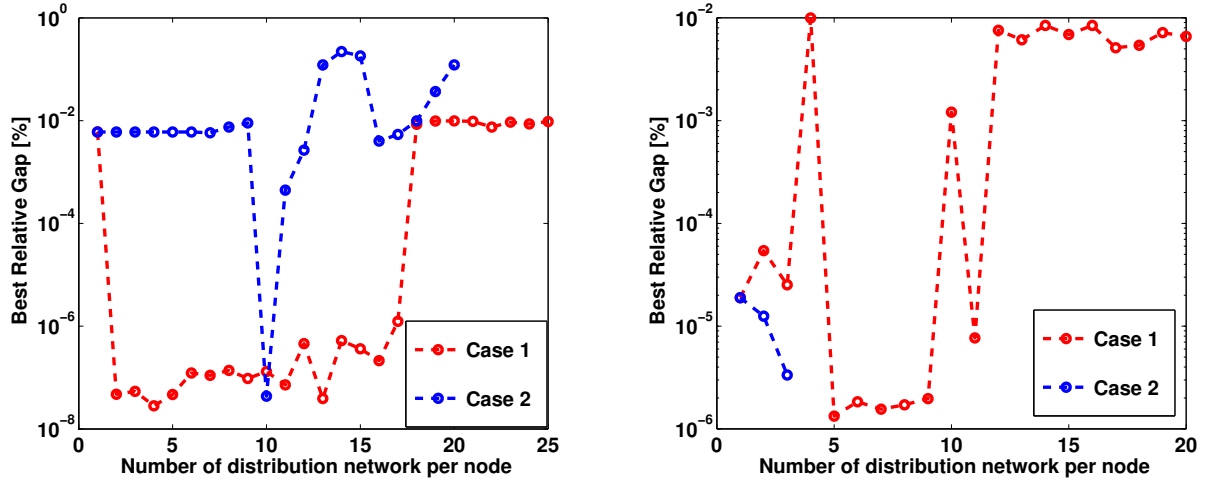


Figure 5.19: Graphical representation of the evolution of the resulting gap obtained after convergence of the ADMM method for different number of distribution networks using the **6 bus** dataset (left) and the **24 bus** dataset (right)

As this can be observed in figures 5.17, the time required by the second formulation to converge is huge for a large number of attached distribution networks compared to the first formulation. Therefore, tests for this method were only realized for a small number of cases for the second formulation of the global optimization program.

Level Method

The results obtained by realizing some tests with the Level method are presented in this section. However, due to the sensitivity of the CPLEX® solver on SOCP programs, the convergence of the Level method can be disturbed as explained previously. Therefore, results for this algorithm are only presented for the **6 bus** dataset with the first formulation of the linking constraints. The solving time and the number of iterations required by the Level algorithm to converge are given in figures 5.20.

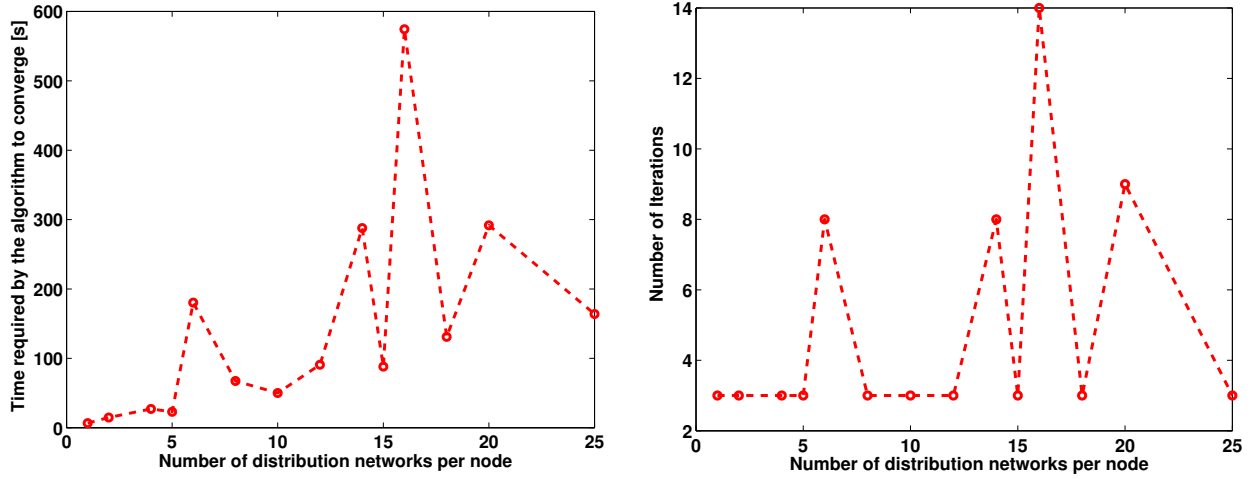


Figure 5.20: Graphical representation of the evolution of the time (left) and the number of iterations (right) required by the Level algorithm to converge for different number of distribution networks using the **6Bus** dataset and the first formulation of the linking constraint

5.3 Discussion

After observing all these results obtained on two different data sets and for two formulations of the linking constraint, some conclusions can be drawn.

First of all, as observed previously, the assumption we have made about the huge computational effort needed by a commercial solver to handle the global optimization problem for a growing number of distribution networks attached to the transmission network is verified by our tests results. Indeed, for one test case, the commercial solver was not able to solve the posed problem. This conclusion proves the necessity to create distributed algorithms to address this new type of problem.

Moreover, as seen previously in the results of every algorithm, the time required for solving the first problem formulation is smaller than for the second formulation for every algorithm apart from the Level method. This may be the consequence of several observations. First, the global optimization problem to solve is smaller than for the second formulation. Indeed, the number of variables and constraints to handle is increased by the second formulation of the problem. Furthermore, for the first formulation, all the distribution networks attached to the same node are considered as a whole in the distribution subproblems compared to the second formulation where each distribution network is considered as a subproblem to solve.

The efficiency of distributed algorithms and commercial solver on the global optimization program seems to be comparable in what concerns the 6 bus dataset. This result was expected due to the small size of these tests cases. However, the ADMM and the Caramanis method seem to be in general more efficient than the commercial solver on the 24 bus dataset. This is actually a really promising result for future research on the global optimization program.

Finally, the number of iterations required by each algorithm to converge is quite small. Therefore, it is a reasonable to choose to solve the recovery problems and the optimization programs allowing to compute a lower bound on the problem at each iteration of each algorithm.

Indeed, this choice will not slow down in the algorithms in a significant measure.

5.4 Further details

Several features of unit commitment problem were not taken into account for this thesis such as technical minimum output of generators, start up costs, etc.. Some other unit commitment characteristics were added to the global optimization problem in order to see their impact on the distributed algorithms presented. These new features are presented in the following section.

5.4.1 Unit Commitment Added Features

Objective

A generation unit has two types of costs. The variable cost depends only on the production decisions at each time period $p_{g,t}$. Furthermore, generators also incur fixed costs. Fixed costs are constituted of a minimum load cost (FC_g) which is induced each time a unit is running. Fixed costs contain also start-up (SU_g) and shutdown costs (SD_g).

In order to minimize the sum of operating costs for each unit over the entire operating horizon, all the aforementioned costs have to be taken into account at each time period. The objective function is therefore given by the following expression:

$$\min_{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}} \sum_{g \in \mathcal{G}, t \in \mathcal{T}} (MC_g p_{g,t} + FC_g u_{g,t} + SU_g su_{g,t} + SD_g sd_{g,t}) \quad (5.2)$$

Technical Minimum Production

While a generator is running, it has to produce at least a certain amount of power (p_g^-). The technical minimum output constraint is therefore given by the following expression:

$$p_{g,t} \geq p_g^- u_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \quad (5.3)$$

5.4.2 Computational Observations

Unfortunately, results for this new global optimization problem were not convincing. Indeed, due to the sensitivity of the commercial solver to the scale of the problem, all algorithms achieved a gap bigger than the minimum level of performance required by the unit commitment literature. Therefore, as a recommendation for future works on the subject, I will advice to change the scale of the solved problems by, for example, using non normalized MVA base power systems characteristics in order to deal with bigger numbers for the distribution networks. Another suggestion could be to use another commercial solver than CPLEX® such as MOSEK which can not be used on conic programs in AMPL. However, this solver has a really good reputation when it comes to using the interior point algorithm.

Chapter 6

Conclusion

6.1 Main Results

This thesis focused on one major idea which is the consideration of the transmission and distribution part of a network as a whole for the analysis of unit commitment and power flow features in a power system. This design was highly suggested by the new types of electrical technologies added to the distribution side of power grids such as rooftop solar panels and electrical vehicles. In the first chapter of this thesis, the optimization models governing the behaviour of the transmission and distribution parts of power networks were stated and the necessity of this new design was highlighted by means of an example. After that, a broad literature review was realized on linked subjects. On the fourth chapter, three distributed algorithms were presented. Finally, the performances of these algorithms were compared with the methods used by the commercial solver CPLEX® to solve the global optimization program. From these results, some important conclusions were drawn and can be summarized as follows:

- The co-optimization problem deserves the implementation because it leads to better results in terms of power systems operations as this can be seen from the example in section 2.5.
- The first formulation of the linking constraint seems to require less computational time compared to the second formulation of the global optimization problem to converge. Therefore, I suggest for future works to use only the first formulation of this global optimization program presented in section. 2.4.
- Distributed algorithms seem to be competitive compared to the commercial solver in terms of computational time required by the algorithms to converge.
- The computational effort required by the commercial solver to solve large global optimization programs seem to increase with the size of the power system. Moreover, a power networks of 24 nodes in the transmission network along with 20 distribution networks attached per transmission demand node is a problem impossible to compute by the commercial solver up to a relative gap of 0.01% due to memory requirements.

The three questions asked in the introduction are then answered by the previous highlighted results. In fact, to summarize, considering the transmission and distribution parts of a network as a whole may lead to superior performances in terms of power system operations. Moreover, even though, the global optimization problem requires a lot of computational effort from commercial solvers to be computed, three algorithms in this thesis were shown to be efficient while solving this optimization program.

6.2 Future Perspectives

As this thesis is focused on an emerging field of research a lot of future perspectives have still to be considered.

First of all, this thesis can be applied to more datasets. However, I suggest to use these power networks in the optimization program with characteristics not expressed using a MVA base in order to decrease bad effects coming from the sensitivity of the solver to small numbers for conic programs. In this thesis, only one type of distribution network was considered and attached several times to the transmission network. Another improvement may be to consider several types of distribution networks in the same global power system to add diversification. Moreover, this will lead to power networks closer to the reality than power networks used in this work.

Then, this thesis only presents three distributed algorithms. However, it may be interesting to use other algorithms to compute efficiently the solution of the global optimization program in a distributed way as for example the Predictor Corrector Proximal Multipliers (PCPM) algorithm [16]. On the other way, new researches could also be directed towards the improvement of the efficiency of the presented algorithms by adding for example cuts or feasible inequality or slightly changing the steps followed by the algorithm (see [56] for an example of improvements realized on the ADMM method) in order to reduce the gap obtained by these algorithms.

Furthermore, parallelism is the new dominant computing paradigm for computers. It allows to process several instructions at the same time. Therefore, parallelism can be applied to our distributed algorithms in order to speed up the computation of the solution of the global optimization program for large-scale power networks.

Additionally, more features of the unit commitment problem can be added to the global optimization program to solve in order to improve even more the accuracy of the model. Among these characteristics, reserve requirements can be taken into account, start-up and shutdown profiles, temperature and so on. Even stochastic unit commitments may be treated on such large networks.

Finally, in order to avoid problems related to the sensitivity of the CPLEX® solver, other commercial solvers could be used such as MOSEK as suggested earlier.

Bibliography

- [1] Anthony Papavasiliou. Introduction. "<http://uclengiechair.be/wp-content/uploads/2017/05/Introduction.pdf>", 2017.
- [2] Joshua Adam Taylor. *Convex optimization of power systems*. Cambridge University Press, 2015.
- [3] Anthony Papavasiliou, Shmuel S Oren, and Barry Rountree. Applying high performance computing to transmission-constrained stochastic unit commitment for renewable energy integration. *IEEE Transactions on Power Systems*, 30(3):1109–1120, 2015.
- [4] European Comission. Renewable energy: Moving towards a low carbon economy. "<https://ec.europa.eu/energy/en/topics/renewable-energy>", 2017. Accessed: 2017-05-30.
- [5] MIT Energy Initiative. Utility of the future: An mit energy initiative response to an industry in transition. "<https://energy.mit.edu/wp-content/uploads/2016/12/Utility-of-the-Future-Full-Report.pdf>". Accessed: 2017-05-27.
- [6] European technology platform for the electricity networks of the future. The smartgrids european technology platform. "<http://www.smartgrids.eu/ETPSmartGrids>", 2013. Accessed: 2017-06-07.
- [7] Hongbin Sun, Qinglai Guo, Boming Zhang, Ye Guo, Zhengshuo Li, and Jianhui Wang. Master-slave-splitting based distributed global power flow method for integrated transmission and distribution analysis. *IEEE Transactions on Smart Grid*, 6(3):1484–1492, 2015.
- [8] Michael Caramanis, Elli Ntakou, William W. Hogan, Aranya Chakraborty, and Jens Schoene. Co-optimization of power and reserves in dynamic t&d power markets with nondispatchable renewable generation and distributed energy resources. *Proceedings of the IEEE*, Vol. 104, No. 4, 2016.
- [9] Anthony Papavasiliou. Economic dispatch. "<http://uclengiechair.be/wp-content/uploads/2017/05/EconomicDispatch.pdf>", 2017.
- [10] Anthony Papavasiliou. Power systems. "<http://uclengiechair.be/wp-content/uploads/2017/05/PowerSystems.pdf>", 2017.
- [11] Masoud Farivar and Steven H. Low. Branch flow model: Relaxations and convexification—part i. *IEEE Transactions on Power Systems*, 2013.
- [12] Masoud Farivar and Steven H. Low. Branch flow model: Relaxations and convexification—part ii. *IEEE Transactions on Power Systems*, 2013.

- [13] Anthony Papavasiliou. Reviews of probability theory and convex analysis. "<http://uclengiechair.be/wp-content/uploads/2017/05/ShortReviews.pdf>", 2017.
- [14] Yurii Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer Science & Business Media, 2013.
- [15] Farid Alizadeh and Donald Goldfarb. Second-order cone programming. *Mathematical programming*, 95(1):3–51, 2003.
- [16] Na Li, Lijun Chen, and Steven H Low. Demand response in radial distribution networks: Distributed algorithm. In *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, pages 1549–1553. IEEE, 2012.
- [17] Anya Castillo, Carl Laird, César A Silva-Monroy, Jean-Paul Watson, and Richard P O'Neill. The unit commitment problem with ac optimal power flow constraints. *IEEE Transactions on Power Systems*, 31(6):4853–4866, 2016.
- [18] Anthony Papavasiliou. Unit commitment. "<http://uclengiechair.be/wp-content/uploads/2017/05/UnitCommitment.pdf>", 2017.
- [19] Jianfeng Liu, Anya Castillo, and Carl D. Laird. Global solution strategies for the network-constrained unit commitment (ncuc) problem with nonlinear ac transmission models. Submitted to Mathematical Programming Computation on Nov. 2016. http://www.optimization-online.org/DB_FILE/2016/12/5751.pdf.
- [20] Hongbin Sun and Boming Zhang. Distributed power flow calculation for whole networks including transmission and distribution. In *Transmission and Distribution Conference and Exposition, 2008. T&D. IEEE/PES*, pages 1–6. IEEE, 2008.
- [21] Ray Daniel Zimmerman, Carlos Edmundo Murillo-Sánchez, and Robert John Thomas. Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems*, 26(1):12–19, Feb. 2011.
- [22] Anthony Papavasiliou. Analysis of distribution locational marginal prices. *IEEE Transactions on Smart Grid*, 2017.
- [23] Burak Kocuk, Santanu S Dey, and X Andy Sun. Strong socp relaxations for the optimal power flow problem. *Operations Research*, 64(6):1177–1196, 2016.
- [24] Dirk Van Hertem, Jody Verboomen, Konrad Purchala, Ronnie Belmans, and WL Kling. Usefulness of dc power flow for active power flow analysis with flow controlling devices. In *AC and DC Power Transmission, 2006. ACDC 2006. The 8th IEE International Conference on*, pages 58–62. IET, 2006.
- [25] Carleton Coffrin and Pascal Van Hentenryck. A linear-programming approximation of ac power flows. *INFORMS Journal on Computing*, 26(4):718–734, 2014.
- [26] Steven H Low. Convex relaxation of optimal power flow—part i: Formulations and equivalence. *IEEE Transactions on Control of Network Systems*, 1(1):15–27, 2014.
- [27] Xiaoqing Bai, Hua Wei, Katsuki Fujisawa, and Yong Wang. Semidefinite programming for optimal power flow problems. *International Journal of Electrical Power & Energy Systems*, 30(6):383–392, 2008.

- [28] Daniel K Molzahn, Bernard C Lesieutre, and Christopher L DeMarco. Investigation of non-zero duality gap solutions to a semidefinite relaxation of the optimal power flow problem. In *System Sciences (HICSS), 2014 47th Hawaii International Conference on*, pages 2325–2334. IEEE, 2014.
- [29] Subhonmesh Bose, Steven H Low, Thanchanok Teeraratkul, and Babak Hassibi. Equivalent relaxations of optimal power flow. *IEEE Transactions on Automatic Control*, 60(3):729–742, 2015.
- [30] Daniel K Molzahn, Jesse T Holzer, Bernard C Lesieutre, and Christopher L DeMarco. Implementation of a large-scale optimal power flow solver based on semidefinite programming. *IEEE Transactions on Power Systems*, 28(4):3987–3998, 2013.
- [31] Javad Lavaei and Steven H Low. Zero duality gap in optimal power flow problem. *IEEE Transactions on Power Systems*, 27(1):92–107, 2012.
- [32] Rabih A Jabr. Radial distribution load flow using conic programming. *IEEE transactions on power systems*, 21(3):1458–1459, 2006.
- [33] Masoud Farivar, Christopher R Clarke, Steven H Low, and K Mani Chandy. Inverter var control for distribution systems with renewables. In *Smart Grid Communications (SmartGridComm), 2011 IEEE International Conference on*, pages 457–462. IEEE, 2011.
- [34] Qiuyu Peng and Steven H. Low. Distributed algorithm for optimal power flow on a radial network. In *53rd IEEE Conference on Decision and Control*, pages 167–172. IEEE, 2014.
- [35] G. Morales-España, J. M. Latorre, and A. Ramos. Tight and compact milp formulation for the thermal unit commitment problem. *IEEE Transactions on Power Systems*, 28(4):4897–4908, Nov 2013.
- [36] Germán Morales-España, Claudio Gentile, and Andres Ramos. Tight mip formulations of the power-based unit commitment problem. *OR Spectrum*, 37(4):929–950, 2015.
- [37] J. Ostrowski, M. F. Anjos, and A. Vannelli. Tight mixed integer linear programming formulations for the unit commitment problem. *IEEE Transactions on Power Systems*, 27(1):39–46, Feb 2012.
- [38] Hrvoje Pandzic, Ting Qiu, and Daniel S Kirschen. Comparison of state-of-the-art transmission constrained unit commitment formulations. In *Power and Energy Society General Meeting (PES), 2013 IEEE*, pages 1–5. IEEE, 2013.
- [39] Deepak Rajan, Samer Takriti, et al. Minimum up/down polytopes of the unit commitment problem with start-up costs. *IBM Res. Rep*, 2005.
- [40] Albert YS Lam, Baosen Zhang, and N Tse David. Distributed algorithms for optimal power flow problem. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*, pages 430–437. IEEE, 2012.
- [41] Balho H Kim and Ross Baldick. Coarse-grained distributed optimal power flow. *IEEE Transactions on Power Systems*, 12(2):932–939, 1997.
- [42] Matt Kraning, Eric Chu, Javad Lavaei, Stephen Boyd, et al. Dynamic network energy management via proximal message passing. *Foundations and Trends® in Optimization*, 1(2):73–126, 2014.

- [43] Zhengshuo Li, Jianhui Wang, Hongbin Sun, and Qinglai Guo. Transmission contingency analysis based on integrated transmission and distribution power flow in smart grid. *IEEE Transactions on Power Systems*, 30(6):3356–3367, 2015.
- [44] Claude Lemaréchal. Lagrangian relaxation. *Computational combinatorial optimization*, pages 112–156, 2001.
- [45] Anthony Papavasiliou. Duality. "<http://uclengiechair.be/wp-content/uploads/2017/05/Duality.pdf>", 2017.
- [46] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato, and Jonathan Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011.
- [47] Anthony Papavasiliou. Lagrangian relaxation: Introduction and applications. "<http://uclengiechair.be/wp-content/uploads/2017/05/LagrangeRelaxationContextApplications.pdf>", 2017.
- [48] Kim C. Border. Supergradients. "<http://people.hss.caltech.edu/~kcb/Notes/Supergrad.pdf>", 2001.
- [49] Anthony Papavasiliou. Lagrange relaxation: Decomposition algorithms. "<http://uclengiechair.be/wp-content/uploads/2017/05/LagrangeRelaxationDualDecomposition.pdf>", 2017.
- [50] Daniel Gabay and Bertrand Mercier. A dual algorithm for the solution of nonlinear variational problems via finite element approximation. *Computers & Mathematics with Applications*, 2(1):17–40, 1976.
- [51] Roland Glowinski and A Marroco. Sur l’approximation, par éléments finis d’ordre un, et la résolution, par pénalisation-dualité d’une classe de problèmes de dirichlet non linéaires. *Revue française d’automatique, informatique, recherche opérationnelle. Analyse numérique*, 9(2):41–76, 1975.
- [52] Yong Fu, Mohammad Shahidehpour, and Zuyi Li. Ac contingency dispatch based on security-constrained unit commitment. *IEEE Transactions on Power Systems*, 21(2):897–908, 2006.
- [53] Cliff Grigg, Peter Wong, Paul Albrecht, Ron Allan, Murty Bhavaraju, Roy Billinton, Quan Chen, Clement Fong, Suheil Haddad, Sastry Kuruganty, et al. The ieee reliability test system-1996. a report prepared by the reliability test system task force of the application of probability methods subcommittee. *IEEE Transactions on power systems*, 14(3):1010–1020, 1999.
- [54] P. M. Subcommittee. Ieee reliability test system. *IEEE Transactions on Power Apparatus and Systems*, PAS-98(6):2047–2054, Nov 1979.
- [55] Anjia Mao and M Reza Iravani. A trend-oriented power system security analysis method based on load profile. *IEEE Transactions on Power Systems*, 29(3):1279–1286, 2014.
- [56] Neal Parikh and Stephen Boyd. Block splitting for large-scale distributed learning. In *Neural Information Processing Systems (NIPS), Workshop on Big Learning*. Citeseer, 2011.

Appendix A

Second Formulation of the Global Optimization Problem

The second formulation of the global optimization problem is given by the following optimization program:

$$\begin{aligned}
& \min_{\substack{su_{g,t}, u_{g,t}, sd_{g,t}, p_{g,t}, \\ f_{k,t}, \theta_{n,t}, P_{i,t,n,t_n}, \\ Q_{i,t,n,t_n}, v_{i,t,n,t_n}, \\ l_{i,t,n,t_n}, \\ p_{i,t,n,t_n}, q_{i,t,n,t_n}, \\ P_{\infty}^{t,n,t_n}, D_{\infty}^{t,n,t_n}}} \\
& \text{subject to} \quad \sum_{g \in \mathcal{G}_n} p_{g,t} + \sum_{k=(\cdot,n) \in \mathcal{K}} f_{k,t} = d_{n,t} + \sum_{k=(n,\cdot) \in \mathcal{K}} f_{k,t} \quad \forall n \in \mathcal{N} : \mathcal{T}_n = \emptyset, t \in \mathcal{T} \\
& \sum_{g \in \mathcal{G}_n} p_{g,t} + \sum_{k=(\cdot,n) \in \mathcal{K}} f_{k,t} = d_{n,t} + D_{\infty}^{t,n} + \sum_{k=(n,\cdot) \in \mathcal{K}} f_{k,t} \quad \forall t \in \mathcal{T}, n \in \mathcal{N} : \mathcal{T}_n \neq \emptyset \\
& f_{k,t} = B_k(\theta_{m,t} - \theta_{n,t}) \quad \forall k = (m,n) \in \mathcal{K}, t \in \mathcal{T} \\
& u_{g,t} = u_{g,t-1} + su_{g,t} - sd_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \sum_{\tau \in [t-UT_g+1, t], \tau \geq 1} su_{g,\tau} \leq u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \\
& \sum_{\tau \in [t-DT_g+1, t], \tau \geq 1} sd_{g,\tau} \leq 1 - u_{g,t} \quad \forall t \in \mathcal{T}, g \in \mathcal{G} \\
& p_{g,t} - p_{g,t-1} \leq R_g^+ \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& p_{g,t-1} - p_{g,t} \leq R_g^- \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \setminus \{1\} \\
& \theta_{hub,t} = 0 \quad \forall t \in \mathcal{T} \\
& v_{i,t,n,t_n} - 2(R_i P_{i,t,n,t_n} + X_i Q_{i,t,n,t_n}) + l_{i,t,n,t_n}(R_i^2 + X_i^2) = v_{A_i,t,n,t_n}, \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& 4(P_{i,t,n,t_n})^2 + 4(Q_{i,t,n,t_n})^2 + (l_{i,t,n,t_n} - v_{i,t,n,t_n})^2 \leq (l_{i,t,n,t_n} + v_{i,t,n,t_n})^2, \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& - \sum_{c_0 \in \mathcal{C}_0} (P_{c_0,t,n,t_n} - l_{c_0,t,n,t_n} R_{c_0}) - p_{0,t,n,t_n} - P_{\infty}^{t,n,t_n} + dp_{0,t,n,t_n} + G_0 v_{0,t,n,t_n} = 0 \\
& \quad \forall t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n \\
& P_{i,t,n,t_n} - \sum_{c_i \in \mathcal{C}_i} (P_{c_i,t,n,t_n} - l_{c_i,t,n,t_n} R_{c_i}) - p_{i,t,n,t_n} + dp_{i,t,n,t_n} + G_i v_{i,t,n,t_n} = 0 \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& - \sum_{c_0 \in \mathcal{C}_0} (Q_{c_0,t,n,t_n} - l_{c_0,t,n,t_n} X_{c_0}) - q_{0,t,n,t_n} + dq_{0,t,n,t_n} + B_0 v_{0,t,n,t_n} = 0 \\
& \quad \forall t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& Q_{i,t,n,t_n} - \sum_{c_i \in \mathcal{C}_i} (Q_{c_i,t,n,t_n} - l_{c_i,t,n,t_n} X_{c_i}) - q_{i,t,n,t_n} + dq_{i,t,n,t_n} + B_i v_{i,t,n,t_n} = 0, \\
& \quad \forall i \in \mathcal{I}, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& P_{\infty}^{t,n,t_n} - D_{\infty}^{t,n,t_n} = 0 \quad \forall t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& u_{g,t}, su_{g,t}, sd_{g,t} \in \{0, 1\} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \\
& -f_k^+ \leq f_{k,t} \leq f_k^+ \quad \forall k \in \mathcal{K}, t \in \mathcal{T} \\
& p_{g,t} \leq p_g^+ u_{g,t} \quad \forall g \in \mathcal{G}, t \in \mathcal{T} \\
& p_i^- \leq p_{i,t,n,t_n} \leq p_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& q_i^- \leq q_{i,t,n,t_n} \leq q_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset \\
& v_i^- \leq v_{i,t,n,t_n} \leq v_i^+ \quad \forall i \in \mathcal{I}_0, t \in \mathcal{T}, n \in \mathcal{N}, t_n \in \mathcal{T}_n : \mathcal{T}_n \neq \emptyset
\end{aligned}$$

Appendix B

Impact of the Penalty Parameter on Primal Variables

In this appendix, additional results about the evolution of the distortion on some primal variables due to different values of the penalty parameter ϵ are presented on figures B.1 and B.2.

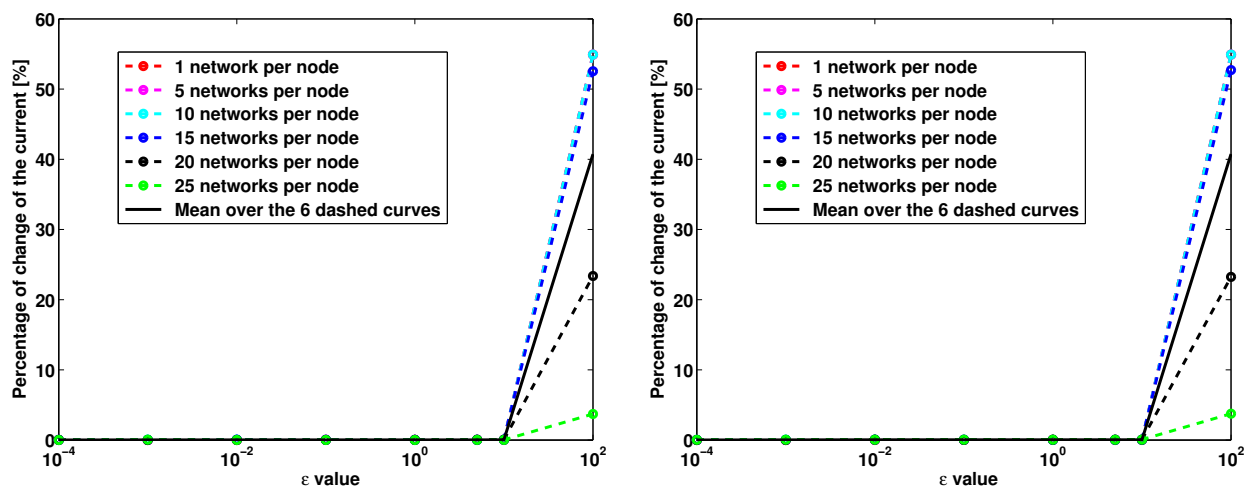


Figure B.1: Evolution of the primal variable $l_{i,t}$ distortion in percent for some values of parameter ϵ for several number of distribution networks per node. The left figure presents results obtained for the first formulation of the linking constraint, the right figure for the second formulation.

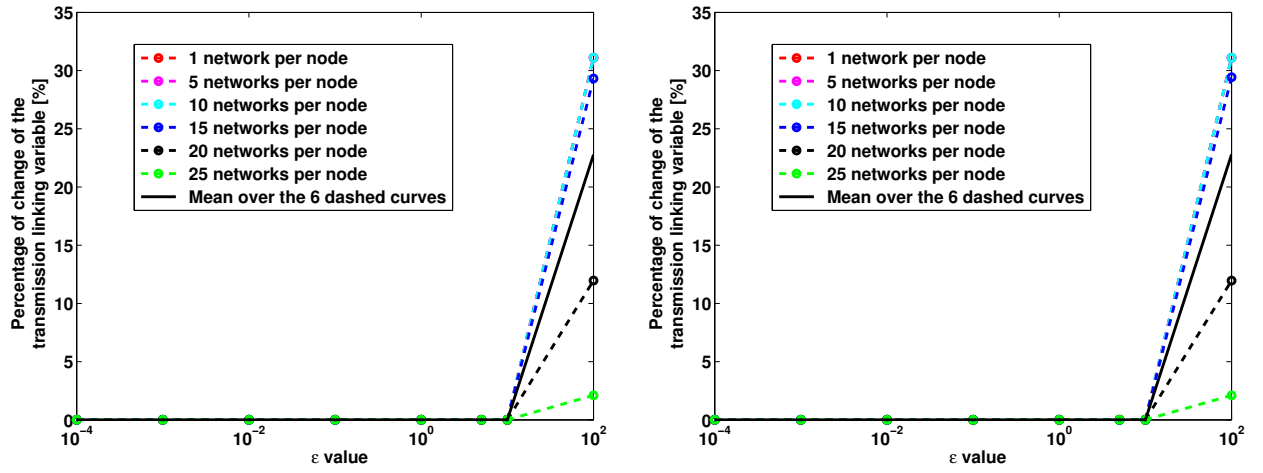


Figure B.2: Evolution of the primal variable $D_{\infty}^{t,n}$ distortion in percent for some values of parameter ϵ for several number of distribution networks per node. The left figure presents results obtained for the first formulation of the linking constraint, the right figure for the second formulation.

